



HAL
open science

Probing Cosmology with the homogeneity scale of the Universe through large scale structure surveys.

Pierros Ntelis

► **To cite this version:**

Pierros Ntelis. Probing Cosmology with the homogeneity scale of the Universe through large scale structure surveys.. Sciences of the Universe [physics]. Astroparticule and Cosmology Group, Physics Department, Paris Diderot University, 2017. English. NNT: . tel-01674537

HAL Id: tel-01674537

<https://hal.science/tel-01674537>

Submitted on 3 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Thèse de doctorat
de l'Université Sorbonne Paris Cité
Préparée à l'Université Paris Diderot
Ecole doctorale STEP'UP No 560
Laboratoire Astroparticule et Cosmologie / Equipe de recherche Cosmologie

Probing Cosmology with the homogeneity scale of the universe through large scale structure surveys

Par [Pierros NTELIS](#)

Thèse de doctorat de Physique

Dirigée par [Dr. Jean-Christophe HAMILTON](#)

Présentée et soutenue publiquement à Paris le

28 Septembre 2017

Devant un jury composé par :

Directeur de thèse :

Rapporteurs :

Examineurs :

Dr. Jean-Christophe HAMILTON	- APC Université Paris Diderot
Dr. Stephanie Escoffier	- CPPM Marseille
Prof. Jean-Paul Kneib	- EPFL Lausanne
Prof. Marc Lashière-Rey	- APC Université Paris Diderot
Dr. Anne Ealet	- CPPM Marseille
Prof. Stavros Katsanevas	- APC Université Paris Diderot

Declaration of Authorship

I, Pierros NTELIS, declare that this thesis titled, 'Probing Cosmology with the homogeneity scale of the universe through large scale structure surveys' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

“ Ἐν οἶδα, οὐδέν οἶδα ”

Σωκράτης 470/469 - 399 BC

This greek phrase translates to "one think I know, that I know nothing". In free translation it means that the amount of knowledge comprehended by an individual is insignificant to the amount existing out there. Some believe that this phrase was attributed to Σωκράτη by the skeptic and philoshopher Αρκεσίλαο the 3rd century BC [1] (Βικιπαιδεία).

PARIS DENIS DIDEROT UNIVERSITY

Abstract

Astroparticle and Cosmology Laboratory
Physics Department

Doctor of Philosophy

**Probing Cosmology with the homogeneity scale of the universe through large scale
structure surveys**

by Pierros NTELIS

In english ...

This thesis exposes my contribution to the measurement of homogeneity scale using galaxies, with the cosmological interpretation of results.

In physics, any model is characterized by a set of principles. Most models in cosmology are based on the Cosmological Principle, which states that the universe is statistically homogeneous and isotropic on a large scales. Today, this principle is considered to be true since it is respected by those cosmological models that accurately describe the observations. However, while the isotropy of the universe is now confirmed by many experiments, it is not the case for the homogeneity. To study cosmic homogeneity, we propose to not only test a model but to test directly one of the postulates of modern cosmology.

Since 1998 the measurements of cosmic distances using type Ia supernovae, we know that the universe is now in a phase of accelerated expansion. This phenomenon can be explained by the addition of an unknown energy component, which is called dark energy. Since dark energy is responsible for the expansion of the universe, we can study this mysterious fluid by measuring the rate of expansion of the universe. The universe has imprinted in its matter distribution a standard ruler, the Baryon Acoustic Oscillation (BAO) scale. By measuring this scale at different times during the evolution of our universe, it is then possible to measure the rate of expansion of the universe and thus characterize this dark energy. Alternatively, we can use the homogeneity scale to study this dark energy.

Studying the homogeneity and the BAO scale requires the statistical study of the matter distribution of the universe at large scales, superior to tens of Megaparsecs. Galaxies and quasars are formed in the vast overdensities of matter and they are very luminous: these sources trace the distribution of matter. By measuring the emission spectra of these sources using large spectroscopic surveys, such as BOSS and eBOSS, we can measure their positions. It is thus possible to reconstruct the distribution of matter in 3 dimensions in gigantic volumes. We can then extract various statistical observables to measure the BAO scale and the scale of homogeneity of the universe. Using Data Release 12 CMASS galaxy catalogs, we obtained precision on the homogeneity scale reduced by 5 times compared to WiggleZ measurement.

At large scales, the universe is remarkably well described in linear order by the Λ CDM-model, the standard model of cosmology. In general, it is not necessary to take into account the nonlinear effects which complicate the model at small scales. On the other hand, at large scales, the measurement of our observables becomes very sensitive to the systematic effects. This is particularly true for the analysis of cosmic homogeneity, which requires an observational method so as not to bias the measurement.

In order to study the homogeneity principle in a model independent way, we explore a new way to infer distances using cosmic clocks and type Ia SuperNovae. This establishes the Cosmological Principle using only a small number of a priori assumption, i.e. the theory of General Relativity and astrophysical assumptions that are independent from Friedmann Universes and in extend the homogeneity assumption.

This manuscript is as follows. After a short presentation of the knowledge in cosmology necessary for the understanding of this manuscript, presented in Chapter 1, Chapter 2 will deal with the challenges of the Cosmological Principle as well as how to overcome those. In Chapter 3, we will discuss the technical characteristics of the large scale structure surveys, in particular focusing on BOSS and eBOSS galaxy surveys. Chapter 4 presents the detailed analysis of the measurement of cosmic homogeneity and the various systematic effects likely to impact our observables. Chapter 5 will discuss how to use the cosmic homogeneity as a standard ruler to constrain dark energy models from current and future surveys. In Chapter 6, we present our results on the determination of the homogeneity of the universe model independently. Finally in chapter 7, we draw our conclusions.

En français ...

Cette thèse présente ma contribution à la mesure de l'échelle d'homogénéité à l'aide de galaxies, avec l'interprétation cosmologique des résultats.

En physique, tout modèle est constitué par un ensemble de principes. La plupart des modèles de cosmologie sont basés sur le principe cosmologique, qui indique que l'univers est statistiquement homogène et isotrope à grande échelle. Aujourd'hui, ce principe est considéré comme vrai car il est respecté par ces modèles cosmologiques qui décrivent avec précision les observations. Cependant, l'isotropie de l'univers est maintenant confirmée par de nombreuses expériences, mais ce n'est pas le cas pour l'homogénéité. Pour étudier l'homogénéité cosmique, nous proposons un postulat d'homogénéité cosmique.

Depuis 1998, les mesures des distances cosmiques à l'aide de supernovae de type Ia, nous savons que l'univers est maintenant en phase d'expansion accélérée. Ce phénomène s'explique par l'ajout d'une composante énergétique inconnue, appelée énergie sombre. Puisque l'énergie noire est responsable de l'expansion de l'univers, nous pouvons étudier ce fluide mystérieux en mesurant le taux d'expansion de l'univers. L'échelle d'oscillation acoustique Baryon (BAO). En mesurant cette échelle à différents moments de la vie de notre univers, il est alors possible de mesurer le taux d'expansion de l'univers et donc de caractériser cette énergie sombre. Alternativement, nous pouvons utiliser l'échelle d'homogénéité pour étudier cette énergie sombre.

L'étude l'échelle de l'homogénéité et l'échelle BAO reclament l'étude statistique de la regroupement de la matière de l'univers à grandes échelles, supérieure à plusieurs dizaines de Megaparsecs. Les galaxies et les quasars sont formés dans les vastes surdensités de la matière et ils sont très lumineuses: ces sources tracent la distribution de la matière. En mesurant les spectres d'émission de ces sources en utilisant de larges études spectroscopiques, telles que BOSS et eBOSS, nous pouvons mesurer leurs positions. Il est possible de reconstruire la distribution de la matière en trois dimensions en volumes gigantesques. Nous pouvons ensuite extraire divers observables statistiques pour mesurer l'échelle BAO et l'échelle d'homogénéité de l'univers. En utilisant les catalogues de diffusion de données 12 de la version 12 de données, nous avons obtenu une précision sur l'échelle d'homogénéité réduite de 5 par rapport la mesure de WiggleZ.

À grande échelle, l'univers est remarquablement bien décrit en ordre linéaire selon le modèle Λ CDM, le modèle standard de la cosmologie. En général, il n'est pas nécessaire de prendre en compte les effets non linéaires qui compliquent le modèle à petites échelles. D'autre part, à grande échelle, la mesure de nos observables devient très sensible aux effets systématiques. Ceci est particulièrement vrai pour l'analyse de l'homogénéité cosmique, qui nécessite une méthode d'observation.

Afin d'étudier le principe d'homogénéité d'une manière indépendante du modèle, nous explorons une nouvelle façon d'inférer des distances en utilisant des horloges cosmiques et SuperNovae de type Ia. C'est la théorie la plus couramment utilisée dans le domaine des hypothèses astrophysiques.

Ce manuscrit est le suivant. Après une courte présentation des connaissances en cosmologie nécessaires à la compréhension de ce manuscrit, présentée au chapitre 1, le chapitre 2 traitera des défis du principe cosmologique ainsi que de la manière de les surmonter. Au chapitre 3, nous discuterons des caractéristiques techniques de la structure à grande échelle dans les enquêtes de BOSS et de eBOSS. Le chapitre 4 présente l'analyse détaillée de la mesure de l'homogénéité cosmique et des différents effets systématiques. Le chapitre 5 traitera de l'utilisation de l'homogénéité cosmique comme règle standard pour limiter les modèles d'énergie sombre des enquêtes actuelles et futures. Dans le chapitre 6, nous présentons nos résultats sur la détermination de l'homogénéité du modèle universel de manière indépendante. Enfin, au chapitre 7, nous tirons nos conclusions.

Keywords: cosmology, galaxies, clustering, surveys, redshift, Λ CDM, matter, energy, fractals, statistics

Acknowledgements

Kalimera!

First of all I would like to thank my supervisor Dr Jean Christophe Hamilton for making it possible for me to complete this PhD thesis in such an outstanding institute as APC, and also for suggesting a very promising topic to work on. He is the best supervisor I could ever had!

Furthermore, I would like to thank the two rapporters Dr Stephanie Escoffier and Prof. Jean Paul Kneib that have read my thesis and bring some insightful comments that improve this manuscript to each best version. Besides, I would also like to express my sincere gratitude to the rest of my thesis committee: Dr Anne Ealet, Prof. Stavros Katsanevas and Prof. Marc Lashière-Rey, for their perceptive comments and the hard questions which urged me to widen my research.

My research in APC has greatly benefitted from fruitful discussions with several colleagues, most importantly with Jean-Marc Le Goff, James Rich Nicolas Busca, Eric Aubourg, Ken Ganga. I would like also to thank Sotiris Loucatos and Yannick Giraud-Héraud for their support throughout the satellite difficulties of this PhD! Thanks!

I would also like to thank Dr Pierre Chanial and the PhD students Pierre Laurent and Davide Poletti for helping me on the computational side of this work!

Even PhD students find free time now and then. Here, I would like to use an opportunity to thank the members of our great basketball team: Bui Van Tuan, Hoang Duc Thuong and Allesandro Traini. I would like also to Eleonora Capocasa for the amazing dinner nights! I would like also to thank Ranajoy Banerji, Camille Perbost and Avgitas Theodore for the rough times of hearing my complains! Thanks guys!

Furthermore, I would like to thank Mikhail Stolpovski and Cyrille Doux for their help in the conduction of the Cosmology Journal Club as well as for our dedicated philosophical discussions!

And finally, I would like to thank my parents Nikos and Ntina and my brother Stathis for the constant support, encouragement throughout and listening all these years!

Thank you very much, everyone!

Pierros Ntelis

Kalimera!

Tout d'abord, j'aimerais remercier mon superviseur, Dr Jean Christophe Hamilton, pour avoir permis de compléter cette thèse de doctorat dans une institution aussi remarquable que l'APC et aussi de proposer un sujet très prometteur. Il est le meilleur superviseur que j'aurais pu!

En outre, j'aimerais remercier les deux rapporteurs, Dr Stephanie Escoffier et Prof. Jean Paul Kneib, qui ont lu ma thèse et apporté des commentaires perspicaces qui améliorent ce manuscrit à chaque meilleure version. En outre, j'aimerais également exprimer ma sincère gratitude au reste de mon comité de thèse: Dr Anne Ealet, Prof. Stavros Katsanevas et Prof. Marc Lashéze-Rey, pour leurs commentaires perceptifs et les questions difficiles qui m'ont exhorté à élargir mes recherches.

Ma recherche dans APC a grandement profité à des discussions fructueuses avec plusieurs collègues, surtout avec Jean-Marc Le Goff, James Rich Nicolas Busca, Eric Aubourg, Ken Ganga. Je tiens également à remercier Stavros Katsanevas, Sotiris Loucatos et Yannick Giraud-Héraud pour leur soutien à travers les difficultés satellite de ce doctorat! Merci!

Je tiens également à remercier le Dr Pierre Chanial et les étudiants de doctorat Pierre Laurent et Davide Poletti pour m'avoir aidé du côté informatique de ce travail!

Même les étudiants en doctorat trouvent du temps libre de temps en temps. Bui Van Tuan, Hoang Duc Thuong et Allesandro Traini. J'aimerais aussi Eleonora Capocasa pour les dîners étonnants! Je tiens à remercier Ranajoy Banerji, Camille Perbost et Avgitas Theodore pour les moments difficiles d'entendre mes plaintes! Merci les gars!

Je tiens à remercier Mikhail Stolpovski et Cyrille Doux pour l'organisation du Cosmology Journal Club ainsi que pour nos discussions philosophiques dédiées!

Et enfin, j'aimerais remercier mes parents Nikos et Ntina et mon frère Stathis pour le soutien constant, l'encouragement et l'écoute de toutes ces années!

Merci beaucoup, tout le monde!

Pierros Ntelis

Contents

Declaration of Authorship	i
Abstract	iii
Acknowledgements	vii
Contents	ix
List of Figures	xiii
List of Tables	xviii
Abbreviations	xix
Physical Constants	xx
Planck Units	xxi
Cosmo Units	xxii
1 Concordance Cosmology	1
1.1 Prelude	1
1.2 Initialising Λ CDM- <i>model</i>	2
1.2.1 A Smooth Universe	2
1.2.2 The cosmic fluid	4
1.2.2.1 Radiation: Photons and Neutrinos?	6
1.2.2.2 Baryonic matter	7
1.2.2.3 Dark matter	8
1.2.3 Fluctuating from equilibrium	9
1.2.4 Expanding universe	11
1.2.5 Evolving Fluctuations	12
1.2.6 Accelerating universe with standard rulers	13
1.2.7 Initial conditions and the Inflation Paradigm	17
1.2.8 Cosmic Microwave Background	20
1.3 Gravitational Waves	23
2 Cosmological Principle	24
2.1 Observational and Historical views	24
2.2 Mathematical definition	27
2.3 Implications of Smoothing	29

2.4	Fractal Universes?	32
2.5	Complications of Celestial 3Dimensional Distributions	33
2.5.1	Biased Tracers	34
2.5.2	Redshift Space Distortions	34
2.6	Current Unanswered questions of Λ CDM-model	36
2.7	Approaching the Challenges	37
3	Large Scale Structure Surveys	40
3.1	Introduction	40
3.2	Photometry vs Spectroscopy	42
3.3	The Sloan Digital Sky Survey	42
3.3.1	(e)BOSS: Main Scientific Goals	44
3.3.2	The Telescope	45
3.3.3	The Camera	45
3.3.4	The spectrographs	46
3.3.5	Wide-field Imaging	47
3.3.6	Galaxy target selection	48
3.3.7	Observations	49
3.3.8	Data reduction pipeline	51
3.4	Sample Selection Strategy	52
3.4.1	Homogeneity study Sample	53
3.4.2	Galaxy weighting scheme	55
3.4.3	Mock Galaxy Catalogues	56
4	Homogeneity Measurement	58
4.1	Homogeneity Observables	58
4.1.1	Counts-in-spheres and Fractal Correlation Dimension	58
4.1.2	$\xi(r)$ - $\mathcal{N}(< r)$ relation	59
4.1.3	Homogeneity scale definitions	60
4.1.4	Homogeneity Estimators	61
4.1.4.1	2ptCF estimators	62
4.1.4.2	$\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$ estimators	62
4.1.5	Theoretical modelling	63
4.1.5.1	Theory for $\xi(r)$, $\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$ observables	64
4.1.5.2	Bias and RSD corrections	64
4.1.6	Core Analysis Algorithm	66
4.2	Results	67
4.2.1	Homogeneity scale for the galaxy distribution	67
4.2.2	Determining the bias and the velocity dispersion	68
4.2.3	Homogeneity scale for the matter distribution	69
4.2.4	Consistency with Λ CDM	71
4.2.5	Constraints on fractal correlation dimension	72
4.3	Selection Criteria and Systematics	73
4.3.1	Scale-Range cuts	73
4.3.2	Covariance and Precision Matrices	74
4.3.3	Spline error robustness	75
4.3.4	Normalization Issue between <i>str cor</i> estimators	75
4.3.5	Fractal correlation dimension bias	77
4.3.6	RSD model Systematic	79
4.3.7	Boosting error Systematics	79
4.3.8	Test of spline fit on QPM mock catalogs	84
4.3.9	Qualitative tests between estimators	84
4.3.10	RSD- \mathcal{R}_H correlations	86

4.3.11	Weight Scheme Systematic tests	88
4.4	Bias Gain	89
4.5	Homogeneity scale at 0.1%	91
4.6	Alternative Definition of \mathcal{R}_H results	92
4.7	3D Homogeneity or 2D Spatial Isotropy?	95
4.8	Conclusion and Discussion	96
5	Cosmic Homogeneity as a Standard Ruler	128
5.1	Does the \mathcal{R}_H bring any new extra cosmological information?	128
5.1.1	Determination of the BAO peak Position	129
5.1.2	Determination of the Homogeneity Scale	129
5.1.3	Blind Cross-Correlation of $\mathcal{R}_H - \mathcal{R}_{BAO}$ plane	130
5.2	Fisher Matrix \mathcal{R}_H vs r_s	131
5.3	MCMC on \mathcal{R}_H	133
5.4	Modelling the Future Surveys	138
5.5	Modelling the sensitivity of \mathcal{R}_H	139
5.5.1	Simulating the Scaling Laws of Precision	141
5.5.2	Reshuffle Method	141
5.5.3	Results of scaling laws	142
5.5.4	Results with the Global Fitting Formula	143
5.6	Conclusion	147
6	Cosmic Homogeneity with Cosmic Clocks and type Ia SN	148
6.1	Introduction	148
6.2	Methodology	149
6.3	3D Space reconstruction	150
6.4	Results	151
6.4.1	Hubble Rate calibration with Cosmic Chronometers	151
6.4.2	Luminosity Distance Calibration using SN	151
6.4.3	Model independent Cosmological distances	152
6.4.4	Fractal Correlation Dimension with Cosmic Chronometers and SN	152
6.5	Conclusion	154
7	Conclusion	156
A	Cosmo-Tips	160
A.1	Cosmography	160
A.1.1	Luminosity Distance revisited	161
A.2	Topological Restrictions	161
A.2.1	Spatial stationary comoving observer	162
A.2.2	Spatial non-accelerating comoving observer	162
A.2.3	Isotropical local expansion	162
A.2.4	Selection of the 3D spatial metric	163
A.3	General n-point Correlation function	164
		164
		164
		164
A.4	CMASS Effective Volume Calculation	165
A.5	Cosmic Variance vs Shot Noise Domination	165
A.6	Systematic Error from the underlying Modelling	165

A.7 $\mathcal{D}_2(r)$ - Error Propagation from s.s.s. to l.s.s.	166
A.8 <i>bias</i> , a gain for \mathcal{D}_2^M	166
A.9 Alternative homogeneity observables	168
A.10 Gravitational Waves	168
A.11 Fisherology	170
A.12 Numerical Error propagation on interpolated point	171
A.13 Why locally we observe Minkowski space while globally we feel expansion?	172
A.14 Can we use ratios as an \mathcal{N} -Estimator?: No	173
A.14.1 Hubble Distance Measure	173
Bibliography	175

List of Figures

1.1	2D Representation of possible variety of the curvature of our 3 dimensional spatial space.	3
1.2	The highly resolved Arp-734 binary galaxy observed by HST[12]. The name of this binary is <i>Chorodidascalous Andromedae</i> , or "Χοροδιδάσκαλος Ανδρομέδα" which means in greek translation "the <i>Dance-teacher Human-protector</i> ", describing the gravitational interaction (dance) between the two galaxies.	7
1.3	Composite image of the Bullet (star) Cluster observed by the Chandra X-ray Observatory [17]. We can see the X-ray(blue) and γ -ray (pink) radiation emitted by the collision of star cluster with another start cluster providing us with direct evidence of Dark Matter through gravitational lensing.	8
1.4	Schematic representation of the linear, coupled Boltzmann-Einstein field equation describing the interplay of physics at large scales, inspired by Fig 4.1 of Dodelson [23].	9
1.5	Measurement of the matter power spectrum observed by combination of different surveys[27].	13
1.6	Scale factor as a function of time, measured by a compilation of the Brightness of samples of Supernova at different redshifts. One can notice that the "always accelerating" scenario is best fitted to these data [Image taken from [30]]	14
1.7	Illustration of the AP-test of large scale structure clustering. [See text for details]	16
1.8	Large scale structure clustering measurements, revealing the accelerating nature of our universe with a combilation of AP-tests [Image drag from [33]]	17
1.9	: Illustration of the inflationary solution to the horizon problem. The comoving physical horizon shrinks during inflationary epoch and expands during the standard Λ CDM-model evolution at least until later time at $a > 0.5$ This allow all large physical fluctuation (p, q points at the surface of last scattering which are outside the physical horizon) to be in causal contact with each other at initial times since in the past they were inside the physical horizon. [Diagram was taken by Baumann [37]	19
1.10	Uniform frequency spectrum (intensity vs frequency) fitted on the theoretical Planck Black Body curve. The errorbars are only a small fraction of the line thickness[41]. [See text for details]	21
1.11	<i>Left</i> : CMB intensity map showing the temperature fluctuations as observed by PLANCK Satellite[42]. <i>Right</i> : Measured angular power spectra of a compilation of experiments denoted with different colours [43]. [See text for details]	22
2.1	<i>Left</i> : The <i>Lianakea</i> structure in the equatorial plane, identified by a velocity Wiener filter algorithm with a compilation of (z, θ, ϕ) coordinates of galaxies[71]. Shaded contours represent density values within the equatorial slice with red at high densities and blue in voids. Individual galaxies from a redshift catalog are given as white dots. Velocity flow streams within the Laniakea basin of attraction are shown in white. The orange contour encloses the outer limits of these streams.	26
2.2	<i>Left</i> : 2D representation of homogeneous (and isotropic) galaxy distribution <i>Right</i> : 2D representation of an isotropic (but not homogeneous) galaxy distribution [See text for explanation][Credit on [99]]	28

2.3	Relative Distance module measurement as a function of redshift of a compilation of Supernova, fitting separate isotropic universe (AvERA) better than homogeneous and isotropic universe (Planck best-fit model).[See text for details]	29
2.4	An alignment of the overdensity and underdensity spots in the cosmic web as discovered by a compilation of supercluster catalogues. [Image taken from [102]]	30
2.5	The first four iterations of the Kock snowflake.[Image taken from [113]]	33
2.6	Schematic illustration of the <i>Redshift Space Distortion</i> effect. In the top the figures represent the actual position of the celestial objects and their velocities. On the left the motions are due to gravitational potential of the overdensity and on the right due to gravitational interaction among the objects. Bottom figures represent the apparent position of the objects and the apparent shape of the overdensity. [See text for details]	35
3.1	SDSS Telescope located in the Apache Point Observatory, in New Mexico, U.S.A.	46
3.2	Photometric Camera Layout and its design overplotted. Optical layout of the focal plane of the SDSS camera. Field 22 (top and bottom) are focus CCDs; Fields 16-21 are astrometric chips, and 1-15 are the photometric array. The TDI drift scan direction is coming from the nose of Y-arrow to its tail, so a star-light traverses this array from right to left. [Image taken from [135] and modified to include the design]	47
3.3	Quantum Efficiency of the telescope camera for the u,g,r,i,z photometric bands. The expected throughput of the optics is included; the lower of each pair of curves includes the expected atmospheric extinction. [Image taken from [135]]	47
3.4	Optical layout of the SDSS spectrographs. Light enters each spectrograph through 320 fibers (180 μm diameter) terminating at a curved slit plate mounted inside the slithead. The slit plate positions the fiber ends on a radius concentric with the spherical collimating mirror (B). The 45 deg dichroic beamsplitter reflects the blue portion of the bandpass ($\lambda < 605 \text{ nm}$) and transmits the red wavelengths ($\lambda > 605 \text{ nm}$). Immediately after the beamsplitter in each channel is a grism (D,E), dispersing the light which exits the gratings and enters all-refractive, eight-element (F,G). Each camera contains a single $2^{11} \times 2^{11}$ CCD with $24\mu\text{m}$ pixels (H). [Image taken from Smee et al. [136]]	48
3.5	Target Selection algorithm, (see details in the text)	49
3.6	BOSS Aluminum plate. Bundles are separated by black bounded edges, and holes are marked in blue to reduce contamination between nearby emission line galaxies or quasars. Holes for guide star fibers are marked in black and denoted by the corresponding number ranging from 1 to 16.	50
3.7	Sample of SDSS spectra. (Top) 2 SDSS-III targeted galaxies spectra (black) and the noise at each pixel (red) with the best redshift measurement (blue). (2nd row to Bottom) are 6 SDSS-IV targeted objects (black) with their identified emission lines (blue).[Image extracted by [103] and modified]	53
3.8	3D representation of the CMASS sample. Different colors represent different redshift slices applied in homogeneity galaxy analysis[144].	54
3.9	R.A. and Dec profiles (top) and redshift distribution (bottom) of the DR12 sample for the NGC and SGC. The vertical dashed lines define the redshift-cut used in the analysis [144].	55
3.10	Qualitative comparison mock catalogues QPM and LPT with N-body simulations (TPM). The advantage of QPM against LPT is clear on the non linear regime.[image taken from White et al. [150]]	57
4.1	Test of necessity of RSD modelling: We show the percentage ratio of the predicted homogeneity scale defined in Eq. 4.7 with and without accounting redshift space distortions according to Eq.4.17. The effect ranges from 15 to 25% and varies according to the bias parameter and is therefore not negligible.	65
4.2	Core analysis pipeline procedure. See text for detailed explanation	67

4.3	Scaled counts-in-spheres, $\mathcal{N}(< r)$, (left) and fractal correlation dimension, $\mathcal{D}_2(r)$, (right). The black points with error bars are the result for the galaxy distribution in the NGC in the $0.430 \leq z \leq 0.484$ redshift bin. The blues lines are the best spline fit of the data. The green lines are the results for the 1000 QPM mock catalogs. The red continuous line is the Λ CDM prediction for $b = 1.95$ and $\sigma_p = 252$ km/s. The vertical dash line shows the BAO feature on the fractal correlation dimension computation.	68
4.4	The correlation function of the CMASS galaxies in the $[0.538 - 0.592]$ redshift range. Data points have both their original error bars and the error bars enlarged to take into account the uncertainty of the RSD model on small scales (see section 4.2.2 and appendix 4.3.7). The red line is the result of the fit performed over the range $[1, 40] h^{-1}$ Mpc (up to the vertical black-dashed line).	69
4.5	Same as figure 4.3 for matter distribution in the redshift interval $0.538 \leq z \leq 0.592$	70
4.6	The homogeneity scale $R_H^{\mathcal{D}_2=2.97}(z)$ measured in the NGC (red) and in the SGC (blue) as a function of redshift. The green line is the Λ CDM model prediction. The shaded areas indicate the 1σ range for the 1000 QPM mock catalogs.	70
4.7	Top: value of the <i>dilatation</i> parameter α resulting from the fit of the NGC (red) and SGC (blue) $\mathcal{D}_2(r)$ data with the model of equation 4.23, in different redshift bins. The shaded areas correspond to the 1σ region for the 1000 QPM mock catalogs. Bottom: The corresponding χ^2	71
4.8	Fractal correlation dimension for the matter distribution, $3 - \langle \mathcal{D}_2 \rangle$, averaged over $300 h^{-1} < r < 1300 h^{-1}$ Mpc, in the different redshift bins.	72
4.9	Scaled number density of random pairs, $r^{-2}dN/dr$, versus the comoving radius of spheres, r , normalized to be unity at small r	74
4.10	Correlation matrices for our estimators $\xi(r)$, $\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$ for the total matter distribution estimators in the north galactic cap at the redshift slice $0.538 \leq z \leq 0.592$	75
4.11	Ratio of the standard deviations of the homogeneity scale of the matter distribution using 3 methods over the standard deviation of the homogeneity scale using the spline method fit of our basis analysis for the different redshift bins for the NGC(red) and SGC(blue). The dot (‘.’) denotes the method of MCMC on the data, the (‘□’) square and (‘^’) wedge denote the standard deviation of the homogeneity scale on 1000 QPM mock catalogues with the interpolation method and the spline fit respectively.	76
4.12	Same as figure 4.3 for the str-estimators in the redshift interval $0.538 \leq z \leq 0.592$	77
4.13	Reconstructed fractal correlation dimension, \mathcal{D}_2 , at the nominal value 2.995 (green dash line) in the redshift interval $z = 0.538 - 0.592$. Yellow line represent each fractal simulation. Green line is the average and error over 500 simulations for str-estimator (left) and cor-estimator (right).	78
4.14	Reconstructed average value of \mathcal{D}_2 as a function of scales, r , with the estimators $\mathcal{D}_2^{str}(r)$ (left) and $\mathcal{D}_2^{cor}(r)$ (right), for 500 simulations. The errors are obtained with the standard deviation over 500 simulations. The dash line corresponds to the initial input value of \mathcal{D}_2	78
4.15	Reconstructed fractal correlation dimension, \mathcal{D}_2 , in the redshift interval $z = 0.538 - 0.592$, averaged over $r \in [15, 1300] h^{-1}$ Mpc.	79
4.16	Measured homogeneity scale in different redshift bins, in the NGC and the SGC, with Kaiser (triangle) and full RSD model (circles).	80
4.17	Empirical 2ptCF, $\xi_{fit}(r; b, \sigma_p, \vec{p}_{cs})$, as a function of scales, r , for the best fit parameters, $(b, \sigma_p, \vec{p}_{cs}) = (b^{fit}, \sigma_p^{fit}, \vec{p}_{cs}^{fit})$ for the mean of the $(0.538 \leq z \leq 0.592)$ -slice. [See text for details]	81
4.18	χ^2 test for selection of best boosting error parametrized by e , for different fitting ranges from left to right and different redshift slices(different colors). See text for details]	81

4.19	Cosmic bias parameter, b , as a function of boosting parameter e , for different fitting ranges from left to right and different redshift slices(different colors). See text for details]	82
4.20	Same as Fig. 4.19 but for the peculiar velocities, σ_p . [See text for details]	82
4.21	Boosting factor on the error in the correlation function as a function of scales, for all redshift bins (different colors).	83
4.22	$R_H(z)$ with (circle) and without (triangle) boosting the errors, for NGC (red) and SGC (blue).	83
4.23	Top: the scaled counts-in-spheres, $\mathcal{N}(< r)$, (left) and the fractal correlation dimension, $\mathcal{D}_2(r)$, (right) for matter distribution, with str(green), lau (black) and cor (blue) estimators, compared to Λ CDM model (red). Bottom: the ratio to Λ CDM model for both estimator.	84
4.24	Ratio of the 1σ of the observables \mathcal{N} (left) and \mathcal{D}_2 (right) for estimators <i>lau</i> (blue) and <i>str</i> (green) against <i>cor</i> as defined in table 4.2 as a function of scales, r . We zoom in at scales $r \in [10 - 1300]h^{-1}\text{Mpc}$	85
4.25	Homogeneity scale for the galaxy distribution measurement, $\mathcal{R}_H^{(G)}$ or matter distribution, $\mathcal{R}_H^{(M)}$, versus bias, b , or peculiar velocities, σ_p , measurements for 1000 QPM mock catalogs on the North Galactic Cap for the different redshift bins color-coded. The contours show the 1σ and 2σ confidence regions. The normalized histograms with the mean and 1σ values are displayed for the two measurements. The correlation coefficient, r , is also displayed.	86
4.26	Same as Fig. 4.25 but for the South Galactic Cap.	87
4.27	Correlation coefficient, r , as a function of redshift, z , between bias, b (left), or peculiar velocities, σ_p (right), and the homogeneity scale, \mathcal{R}_H for the North or South galactic caps for the matter (M) or galaxy (G) distribution, as measured for 1000 QPM mock catalogs.	88
4.28	Systematic measurement of the bias, b and peculiar velocities, σ_p as a function of redshift bins, z for the North galactic cap. Different colors correspond to different weighting scheme. Top figures show the measurement and bottom figures show the percentage ratio from the measurement applying all the weights. Shaded regions shows the 1σ error of the measurement applying all the weights.	89
4.29	Systematic measurement for the homogeneity scale, \mathcal{R}_H , of the galaxy (left) and matter (right) distributions as a function of redshift, z , for the North galactic cap. Different colors correspond to different combination of the weighting scheme. Top figures show the measurement and bottom figures show the percentage ratio from the measurement applying all the weights. Shaded regions shows the 1σ error of the measurement applying all the weights.	90
4.30	Same figure as Fig. 4.28 for the south galactic cap	91
4.31	Same figure as Fig. 4.29 for the south galactic cap	92
4.32	<i>Top</i> : \mathcal{D}_2 -precision <i>Bottom</i> : \mathcal{N} -precision for galaxy(blue) , matter(red) and the prediction-relation (green)	93
4.33	The homogeneity scale at 0.1% level, $R_H^{\mathcal{D}_2=2.997}(z)$, measured in the NGC (red) and in the SGC (blue) as a function of redshift. The purple point is the result obtained with quasars in the NGC, in the redshift range $2.2 \leq z \leq 2.8$ by Laurent et al. [166]. The green line is the Λ CDM model prediction. The shaded areas indicate the 1σ range for the 1000 QPM mock catalogs.	93
4.34	\mathcal{R}_H^V -estimator defined by equation 4.39 for 5 z -bins. <i>Left</i> : Galaxy <i>Right</i> : Matter.	94
4.35	Mean and dispersion of \mathcal{R}_H^E -estimator for NGC and SGC.	94
5.1	Contour plot with $1\sigma(2\sigma)$ in dark blue (light blue) for the Homogeneity scale versus the BAO scale for the $0.538 \leq z < 0.592$. See text for details.	130
5.2	<i>Upper</i> : Correlation coefficient (r) as a function of z for methods measuring \mathcal{R}_H and r_s . <i>Bottom</i> : \mathcal{R}_H and r_s measurement as a function of z for 1000 QPM mock catalogues. See text for explanation.	131

5.3	Contours of 68% (shaded) and 95% (light) C.L. of the CMB observable (blue), CMB combined with the BAO scale observable (red), CMB combined with the Homogeneity scale observable (blue) and CMB combined with the Homogeneity scale observable and the BAO scale observable (green) for the 7 cosmological parameters of $w_0wa\Lambda$ CDM-model.	132
5.4	Contours of 68% (shaded) and 95% (light) C.L. of the homogeneity scale of the galaxy distribution for the 3 cosmological parameters of Λ CDM-model at the intermediate redshift slice of North Galactic Cap. Different colors represent different configurations of the measurement [See text for details.]	135
5.5	Contours of 68% (shaded) and 95% (light) C.L. of the homogeneity scale of the galaxy distribution for the 3 cosmological parameters of Λ CDM-model for the 5 redshift slices (color-coded) of the North Galactic Cap. [See text for details.]	136
5.6	Contours of 68% (shaded) and 95% (light) C.L. of the homogeneity scale of the galaxy distribution for the 3 cosmological parameters of Λ CDM-model ($\Omega_m, \Omega_\Lambda, h$) and the 5 bias parameters of the 5 redshift slices of the North Galactic Cap. With blue is the CMB measurement of Planck 2013 combined with BAO, red CMB measurement combined with the \mathcal{R}_H and with green combine CMB, BAO and \mathcal{R}_H	137
5.7	Normalized fractional error $\delta R/R$ as a function of volume, V , for the bao scale (blue) and the homogeneity scale (green) for the the galaxy distribution. [See text for details.]	143
5.8	Normalized fractional error $\delta R/R$ as a function of density, n , for the bao scale (blue) and the homogeneity scale (green) for the the galaxy distribution. [See text for details.]	144
5.9	Normalized fractional error $\delta R/R$ as a function of photometric redshift error, σ_z , for the bao scale (blue) and the homogeneity scale (green) for the the galaxy distribution. [See text for details.]	145
5.10	Precision of the scale, \mathcal{R}_X versus the redshift, z for the different experiments which are color-coded. The dash lines correspond the the BAO scale while the continues lines correspond to the homogeneity scale for the matter distribution. See text for details	146
6.1	<i>Left:</i> Calibrating the Hubble expansion rate using cosmic chronometers. [See text for details]	152
6.2	Supernovae distance module, μ , as a function of redshift, z . Data points are denoted with blue, red line is the function of Λ CDM-model and orange is a simple relation between the distance module and the redshift.	153
6.3	(Top) Cosmological distance versus redshift for the FLRW comoving distance, d_C (blue) , model-independent transverse comoving distance, d_M (orange), and the two model independent cosmic clock comoving distances, d_{clock} (green and red). The black vertical lines indicate the redshift cuts of our survey. (Bottom) Different cosmological distances over comoving distance versus redshift.	154
6.4	Fractal Correlation Dimension as a function of scales, $D_2(r)$, for the CMASS galaxy sample using a model independent distance measure and confirming the <i>Cosmological Principle</i> . [See text for details]	155
A.1	(LIGO) Laser Interferometer Gravitational Wave Observatory.	169
A.2	<i>Left:</i> Gravitational Wave (GW) Signal as a function of time <i>Right:</i> Physical interpretation of the GW signal, which correspond to a coalescence of two Black Holes	169
A.3	Λ CDM-model Distance minus hubble distance as a function of redshift, z , for different fiducial cosmologies.	174

List of Tables

3.1	DR12 data sample in 5 redshift intervals in the NGC and SGC	54
4.1	The current list of 2pt correlation function estimators, $\xi(r)$, and their rescaled RMS, $r^2\sigma_\xi(r)$, in the region $r \in [80 - 120] h^{-1}\text{Mpc}$ as was estimated by Vargas-Magaña et al. [159]. Clearly the best option are between the three last estimators.	62
4.2	The three different estimators for the <i>scaled count-in-spheres</i> , $\mathcal{N}(< r)$ and <i>fractal correlation dimension</i> , $\mathcal{D}_2(r)$, studied in this thesis.	63
4.3	Fitted values of bias, b , and velocity dispersion, σ_p , in the different redshift bins, together with the corresponding reduced χ^2 for 24 degrees of freedom.	69
4.4	Homogeneity scale, $R_H^{\mathcal{D}_2=2.97}(z)$, for the galaxy and matter distributions in the north and south galactic caps.	71
4.5	Mean, error and χ^2 of the consistency test fit given by Eq. 4.23 with 23 degrees of freedom, in the NGC and SGC for the five redshift bins.	72
4.6	Fractal correlation dimension, $3 - \langle \mathcal{D}_2 \rangle_r(z)$, averaged over $300 < r < 1300 h^{-1}\text{Mpc}$, with 1σ errors, in the NGC and SGC.	72
4.7	Mean and error over the 1000 QPM mock catalogs for the χ^2 of the spline fit with 1 degree of freedom, in the NGC and SGC for the five redshift bins.	84
4.8	KS-test for North and South Galactic Cap. [See text for details]	95
4.9	Different Universe and their metrics	95
5.1	Weighted average over 5 redshift bins for the Homogeneity and BAO scale for the two different methods.	131
5.2	Prior measurements of the bias and the cosmological parameters space that we explore with the homogeneity scale of the galaxy distribution \mathcal{R}_H^G	134
5.3	Configuration of the simulation of the redshift density profile and survey area of the different experiments.	144

Abbreviations

CP	Cosmological Principle
GR	General Theory of Relativity
QFT	Quantum Field Theory
ΛCDM	Λ constant Cold Dark Matter
LSS	Large Scale Structures
SLS	Surface of Last Scattering
CMB	Cosmic Microwave Background
BBN	Big Bang Nucleosynthesis
BAO	Baryon Acoustic Oscillations
SN-Ia	Type Ia SuperNovae
CMASS	Constant stellar MASS objects
LRG	Luminous Red Galaxies
ELG	Emission Line Galaxies
QSO	Quasi-Stellar Objects or Quasars
LyaF	Lyman-a Forest absorption lines
CCD	charged coupled device
SDSS	Sloan Digital Sky Survey
BOSS	Baryon Acoustic Oscillation Survey
CFIS	Canadian-France Imaging Survey
DESI	Dark Energy Spectroscopic Instrument
DES	Dark Energy Survey
LSST	Large Synoptic Survey Telescope
EUCLID	Energy Up-to Colorful Luminocities In Darkness
2ptCF	Two point Correlation Function
RSD	Redshift Space Distortion
AP-test	Alcock-Paszynski Test
NGC	North Galactic Cap
SGC	South Galactic Cap
XGC	Either Galactic Cap

Physical Constants

Speed of Light	c	=	2.997 924 58 (exact)	$\times 10^8$	ms^{-1}
Planck constant	\hbar	=	1.054 571 800(13)	$\times 10^{-34}$	$\text{m}^2\text{kg}\text{s}^{-1}$
Gravitational constant	G	=	6.674 08(31)	$\times 10^{-11}$	$\text{m}^3\text{kg}^{-1}\text{s}^{-2}$
Coulomb constant ¹	k_e	=	8.987 551 787 368 176 4	$\times 10^9$	$\text{kgm}^3(\text{sC})^{-2}$
Boltzmann constant ²	k_B	=	1.380 648 52(79)	$\times 10^{-23}$	$\text{km}(\text{sK})^{-1}$

¹ $k_e = (4\pi\epsilon_0)^{-1}$ Most precised measured constant in physics for Electrostatic Effects. For electrodynamics, Q.E.D. prediction has the best measured prediction for a constant measurement [2].

² Physical Constants in S.I. units. Constants are quantities that remain unchanged in time (but what is time?), <http://physics.nist.gov/cuu/Constants/>. Many constants were shown to change in extreme conditions of nature.

Planck Units

Length	l_P	$= \sqrt{\frac{\hbar G}{c^3}}$	$= 1.616\ 229(38)$	$\times 10^{-35}$	m
Mass	m_P	$= \sqrt{\frac{\hbar c}{G}}$	$= 2.176\ 470(51)$	$\times 10^{-8}$	kg
Time	T_P	$= \sqrt{\frac{\hbar G}{c^5}}$	$= 5.391\ 16(13)$	$\times 10^{-44}$	s
Temperature	Θ_P	$= \sqrt{\frac{\hbar c^5}{G k_B^2}}$	$= 1.875\ 545\ 956(41)$	$\times 10^{-18}$	K
Electric Charge	q_P	$= \sqrt{\frac{\hbar c}{k_e}}$	$= 1.416\ 808(33)$	$\times 10^{+32}$	C

3

³Planck units are a system of natural units that is not defined in terms of properties of any prototype (Μέτρο kept in *Bureau International des Poids et Mesures, Se'vres, France* [3]), physical object, or even elementary particle. They defined by $c = \hbar = G = k_e = k_B = 1$.

Cosmo Units

Astronomical Unit	<i>A.U.</i>	= 1.495 978 71	$\times 10^{11}$	m
Lightyear	ly	= 9.460 528 4	$\times 10^{15}$	m
Parcesec	pc	= 3.085 677 58	$\times 10^{16}$	m
Degree	$^{\circ}$	= $\pi/180$	$\times 10^0$	<i>rad</i>
Right Ascession	R.A.	= 1	$\times 10^0$	$^{\circ}$
Declination	Dec	= 1	$\times 10^0$	$^{\circ}$
Arcmin	'	= 1/60	$\times 10^0$	$^{\circ}$
Arcsecond	"	= 1/360	$\times 10^0$	$^{\circ}$
Janskiy	<i>Jy</i>	= 1	$\times 10^{-23}$	$\frac{W}{m^2 \cdot Hz}$

Dedicated to my family

Chapter 1

Concordance Cosmology

1.1 Prelude

The current concordance picture of our Universe is described by the evolution and the interaction between matter and space. Two most well established theories exist that describe the fundamental laws of the universe from the smallest possible scales to the largest possible scales. The first one is theory of General Relativity (GR) which is the fundamental theory that we usually use to describe the universe at largest possible scales¹, i.e. up to the comoving diameter of the observable universe which is estimated to be $\sim 19\text{Gpc} \simeq 10^{25}m^{2,3}$. The second one is the Quantum Field Theory (QFT) which describes usually the behaviour of our universe at the smallest possible scales (down to the order of planck scale, $l_P \simeq 10^{-35}m^4$) which is about 20 orders of magnitudes smaller than the diameter of the known proton. Associating those theories⁵, a universal mathematical framework is built to explain the phenomena and the history of nature, with the name, across the physics community, $\Lambda\text{CDM-model}$. Along those 40-orders of magnitudes a very interesting and beautiful journey begins, that lasts for more than 13.799 ± 0.021 billion years [6] and it has many possible futures! The main pillars of the observational framework is the expanding nature of the universe, the existence of dark matter, the growth of structures, and the cosmic microwave background. The mathematical assumption that glues every piece together is the Cosmological Principle which states that:

The universe is homogeneous and isotropic, or smooth, at large enough scales.

and is the main topic of this thesis.

¹Some physicists combine GR with QFT in order to describe also the smallest possible scales, among other phenomena, known as *string theory*, in the framework of quantum gravity, predicting extra dimensions [4].

² Is the angular comoving distance, as described in section 1.2.6, at times corresponding to the emission of the first light at redshift $z \simeq 1090$.

³ Parsec is a unit of length used to measure distances of celestial objects outside our solar system with a technic called parallax

⁴The planck length scales, l_P , which are not yet observable, are the scales theoritized to be the quantization of space which makes up the fabric of the universe referred to as a quantum foam [5]

⁵Not yet a satisfactory combination of the two having the name *Quantum Gravity*

1.2 Initialising Λ CDM-*model*

At the core of the simplest Λ CDM-model, we find several assumptions that are in accordance with our current observations. The interpretation of these observations is structured around the assumption that the Theory of General Relativity is the correct theory of gravity on cosmological scales ($r \gg 10h^{-1}\text{Mpc}$) and gave birth to *Relativistic Cosmology*. This theory is based on several principles. One of them, the idea of relativity, is based on the equivalence of free-falling frames which led Einstein to conclude that gravity is not a force during 1915 [7]. Gravitational force, in the context of this theory, is explained as follows: one may consider a massless particle which free-falls and follows a trajectory in the space-time, called *geodesic*. The space-time continuum is locally curved in the presence of a gravitational field and is flat in the absence of it. This statement is formulated by a metric tensor⁶, $g_{\mu\nu}$, which characterises the invariant line element $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$ for a local 4-space coordinates of a non-euclidean geometry⁷, $x = x^\mu = (t, \vec{x}) = (t, x, y, z)$. Thus, the behaviour of this metric depends on the energy filling the spacetime. We can summarise this relationship in the Einstein's fields equations (EFE) as follows:

$$G_{\mu\nu}(x) = \frac{8\pi G}{c^4} T_{\mu\nu}(x) \quad (1.1)$$

where $G_{\mu\nu}$ is the *Einstein tensor* quantifying the behaviour of the metric $g_{\mu\nu}$, i.e. $G_{\mu\nu}(x) = G_{\mu\nu}[g_{\mu\nu}(x)]$. $T_{\mu\nu}$ is the *energy-momentum tensor* representing the energy, momentum density, pressure and shear. G is the Newton's gravitational constant and c is the speed of light in a vacuum. The aforementioned non-linear, second order differential equations of $g_{\mu\nu}$ (Eq. 1.1) are simplified by differential geometry considerations that are called Bianchi identities. These are expressed as $G^{\mu\nu}{}_{;\mu} = 0$, where ";" implies a covariant derivative⁸. These identities are related to the physical independence of Einstein's equations on the choice of a coordinate system which we call *gauge invariance*. Thus, the solutions of these equations presquest a fixing of a coordinate system.

1.2.1 A Smooth Universe

Friedmann⁹ has shown in the 20's that one can use the cosmological principle to build such a coordinate system in order to solve Einstein equations for a dynamical model that describe an expanding, homogeneous and isotropic universe, or *smooth universe*. The cosmological principle (homogeneity and isotropy) restricts the range of possibilities for such a metric. The isotropy condition restricts us to define a spatial metric through the expression of an infinitesimal 4-line element ds that is independent of a spatial component. The homogeneity condition allow us to define this metric with a radial dependence and parametrise it with a parameter k for a possible variation of the spatial curvature. Moreover, the cosmological principle allows a possible time-dependence of the space component of the metric and we parametrise it by an arbitrary function

⁶Greek letters are used as indices for all four spacetime components, while Latin letters are used only for the angular component. Furthermore, we use the einstein convention that repeating indices imply a sum.

⁷The partial derivative in respect of the coordinate system is defined as $\partial_\mu = (\partial_t, \partial_i) = (\frac{\partial}{\partial t}, \frac{\partial}{\partial \vec{x}})$

⁸which is a well-defined derivative in a curved space-time.

⁹Lemaître, Robertson and Walker, independently of Friedmann, have developed the same model during same epoch, 1920-1930.

of time $a(t)$, commonly called as *scale factor*, that works as a space dilator. For convenience, we write ds in spherical coordinates. As stated before, the gauge freedom imposes us to arbitrarily fix some terms of the metric. Usually the choice is a diagonal metric such that $g_{00} = 1$ and $g_{0i} = 0$, which simplifies equations. Thus the homogeneous and isotropic metric, namely FLRW metric, is thus written as :

$$ds^2 = c^2 dt^2 + a^2(t) \left[\frac{1}{1-kr} dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right] \quad (1.2)$$

where $d\Omega = d\theta^2 + \sin^2\theta d\phi^2$ reflects the isotropy condition and $\gamma_{ij} = \frac{1}{1-kr^2} dr^2 + r^2 d\Omega$. If $k = 0$, space is flat and infinite (critical). If $0 < k < 1$ space is spherical and finite (closed), while $-1 < k < 0$ correspond to a hyperbolic and infinite space (open). The 2D representation of the different curvatures are shown in Fig. 1.1. A more detailed description of the construction of the FLRW metric is given in appendix A.2. One notices that:

- So far, we have only assumed geometrical properties and the Cosmological Principle to define the behaviour of the coordinate.
- The FLRW metric allows us to define the *comoving observers* which are observers that measure the distances of objects that are not gravitationally bound to them. This defines us the *comoving distances*, d_C which factor out the expansion of the universe model by $a(t)$ and the *physical distances* which are the actual distances, $d_{ph}(t) = a(t)d_C$, measured by the usual observers, which depends on time.

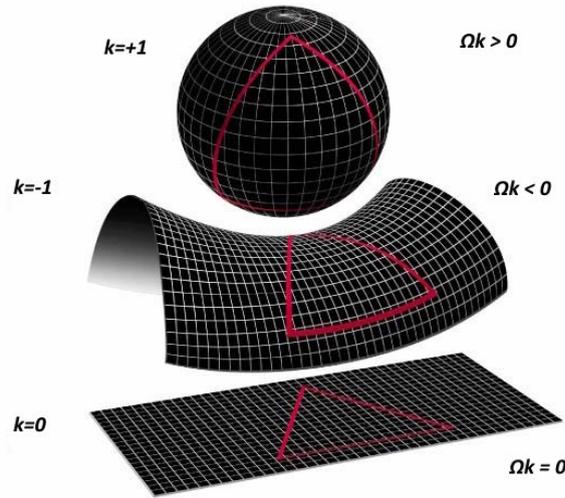


FIGURE 1.1: 2D Representation of possible variety of the curvature of our 3 dimensional spatial space.

The FLRW metric is used as an input on the left hand side of equation Eq. 1.1 for the computation of the scale factor as a function of the geometrical properties of the universe. Thus one can easily show that the Einstein Tensor, for an FLRW metric (Eq. 1.2), reduces to a tensor

with the following non-zero components:

$$G_{00} = 3 \left(\left[\frac{\dot{a}(t)}{a(t)} \right]^2 + \frac{kc^2}{a^2(t)} \right), \quad G_{ij} = -\gamma_{ij} [k + 2a(t)\ddot{a}(t) + \dot{a}^2(t)] \quad (1.3)$$

where dot "·" represent the derivative in respect of time t ¹⁰.

The right hand side of equation Eq. 1.1 describes the energy content of the universe as a perfect fluid in thermodynamic equilibrium, thus the energy-stress tensor takes the simplified form:

$$T_{\mu\nu} = \left[\rho(x) + \frac{P(x)}{c^2} \right] u_\mu u_\nu + P(x) g_{\mu\nu} \quad (1.4)$$

where $\rho(x)$ is the energy density, $P(x)$ is the pressure and u_μ is the 4-velocity. The cosmological principle implies that $u^\mu = (1, 0, 0, 0)$, meaning that the fluid is locally at rest with respect to the chosen frame. Furthermore, the cosmological principle restricts the energy density and pressure to be constant over space but allows a possible time dependence. These considerations model the stress energy tensor, with only non-zero components, as follows:

$$T_{00} = \rho(t), \quad T_{ij} = \frac{P(t)}{c^2} a^2(t) \gamma_{ij} \quad (1.5)$$

The gauge invariance allows us to add a constant on the Eq. 1.1, *cosmological constant*, Λ . By taking all the above considerations into account, the 00-component and the trace of Eq. 1.1 are written as:

$$H^2(t) \equiv \left(\frac{\dot{a}}{a} \right) = \frac{8\pi G}{3} \rho(t) - \frac{kc^2}{a^2(t)} + \frac{\Lambda c^2}{3} \quad (1.6)$$

$$- \left(\frac{\ddot{a}}{a} \right) = \frac{8\pi G}{2} \left[\rho(t) + \frac{3P(t)}{c^2} \right] - \frac{\Lambda c^2}{3} \quad (1.7)$$

where the $H(t) = \dot{a}/a$ is the Hubble expansion rate. The above differential equations are not enough to completely specify the system, i.e. $a(t)$, $\rho(t)$ and $P(t)$. Thus, either by combining the above equations or by using the local conservation of the stress-energy tensor ($T^{\mu\nu}{}_{;\mu} = 0$), we have that:

$$\dot{\rho}(t) = -3H(t) \left[\rho(t) + \frac{P(t)}{c^2} \right] \quad (1.8)$$

The set of the 3 latter equations (Eq. 1.6, Eq. 1.7 and Eq. 1.8) are used to describe the evolution $a(t)$ of the cosmic fluid with properties $\rho(t)$ and $p(t)$. This set of equations are called *Friedmann* equations. However, in the Λ CDM-modelling there are several species of the total cosmic fluid and a more detailed description is given in the following section.

1.2.2 The cosmic fluid

The most important matter content of the universe describing the large scale structures is the mixture of photons, baryons, cold dark matter and neutrinos which we call the *cosmic fluid*. Their properties are described by the quantum definition of the energy density and pressure.

¹⁰For a detailed derivation one may read the 1st chapter of the manuscript of Pieroni [8].

In the early hot universe, reactions occurred at such rates that the different species were in the thermal equilibrium, sharing a time-evolving temperature $T(t)$. Thus, their energy densities and pressure can be described by the statistical averages of the total relativistic energy $E(p) = \sqrt{(pc)^2 + (mc^2)^2}$ (0th component of the 4-momentum) and 3dimensional momentum P weighted by the distribution of particles in phase-space $f_X(\vec{x}, \vec{p})$, where the subscript X denotes the different species. These are formulated by:

$$\rho_X(\vec{x}, \vec{p}) = N_X \int \frac{d^3p}{(2\pi\hbar)^3} f_X(\vec{x}, \vec{p}) E_X(p) \quad (1.9)$$

$$P_X(\vec{x}, \vec{p}) = N_X \int \frac{d^3p}{(2\pi\hbar)^3} f_X(\vec{x}, \vec{p}) \frac{p^2}{3E_X(p)} \quad (1.10)$$

where N_X denotes the degeneracy of individual species (e.g. $N_X = 2$ polarisations of the photons).

Assuming, now, the Cosmological Principle, we can model the phase-space distribution of particles independent of position \vec{x} and orientation \hat{p} . This is described by standard Bose-Einstein ("+" sign) or Fermi-Dirac statistics ("- sign):

$$f_X(\vec{x}, \vec{p}, t) \rightarrow f_X(p, t) = \left[e^{(E_X(p) - \mu)/k_B T_X(t)} \pm 1 \right]^{-1} \quad (1.11)$$

where μ is the chemical potential, which to a good approximation is much smaller than the temperature over almost all times and species and hence $\mu \rightarrow 0$. Note that this also implies that the number of particles is equal to that of anti-particle, which is not true for baryons.

The cosmic species are divided into two general categories, i.e. the *relativistic* and *non relativistic* species, according to the level of their rest mass energy mc^2 . The former have a rest mass energy which is insignificant against their average kinetic energy $mc^2 \ll k_B T$. This leads to analytical solutions of Eq. 1.9 and Eq. 1.10 giving a simple relation between energy and pressure: $P_{rel} = \rho_{rel}/3$. The latter are those whose momentum is negligible to their rest energy ($mc^2 \gg k_B T$), and therefore $P_{n.rel} \simeq 0$. However, one may generalise those two approximated relations for the two categories of species with a parameter

$$w = \frac{P}{\rho c^2} \quad (1.12)$$

namely *equation of state* parameter. This allow for a class of solutions of Eq. 1.8, i.e.

$$\rho_X(t) \propto [a(t)]^{-3(w_X+1)} \quad (1.13)$$

for each species X .

It is convenient, now, to define the critical energy density as the energy density for a universe of zero curvature ($k=0$) and no cosmological constant ($\Lambda = 0$):

$$\rho_c(t) = \frac{3H^2(t)}{8\pi G} \quad (1.14)$$

Then by dividing Eq. 1.6 with $H^2(t)$, we have:

$$1 = \frac{8\pi G}{3H^2(t)}\rho(t) - \frac{kc^2 8\pi G}{8\pi G a^2(t) H^2(t)} + \frac{\Lambda c^2 8\pi G}{3 \times 8\pi G H^2(t)} \quad (1.15)$$

Now substituting Eq. 1.14 to Eq. 1.15 we have:

$$1 = \frac{\rho(t)}{\rho_c(t)} - \frac{1}{\rho_c(t)} \frac{kc^2}{8\pi G a^2(t)} + \frac{1}{\rho_c(t)} \frac{\Lambda c^2}{3 \times 8\pi G} \quad (1.16)$$

Last but not least, we introduce the ratio of energy densities of the possible species (X) of our universe against the critical density as:

$$\Omega_X(t) = \frac{\rho_X(t)}{\rho_c(t)} \quad (1.17)$$

where $X = \{\gamma, \nu, b, cdm, \Lambda\}$ correspond to photons, baryons, neutrinos, cold dark matter and dark energy, respectively. One may define as well the energy density ratio of curvature as:

$$\Omega_k(t) = -\frac{kc^2}{8\pi G a^2(t)}. \quad (1.18)$$

Therefore, all those species must satisfy the local energy conservation equation at all times:

$$\Omega_k(t) + \sum_X \Omega_X(t) = 1. \quad (1.19)$$

Thus in the field of concordance cosmology, we use the above simple parametrization (Eq. 1.17) to measure the ratio of energy densities of the different species in our universe. The convention we adopted is that when we drop the time dependence, we talk about the energy density ratio today $\Omega_X = \Omega_X(t=0)$.

1.2.2.1 Radiation: Photons and Neutrinos?

The early universe was dominated by radiation. Radiation is composed of primordial photons and neutrini. Photons being massless, they are always relativistic. These are the carriers of the electromagnetic interactions of charged particles. Most of the primordial photons, which we detect today, are in the form of the cosmic microwave background (CMB) and were emitted 380000 years after the planck epoch. These photons are thermalized and follow very close the black-body distribution with a temperature $T_\gamma = 2.7255 \pm 0.0006K$ as measured by Fixsen [9]. Therefore, the evolution of the energy density of CMB photons is written as:

$$\rho_\gamma(t) = 2 \int \frac{d^3p}{(2\pi\hbar)^3} \frac{pc}{e^{pc/k_B T_\gamma(t)} - 1} = \frac{\pi^2 k_B^4}{15c^3 \hbar^3} T_\gamma^4(t) \quad (1.20)$$

Since we have a relativistic species, $w_\gamma = 1/3$, we have that $\rho \propto a^{-4}(t)$ which means that the CMB photon temperature is inverse proportional to the scale factor, $T_\gamma(t) \propto a^{-1}(t)$. Intuitively, the temperature of photons, instead of the scale factor or time, is used to define the different epochs of the Universe. However this is not valid in very early epochs due to the change in the

photon distribution when leptons and antileptons annihilate. The density of entropy is another intuitive time indicator for the early universe since $s(t) \propto a^{-4}(t)$.

On the other hand, neutrini are relativistic fermions that interact weakly with matter. Recent experiments show that the three observable families of neutrini (electronic, muonic and taunic) are possibly a quantum linear combination of the three non-observable neutrino eigenstates. A measurable effect by Ahmad et al. [10], possibly explaining the neutrino mass, implies that neutrini oscillate between different families. These experiments have access to the difference between the square of two neutrino eigenstate masses, $\Delta m^2 \geq 0.0027 eV^2$, meaning that at least one neutrino would have a mass of at least $0.04 eV$. However, in this study, in the radiation domination era neutrinos are treated as massless particles to simplify our modelling.

1.2.2.2 Baryonic matter

Observing the night sky in the optical spectrum, we came to the realisation that the universe consist of baryonic and leptonic matter fluid. Besides that matter, the universe is consist of Dark Matter (described in next section). In this framework, we consider as baryonic matter the set of all the luminous matter i.e. the matter that interacts with the photons (baryons and leptons) and it is formed in the large scales in structures that we call galaxies, i.e. collections of gravitationally bound stars, objects which were formed at the late times of our universe $t = 8\text{Gyr}$ after the planck time. The baryonic matter is observed in increasing scales in different gravitationally bound objects such as planetary systems, stars, galaxies or in the form of dust in the intergalactic medium, composing totally about $\Omega_b = 4.9\%$. The baryonic matter interacts with the photons, thus it is easily observed through the photodetectors, such as the HST observatory[11]. An example of baryonic matter is the binary-galaxy, called *Chorodidascalous Andromedae*, belonging to the vast galaxy zoo. This binary is shown in Fig. 1.2, where the HST observatory observed the gravitational interaction of the two close-by galaxies.



FIGURE 1.2: The highly resolved Arp-734 binary galaxy observed by HST[12]. The name of this binary is *Chorodidascalous Andromedae*, or "Χοροδιδάσκαλος Ανδρομέδα" which means in greek translation "the *Dance-teacher Human-protector*", describing the gravitational interaction (dance) between the two galaxies.

Furthermore, observations show the existence of the leptonic matter such as electrons and muons which observed by numerous observatories. One of them is the HESS telescope based on cherenkov radiation detection[13]. The last element of large cosmic structures are the neutrini, which are detected by experiments such as the Hyper-Kamionkande water tank multi-photomultiplier detectors located in Japan[14]. Although we are going to, briefly, explain in the chapter 1.2.3 the physics of the interaction of baryonic, leptonic and neutrino matter with the photons and the metric, the budget of those species is negligible in respect of baryons and dark matter for this study. However, in the near future precision measurements of large scale structure observations, neutrini could play an important role at non-linear scales($r < 20h^{-1}\text{Mpc}$) on structure formation models. At the scales we are considering ($\sim 100h^{-1}\text{Mpc}$) the baryonic matter is usually formed in large scales in *galaxies* and collections of galaxies, *galaxy clusters*. The baryonic matter currently is mapped with various large scale structure surveys that we are going to discuss in chapter 3. One needs to note that there is also a small fraction $< 0.01\%$ of cold hydrogen that is a very faint source obviously and we refer to it as *dark baryons*. This was detected by the study of Fukugita et al. [15].

1.2.2.3 Dark matter

There is evidence for an additional matter fluid called, Cold Dark Matter (CDM) which does not interact with the photons and consist of $\Omega_{cdm} = 26.3\%$ of our universe today. The first strong astronomical evidences of this CDM were observed by the study of the rotational curve of the Coma galaxy cluster, shown in Fig. 1.3, by Swiss-American astronomer Fritz Zwicky [16]. In his study, he used, among other methods, the *virial theorem* of classical mechanics to determine the velocities and masses of celestial objects. With this method he revealed a missing mass component, naming it *Dark Matter*. Later on 1970, Vera Rubin[18] studying the rotational curve

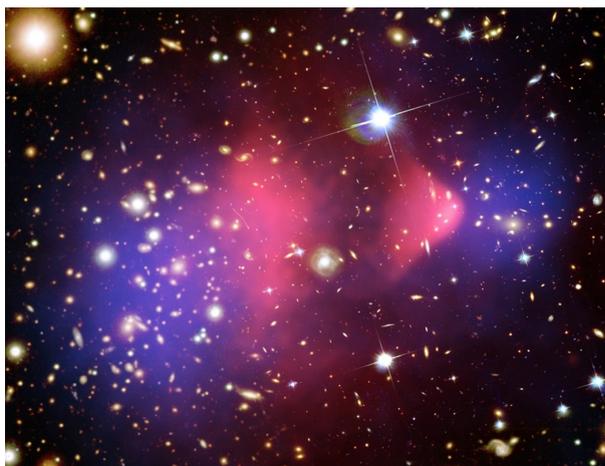


FIGURE 1.3: Composite image of the Bullet (star) Cluster observed by the Chandra X-ray Observatory [17]. We can see the X-ray(blue) and γ -ray (pink) radiation emitted by the collision of star cluster with another star cluster providing us with direct evidence of Dark Matter through gravitational lensing.

of M31 galaxy, i.e. studying the rotational velocities of the galaxy $v_{rot}(r)$ against the inferred mass from X-ray luminosity observations, found strong evidence for Dark matter as well. The

gravitational lensing studies of smaller structures, such as the Bullet Cluster[17] shown in fig 1.3, support the Dark Matter essence existence. Evidence of Dark matter has inspired many scientists to characterize the cosmic web of our universe. Thus the classification of the largest possible structures of our universe is a hot research subject at the time and there are many efforts of classifying the large cosmic web with the notions of nodes, filaments, sheets and voids which is a dynamical classification according to the tidal forces acting on the different structures due to the gravitational potential of the clustering at different corners of our universe¹¹.

The nature of Dark Matter is not yet determined, however many theoretical explanations exist, which model the Dark Matter as a hot neutrini or warm neutrini, a MAAssive Compact Halo Objects (MACHOS), a Weakly Interacting Particles (WIMPs), an oscillatory scalar field coupled to Higgs Boson[20] or other exotic objects.

1.2.3 Fluctuating from equilibrium

In the previous section, a homogeneous and isotropic model for our universe was introduced. However, the evolution of a homogeneous cosmic fluid cannot lead to the development of galaxies and the peculiar large scale structures that we observe on the night sky. The complexity of GR equations Eq. 1.1 does not allow for an analytical treatment of the evolution of the matter density field, $\rho(x)$, that describes the observed large scale structures. Therefore, we use the simplest linear perturbation modelling, considering small spatial density fluctuations, $\delta(x)$, compared to the spatial average matter density $\bar{\rho}(t) = \langle \rho(t, \vec{x}) \rangle_{\vec{x}}$ which allows us to expand equations to first order¹². Those differential equations become linear in this approximation. The agreement between this linear approximation and observations is a remarkable achievement, as shown in the latest large scale structure observations[21, 22].

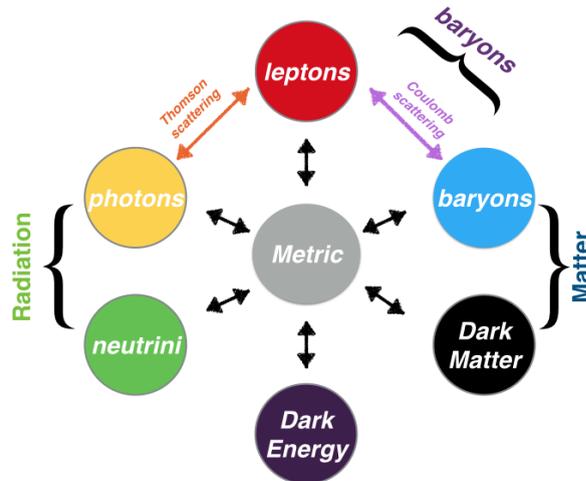


FIGURE 1.4: Schematic representation of the linear, coupled Boltzmann-Einstein field equation describing the interplay of physics at large scales, inspired by Fig 4.1 of Dodelson [23].

¹¹ An arising famous technique is the classification of the cosmic web using information theory [19].

¹²This is a good approximation for early times

The small spatial overdensities around the mean field of the different species of the cosmic fluid, X , are modelled, simply, via:

$$\rho_X(x) = \bar{\rho}_X(t) [1 + \delta_X(x)] \quad (1.21)$$

where the evolution of the mean density, $\bar{\rho}_X(t)$, is described via the smooth cosmology.

Two ingredients are necessary to describe the evolution of those perturbations and their corresponding velocity field $\vec{v}_X(x)$. The first ingredient is the perturbation of the FLRW space-time metric as:

$$ds^2 = - [1 + 2\Psi(\vec{x})] dt^2 + a^2(t) [1 + 2\Phi(t)] d\vec{x}^2 \quad (1.22)$$

where we have considered scalar perturbations defined via the $\Phi(t)$ spatial curvature field and the $\Psi(\vec{x})$ Newtonian potential field¹³, which is called the *synchronous gauge*. By neglecting Ψ and Φ scalar perturbations, we retrieve the homogeneous and isotropic, FLRW metric.

The second ingredient is the Boltzmann equations, describing the nature of the interactions and the evolution between the different species of the cosmic fluid, beyond the equilibrium. These are defined generally via:

$$D_t f_X(\vec{x}, \vec{p}, t) = C[f_X(\vec{x}, \vec{p}, t)] \quad (1.23)$$

where the left hand side describes the time evolution of the distribution of the temperature fluctuations of each species, $T_X(x) = T_X(t) [1 + \frac{\delta T}{T}|_X(\vec{x}, \hat{p}, t)]$, which we have developed in first order approximation¹⁴, while the right hand side describe the collision treatment between the different species X , $C[f_X]$. For the interaction between photons and leptons, we consider the classical *Thomson scattering* non-relativistic approach, $l^\mp + \gamma \leftrightarrow l^\mp + \gamma$ with an interaction rate $\Gamma \simeq n_l \sigma_T$, where $\sigma_T \simeq 2 \times 10^{-3} MeV^{-2}$ is the Thomson cross section. For cold dark matter, we consider a collisionless non-relativistic approach, as done in various famous structure formation history models. This are the simplest models that agree with observational large scale structure data. For baryons and leptons interactions, we assume a *Coulomb Scattering*, $b^\pm + l^\mp \leftrightarrow b^\pm + l^\mp$ in the Quantum ElectroDynamic (QED) approach. While for neutrini, we only consider them as a massless relativistic particle fluctuation overdensity and therefore we assume that they do not interact with matter. This is true only in the linear regime at large scales. Adopting a Fourier transform framework to simplify the equations in question, we end up with a set of 6 linear differential equations describing the non linear evolution of the 3 different species of density fluctuations (baryons, photons and neutrinos and Dark Matter) and their corresponding velocities at large scale as a function of conformal time¹⁵, η , and wavenumber, \vec{k} . However, this system is coupled to the 2 degrees, $\Phi(\eta)$ & $\Psi(\vec{k})$, of freedom defined by the perturbations of the curved metric. Thus, in order to completely specify the system one may solve the time-time component and the spatial trace of the Einstein equations using the perturbed metric defined via Eq. 1.22. Thus we end up with the coupled *Boltzmann-Einstein equations* that completely specify the system on large scale structures, i.e. the evolution of the density and temperature fluctuations,

¹³For a non-scalar perturbation treatment I refer the reader to the 5th chapter of [23].

¹⁴where $D_t = \partial_t + a^{-1}(t)\hat{p}^i + \partial_t\Phi(t) + a^{-1}(t)\hat{p}^i\partial_i\Psi(\vec{x})$ is the well defined derivative of the perturbed metric.

¹⁵The conformal time $\eta = \int_0^t dt'/a(t')$ defines the time needed for particles that travel in the speed c to reach an observer from the maximum distance existing in the universe (*observable universe*) which we call *particle horizon*.

$\delta_X(t, \vec{x})$ & $\frac{\delta T}{T}|_X(t, \vec{x})$, of the different species X. This interplay¹⁶ between the different species and the metric is represented schematically in Fig. 1.4.

Finally, all this special interaction leaves a special pattern on the structured density field, both temperature and matter. If the mean density of the structures (matter or temperature) is n , then we can characterise these inhomogeneities with the overdensity variable, $\delta(\vec{x}) = [n(\vec{x}) - n]/n$, or its Fourier transform $\tilde{\delta}(\vec{k})$. However observationally, we are interested in both $\delta(\vec{x})$ and also to the power spectrum $P(\vec{k})$ which is defined as:

$$\langle \tilde{\delta}(\vec{k}) \tilde{\delta}^*(\vec{k}') \rangle = (2\pi)^3 \delta_D(\vec{k} - \vec{k}') P(\vec{k}) \quad (1.24)$$

where $\delta_D(\vec{k} - \vec{k}')$ is a Dirac distribution. This way, we can characterise all the modes, \vec{k} that have interacted within the observable universe.

1.2.4 Expanding universe

The expanding of the universe was first evidenced by the famous astronomer, Edwin Hubble [25], who determined, in 1929, a linear relation between the recessional velocities (v_{rec}) and the radial distances (d) of galaxies¹⁷, at low redshift ($z < 0.01$ observations):

$$v_{rec} := cz = H_0 d \quad (1.25)$$

where H_0 is the hubble constant measured today, hence the subscript 0 and z is the redshift variable. This defines the hubble horizon today, $d_H \equiv d_H(z=0) = c/H_0$, which is the maximum distance a photon can travel during a time as long as the age of the universe. The recessional velocities, or redshift z , can be obtained by the spectra information. The spectra information for each source emitter is simply determined by the relative difference between the emitting source wavelength at the rest frame of the emitter, $\lambda_{em}(t_e)$, and the wavelength that we observe, $\lambda_{obs}(t_o)$. This leads to the relation for the expansion rate of the universe, since:

$$1 + z = \frac{\lambda_{obs}(t_o)}{\lambda_{em}(t_e)} \equiv \frac{\lambda_{comov} a(t_o)}{\lambda_{comov} a(t_e)} = \frac{1}{a(t)} \quad (1.26)$$

where we have adopted the convention that $a(t_0) = 1$. The distances were determined by the *standard candle technique*, which is basically the distance measurement through measuring the relative fluxes of those objects in respect of their rest frame spectra, which are obtained by studying chemical element emission in the laboratories. This allows us to rewrite the first Friedmann Eq. 1.6 as:

$$H(z; H_0, \vec{\Omega}) = H_0 \sqrt{\Omega_k (1+z)^2 + (\Omega_{cdm} + \Omega_b)(1+z)^3 + \Omega_\Lambda f(z) + (\Omega_\nu + \Omega_\gamma)(1+z)^4} \quad (1.27)$$

¹⁶For a simplistic derivation of Boltzmann-Einstein equations one may revise the 4th and 5th chapters of Dodelson [23] or the lecture notes of Baumann [24].

¹⁷Cepheids Variables are stellar objects within galaxies, undergoing pulsations in very regular periods on the order of days or months and are established by Hubble for determination of galaxy distance within our local group.

where $f(z)$ quantifies the different flavours of Dark Energy models[26], with $f(z) = 1$ corresponding to the *cosmological constant* Λ . This defines as the Hubble expansion rate, $H(z; \vec{\Omega})$ as a function of redshift information, z , and the various energy density parameters quantified as $\vec{\Omega} = (\Omega_k, \Omega_{cdm}, \Omega_b, \Omega_\nu, \Omega_\gamma, \Omega_\Lambda)$. This equation is useful for all cosmological observations since we have only access on the redshift information z of the celestial objects.

1.2.5 Evolving Fluctuations

The solutions of the Boltzmann-Einstein equations depend on the choice of the initial conditions. The choice of initial condition is related to the different hypotheses for the origins of the universe, one of the most intriguing questions in physics today. We are going to give a brief discussion on the initial conditions in chapter 1.2.7, while here we are going to describe the current picture of the solutions. Observations have shown that the Universe was so hot and dense in the past that baryons (and leptons) and photons acted as a single fluid, a plasma. During this hot period, each plasma overdensity was growing due to the interplay between gravity, pressure and expansion. The number density of free leptons was so high that the mean free path of the Thomson scattering of photons was extremely hot. On the other hand, this plasma was being pulled away of these potentials due to thermal pressure. This interplay between gravity and pressure around these density fluctuations lead to the propagation of density waves outwards of the potential wells. These sound waves propagate in space until the pressure falls to zero after the recombination epoch. At that time the universe cools down to energy levels, $E < 13.6eV$, letting the baryons recombine with leptons, at a time corresponding to $z_* = 1090$. When the baryons and leptons stop feeling the pressure of photons, defined as drag epoch, $z_{drag} = 1060$, the sound waves stops propagating and leave a slightly over-dense shell, $r_d(z)$, of baryonic matter around each original perturbation. This drag shell is described by the oscillatory solutions, $kr_s(\eta)$, of the damped Boltzmann-Einstein equations for the baryon-lepton-photon fluid, where r_s is the comoving distance traveled by the sound wave at conformal time η , i.e. :

$$r_s(z) = \int_0^\eta d\eta' c_s(\eta') = \int_\infty^{z(\eta)} dz' \frac{c_s(z')}{H(z')} \quad (1.28)$$

where the sound speed in the plasma is simply given by:

$$c_s^{-1}(z) = \sqrt{3 \left[1 + \frac{3\rho_b(z)}{4\rho_\gamma(z)} \right]} \quad (1.29)$$

This phenomenon is called baryon acoustic oscillations (BAO). Since drag epoch, the radius of these shells is only affected by the expansion of the Universe, with negligible gravitational collapse into the central region of original perturbations. Those oscillations leave a pattern that can be described in the Fourier space by the power spectrum $P(\vec{k})$ of the density Fourier modes, $\delta(\vec{k})$. Since dark matter fluid is practically not affected by other species or thermal pressure(except at small scales), simply falling into gravitational potential wells, it is common practice to use the dark matter power spectrum as a reference to study large scale structures at all times. However, dark matter fluctuations, during radiation dominated era, are indirectly affected by radiation. The potential wells are subdominant in respect to dark matter itself during

matter dominated era. On the other hand, the effect of baryonic fluctuations is negligible on the overall shape of the power spectrum apart from the small wiggles at high modes due to BAO. Linear perturbation theory predicts that large scale modes keep their gaussianity until later times. Non-linear physics affect small scale modes, coupling different wavelengths, and creating non zero odd point correlation functions which is the fourier transform of the power spectrum.

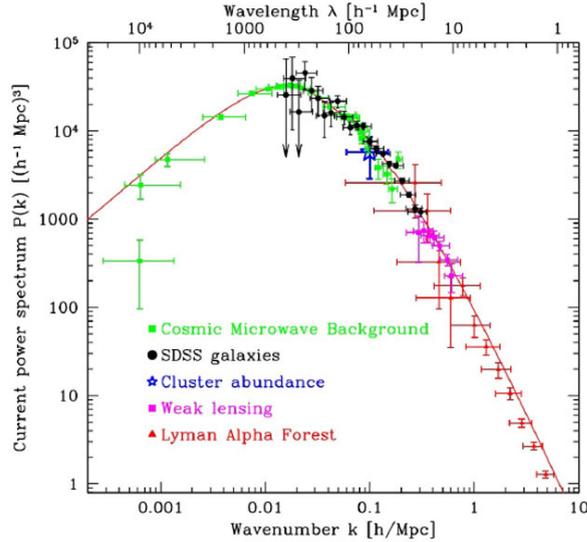


FIGURE 1.5: Measurement of the matter power spectrum observed by combination of different surveys[27].

The initial conditions¹⁸, observationally, are set with a scale-independent power spectrum $P(k) \propto k^{n_s-1}$ that evolves with time or redshift and $n_s \simeq 1$. Once the wavelength of a given density mode $\lambda_k = 2\pi k^{-1}$, is smaller than the hubble horizon, $d_H(z) = c/H(z)$, at a given time, the Dark matter collapses and this mode is in causal contact. Matter density perturbations grow by the gravitational potential with time. However, during radiation dominated era the rate of growth is insignificant due to the thermal pressure. Outside the hubble horizon modes are growing since they are not causally connected and they are not affected by the pressure, this will become clearer in chapter 1.2.7. This leads to a break in the observed power spectrum slope at this wavelength, as shown in Fig. 1.5.

1.2.6 Accelerating universe with standard rulers

In 1998, Riess, independent from Schmidt and Perlmutter[28] led teams to performed a sophisticated analysis using Type Ia Supernovae (SNIa), i.e. explosions of stars,¹⁹ and revealed the accelerating nature of the universe. This measurement was recently updated by a compilation of different projects[29] using SN up to redshift $z \lesssim 1$. Figure 1.6 shows the scale factor of the universe as a function of time which is the analogue of the measured brightness and redshift of the supernovae revealing the preference of an accelerating universe against a forever expanding

¹⁸ We will explain this choice of the initial conditions in more details in chapter 1.2.7

¹⁹ Explosion of stars with generated luminosities equivalent of the host galaxy. This explosion occurs only after the mass of the star reaches the Chandrasekhar limit due to gravitational attraction of their massive environment releasing energies up to $E_{SN} \simeq \frac{3}{5} \frac{G M_{SN}^2}{R_{SN}} = 2 \times 10^{46} J$ for a mass $1.4M_{\odot}$ and a radius 15km.

or decelerating or collapsing. These methods for the determination of the acceleration of the universe are called *standard candles* methods. Standard candles are objects whose luminosity is known, for instance exploding stars (Supernovae) that have the same luminosity at the maximum of their explosion. This is true for Type Ia supernovae with a spread of their maximum luminosity (after some corrections) of order 12%. Having the apparent luminosity of the object, and knowing its absolute luminosity, one can directly measure the luminosity distance and compare it with the redshift of the object, measuring a Hubble diagram which depends on the cosmological parameters.

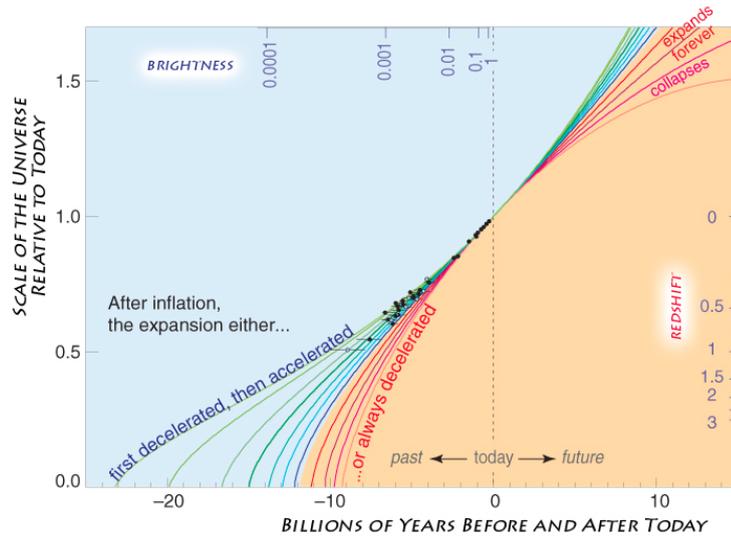


FIGURE 1.6: Scale factor as a function of time, measured by a compilation of the Brightness of samples of Supernova at different redshifts. One can notice that the "always accelerating" scenario is best fitted to these data [Image taken from [30]]

A similar concept, uses standard rulers, e.g. special rods that are objects with a known size. In particular, the BAO features of the universe which are imprinted on the large scale structures remain constant in time. Therefore, they can be used as a standard ruler to measure mainly the acceleration, and expansion but also a possible curvature and matter to baryon ratio of the universe. This interesting feature does not change during the whole period from recombination to nowadays, which means that the sound horizon is only affected by the expansion of the universe. This is due to that the large size of the sound horizon, of about $110h^{-1}\text{Mpc}$ (larger than any collapsed structure) protecting this feature from non-linear effects of structure formation. One of the main differences with respect to *standard candles* is that the BAO signal is so weak that statistical measurements over large volumes are needed.

Alcock and Paczynski [31] derived a test, namely AP-test, to optimally and independently extract the cosmological information from the BAO features. To do so in practice we need two ingredients:

- a A way to measure distances of the underlying distribution
- b A way to study this distribution

To measure radial comoving distances, we assume an FLRW metric (Eq. 1.2) and by adopting $ds = d\theta = d\phi = 0$, we are lead to:

$$d\chi(z) = \frac{c}{a(t)} dt = -\frac{c}{H(z)} dz \quad (1.30)$$

Then, we can express the comoving radial distance at redshift z of an object by integrating the redshift information of photon emitted by the source at z to now ($z=0$) when it was observed, as:

$$d_C(z) = c \int_0^z \frac{dz'}{H(z')} \quad (1.31)$$

Then, by having the 2 dimensional angular information on the sky, i.e. (θ, ϕ) of every galaxy, we reconstruct the 3 Dimensional distribution in comoving cartesian coordinates (x,y,z) . However, this comoving distribution is only valid for a flat ($k = 0$) universe. If we need to measure the transverse size of this distribution in the sky at a specific redshift we need to assume FLRW metric Eq. 1.2. However, since we are interested in the angular coordinates, the Eq. 1.2 fail to describe[32] the whole 3D-Sphere with curvature $k=+1$. Therefore, the metric is shown to be rewritten as:

$$ds^2 = -c^2 dt^2 + a^2(t) [dr^2 + S_k^2(r)(d\theta^2 + \sin^2\theta d\phi^2)] \quad (1.32)$$

where

$$S_k(r) = \left\{ \begin{array}{ll} \sqrt{k}^{-1} \sin(r\sqrt{k}) & , k > 0 \\ r & , k = 0 \\ \sqrt{|k|}^{-1} \sinh(r\sqrt{|k|}) & , k < 0 \end{array} \right\} . \quad (1.33)$$

Now, if we adopt $dt = dr = d\phi = 0$ on the new version of FLRW metric (Eq. 1.32), we end up to:

$$ds = a(t)S_k(r)d\theta . \quad (1.34)$$

The above equation tell us that the ds is the proper transverse size of the distribution (or the transverse size of BAO), $d\theta$ is the comoving angular size and $d_A(t) = a(t)S_k(r)$ is the physical angular distance to the object. Then by using the time-redshift relation (Eq. 1.26), the measured angular size of an object is rewritten as:

$$d\theta = \frac{ds}{a(z)S_k^2[d_C(z)]} := \frac{ds}{d_A(z)} \quad (1.35)$$

Notice, that we substituted $r = d_C(z)$ at the specific redshift, z , that we are measuring. Therefore, the angular physical distance of any object, $d_A(z)$, is defined via:

$$d_A(z) = \frac{1}{1+z} \left\{ \begin{array}{ll} \frac{d_H}{\sqrt{\Omega_k}} \sinh\left(\sqrt{\Omega_k} \frac{d_C(z)}{d_H}\right) & , \Omega_k > 0 \\ d_C(z) & , \Omega_k = 0 \\ \frac{d_H}{\sqrt{\Omega_k}} \sin\left(\sqrt{\Omega_k} \frac{d_C(z)}{d_H}\right) & , \Omega_k < 0 \end{array} \right\} \quad (1.36)$$

To study now the distribution, we need in our disposal, the excess probability dP of finding an object within a volume $d\vec{r}$ around another object of a density field with mean density $\bar{\rho}$. This

expression is quantified as follows:

$$P \equiv \int dP = \int \bar{\rho} [1 + \xi(\vec{r})] d\vec{r}. \quad (1.37)$$

A detailed derivation of the 2-point correlation function is given in appendix A.3. Thus, we study the BAO features statistically via the the 2-pt correlation function, $\xi(\vec{r})$. Since the BAO feature is a 3Dimensional object, we decompose $\xi(\vec{r})$ into comoving line-of-sight (l.o.s. hereafter) dependence, comoving parallel in the l.o.s. dependence.

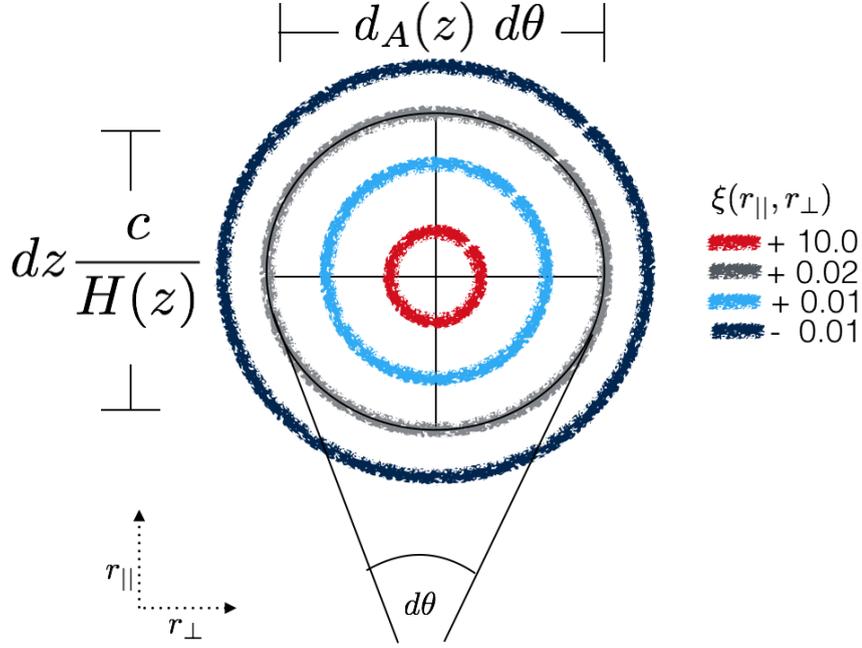


FIGURE 1.7: Illustration of the AP-test of large scale structure clustering. [See text for details]

Since, in the radial (redshift) direction, one measures comoving sizes of objects, their physical radial size dl will be related to the expansion rate of the universe via $dl = cdz/H(z)$. Therefore the physical BAO size parallel to l.o.s. will be:

$$dl_{BAO,\parallel}(z) = dz_{BAO}c/H(z) \quad (1.38)$$

While its transverse comoving size one measures the projected expansion of the universe, therefore measures a dl which is proportional to the integrated expansion rate of the universe. For any universe of curvature Ω_k , the transverse comoving size is proportional to the comoving angular distance, $d_{A,c} = d_A(z)/(1+z)$, i.e. :

$$dl_{BAO,\perp}(z) = (1+z)d_A(z)d\theta. \quad (1.39)$$

If the measurement is completely spherical, then $dl_{BAO,\perp} = dl_{BAO,\parallel}$ holds²⁰ and therefore:

$$\frac{dz}{zd\theta} = \frac{1+z}{cz}d_A(z)H(z) \quad (1.40)$$

²⁰Notice that this holds for every curvature of the FLRW-metric Eq. 1.2.

which is the definition of the AP -test. The test is illustrated in Fig. 1.7, where we can observe the measurement of $\xi(r_{\parallel}, r_{\perp})$, which is represented with different colors. One can see that the peak is illustrated with the grey color.

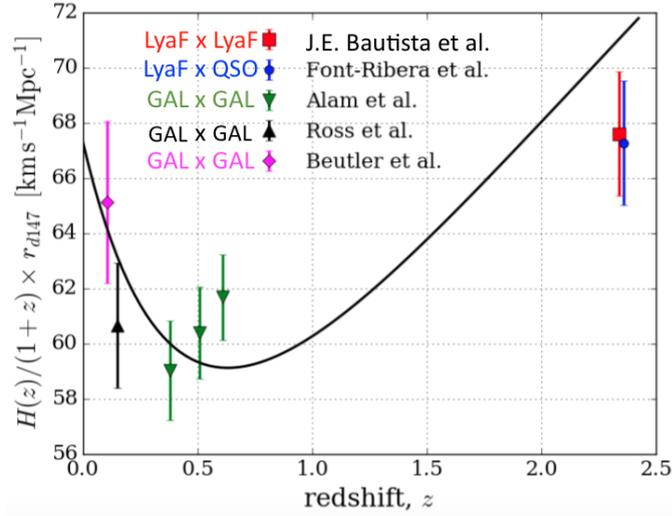


FIGURE 1.8: Large scale structure clustering measurements, revealing the accelerating nature of our universe with a combination of AP -tests [Image drag from [33]]

Currently, the most sophisticated AP -test try to take the advantage of the analysis of Lyman- α Forests[33], which are the absorption lines coming from the gas that fill the space in front of the line of sight of observations of quasars²¹ (QSO). The state of the art measurement of accelerated expansion of the universe from large scale structure clustering measurements are summarised in the right panel of Fig. 1.8. They have used a compilation of different total matter tracers, i.e Gal: galaxies QSO quasars and Lyman-a Forest, as long as the combination between Lyman-a and Quasars. In each compilation they have measured the BAO feature which underlies at each tracer. Measuring this feature they determined the Hubble expansion rate, which was normalized by the scale factor and calibrated with the drag shell, $r_{d147} = 147.33\text{Mpc}$ as measured by current CMB experiments. This amazing result shows that the rate of the expansion of the universe $H(z)/(1+z) = a(z)\dot{a}(z)/a(z) = \dot{a}(z)$ was decreasing in epochs, $z > 1$, while later on, $z < 1$ the universe started an accelerating expansion.

1.2.7 Initial conditions and the Inflation Paradigm

In order to understand the evolution of the structures in the universe, we have explored the equations of the perturbations around a smooth background. However, in order to solve those *Boltzman-Einstein* equations, we need a proper choice of initial conditions. As discussed in Dodelson [23], the quest of the initial conditions has lead to a new realm of physics, the theory of *Inflation*. Historically, inflation was introduced by Guth [34], Linde [35], Albrecht and Steinhardt [36] to explain basically three issues of ΛCDM -model:

²¹Quasars or QSO (Quasi stellar objects) are galaxies that form in the early universe, $z \geq 2$ with a high mass, usually $M_{qso} = 10^9 M_{\odot}$. Quasars have a supermassive black holes in their center having very bright accretion discs.

1. The Horizon Problem
2. The Flatness Problem
3. The Magnetic Monopole problem

Let's consider the following paradigms to understand how this mechanism solve all the above problems. In the context of inflation, the universe initially, during planck time²², $t \sim t_P$, passed through a rapid exponential expansion phase, $a(t) \propto t^{Ht}$, until the end of its phase t_f , corresponding to $z \sim 1090$. For the *horizon problem* consider two facts:

- a the *physical wavelength of fluctuations*, λ_p , increase due to the expansion of the universe
- b the *physical horizon*, $d_{p,hor}$, depends on time.

The physical horizon, constitutes of all the spacetime regions where individual points can interact with each other and it is expressed as:

$$d_{p,hor}(t) = a(t)d_{hor}(t) = a(t) \int_0^t \frac{cdt'}{a(t')} = a(t) \int_{\log a(t=0)}^{\log a(t)} \frac{cd \log a}{aH(a)} \quad (1.41)$$

In standard Λ CDM-model, i.e. without the extension of theory of inflation, with a cosmic fluid which has constant in time equation of states²³ we end up to a situation where the physical horizon is proportional to the Hubble radius, $d_{p,hor}(t) \propto (aH)^{-1}$. While the physical wavelength of fluctuations is only $\lambda_p(t) = a(t)\lambda_c$. This means that the *physical horizon grows faster than the physical wavelength of fluctuations*. Thus, the largest observed scales today were outside the horizon at early times. Hence, the individual points on the spacetime fabric were not causally connected. Qualitatively, this translates to about 10^4 causally disconnected regions at times, $z \sim 1100$. This is illustrated in Fig. 1.9 in which the p and q point are outside the physical horizon before reaching our world line (w.l.). However, observations have shown that the temperature at that time was everywhere uniform, isotropic (see next chapter 1.2.8), and therefore should be causally connected. Thus, by introducing a mechanism of a fast expanding universe at early times, the *physical wavelengths of fluctuations grow faster than the physical horizon*. This means that the fluctuations were inside the physical horizon at initial times, but they stretched outside of the horizon during slightly later times, inflationary era, and re-enter the horizon in the later universe, at matter-radiation era, $a \simeq 0.5$, This mechanism force individual points of space interact with each other in the past even though now they are observed far away from each other.

In parallel, the *flatness problem* is considered by the fact that while standard Λ CDM-models can have any curvature, observations have shown a value of the curvature which is very small, $\Omega_k(t=0) \simeq 10^{-62}$. Nonetheless, standard Λ CDM-model solutions with zero curvature, i.e. $\sum_X \Omega_X = 1$, are unstable. Any slight difference of the total energy density ratio value can

²²Alternative theories exist for prior planck time, but none of them were verified observationally

²³If we substitute Eq. 1.13 to Eq. 1.6 for $\Lambda = k = 0$, we have: $[aH(t)]^{-1} = H_0^{-1}a^{(1+3w)/2}(t)$ and hence $d_{hor}(t) = \frac{2cH_0^{-1}}{1+3w} = a^{(1+3w)/2}(t) - a^{(1+3w)/2}(t=0)$. But, the comoving horizon receives its largest contributions from late times $\frac{1}{1+3w}a^{(1+3w)/2}(t=0) \xrightarrow{w > -1/3} 0$ and therefore $d_{ph}(t) = \frac{2}{1+3w} [a(t)H(t)]^{-1}$.

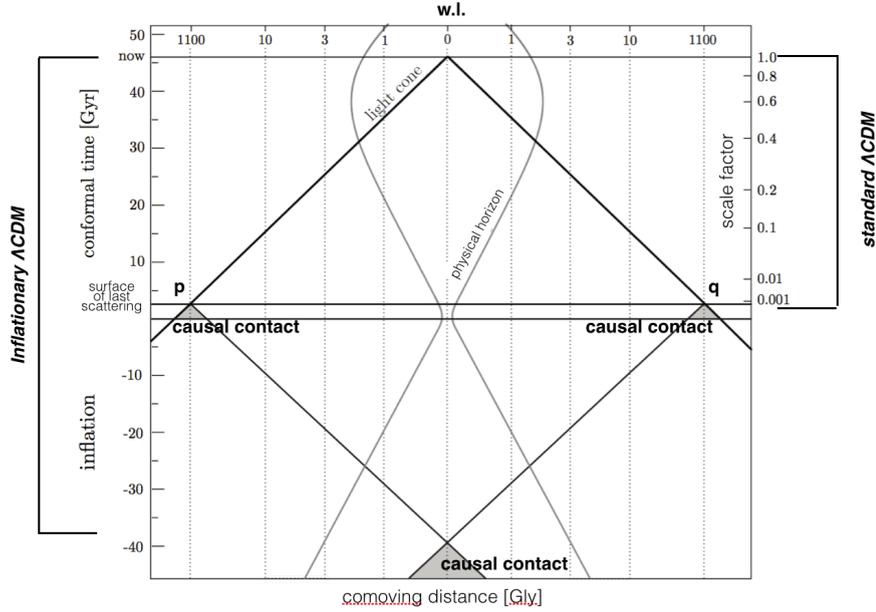


FIGURE 1.9: : Illustration of the inflationary solution to the horizon problem. The comoving physical horizon shrinks during inflationary epoch and expands during the standard Λ CDM-model evolution at least until later time at $a > 0.5$. This allows all large physical fluctuations (p, q points at the surface of last scattering which are outside the physical horizon) to be in causal contact with each other at initial times since in the past they were inside the physical horizon. [Diagram was taken by Baumann [37]]

cause a rapid expanding universe at early times. For example, if at early times $\sum_X \Omega_X(t = 10^{44}t_P = 1\text{sec}) = 0.9$, the universe would grow to a value of 10^{-14} today. However, if at early times $\sum_X \Omega_X(t = 10^{44}t_P = 1\text{sec}) = 1.1$, the universe would grow so rapidly that it would have recollapsed just 45 seconds after Planck times. The resolution comes to this problem by forcing a rapid accelerating period at early times, $\ddot{a}(t \simeq t_P) > 0$ and therefore the $\sum_X \Omega_X = 1$ becomes an attractor solution. If, for example, the scale factor grew by a factor of e^{60} , namely *inflation lasted for at least 60 e-folds*, then $\sum_X \Omega_X$ is driven so close to 1 that we will still observe it near 1 today.

The inflation theory could also be a solution for magnetic monopoles problems, using the baryon/antibaryon asymmetry. Tiny perturbations that happened during inflation could create little regions of more baryons than antibaryons and destroy the production of any magnetic monopoles that were introduced by standard homogeneous and isotropic Λ CDM-models.

To explain briefly the inflation mechanism, consider that during early times, $t \sim t_P$, the universe is dominated by a very different form of energy. This is often modelled by a scalar field ϕ , *inflaton*, with potential energy density $V(\phi)$. During the inflationary era, this field is displaced from a global vacuum in a state of high energy density. Assuming, further, that this field in this vacuum state dominates all contributions of the total energy density, the stress energy tensor of Eq. 1.1 suggests in this paradigm that the energy density of the universe is determined by the kinetic energy $\frac{1}{2}\dot{\phi}^2$ and each potential energy $V(\phi)$:

$$\rho(\phi) = \frac{1}{2}\dot{\phi}^2 + V(\phi), \quad (1.42)$$

while its pressure is determined by space component of the stress energy tensor, e.i.:

$$P(\phi) = \frac{1}{2}\dot{\phi}^2 - V(\phi) . \quad (1.43)$$

An enough slowly evolving field, i.e. $V(\phi) \gg \frac{1}{2}\dot{\phi}^2$, results in a source of negative pressure, since $P(\phi) \simeq -V(\phi) \simeq -\rho(\phi)$. Under negative pressure fluid condition, the solution of Friedmann equations, ignoring curvature and Λ terms result to $H^2(\phi) \simeq \frac{8\pi G}{3}V(\phi)$. Therefore, solving under these conditions Eq. 1.8, we have :

$$a(t) \propto e^{Ht} . \quad (1.44)$$

This mechanism explaining the exponential expansion during early times is called *slow-roll inflation*.

Finally, inflation mechanism give rise to quantum fluctuations that become macroscopic during that time²⁴. Considering that the primordial inflaton, $\phi(t, \vec{x})$, is splitted into a homogeneous background, $\bar{\phi}(t)$ and a spatial varying perturbation, $\delta\phi(t, \vec{x})$, then $\delta\phi(t, \vec{x})$ satisfy the equations of motion of a harmonic oscillation with a time-dependent mass. The quantum treatment of the perturbations, translates to a collection of one-dimensional harmonic oscillators spatially distributed. The zero-point quantum fluctuations during inflation induce a non-zero variance for the inflation perturbations. Therefore, the end of inflation is modulated by the fluctuations on the inflaton. This means different fluctuations stop at different times. This process can convert the inflaton perturbations into energy density $\delta\rho(t, \vec{x})$. The efficiency of the production of density perturbations depends on the slope of the inflaton potential, while the size of quantum fluctuations on its amplitude. The full calculation result to a parametrization on the primordial power spectrum, i.e. the power spectrum of the metric:

$$P_{prim}(k) = A_s k^{n_s - 1} \quad (1.45)$$

where A_s is the *amplitude of the primordial scalar perturbations* and n_s is called the *spectral index*. Several models of inflation predict that n_s is less than one. Currently, Planck Mission measure its value at near percentage level, i.e. $n_s = 0.9667 \pm 0.0040$.²⁵

1.2.8 Cosmic Microwave Background

In 1965 A. Penzias and R. Wilson[38], found an excess noise on the measured temperature of their antenna. In particular they have been using a Horn-reflector antenna, 6m long, which they were pointing it to the sky at various directions away from the galactic plane which covers 16% of the sky. They have found an unpolarized, isotropic temperature excess of $3.5K$ at around 4GHz frequency that remain constant at seasonal variations. They verified that this temperature was not coming from the receiver of the antenna or other observed known radio galactic sources on the sky. This temperature had an unknown source and a possible theoretical

²⁴For a detailed discussion, I refer the reader to Baumann [37].

²⁵Note that the power spectrum of the densities is defined as $P_\delta \propto \langle \rho^2 \rangle \propto k^{n_s}$. Since the poisson equation tells us that $\nabla\Phi \propto \rho$ or in fourier space $k^2\Phi = \rho$. From eq 17 of Baumann [37] of the definition for the Primordial Power spectrum P_ϕ , we have that the spectrum of the metric is $P_\Phi \propto k^3 \langle \Phi^2 \rangle$. Substituting Poisson equation we have $P_\Phi \propto \langle \rho^2 \rangle \propto k^{-1}$. But since $\langle \rho^2 \rangle \propto k^{n_s}$, we have that the primordial power spectrum of the metric is $P_\Phi(k) \propto k^{n_s - 1}$.

explanation and a confirmation of the measurement was given by Dicke et al. [39]. In their letter, they explained that the primordial universe due to thermal considerations possibly emits a temperature at microwave frequency matching the observations of Penzias and Wilson. This emission was named Cosmic Microwave Background (CMB) since it originates from the thermal component of the early universe. This thermal component was predicted by Alpher et al. [40] in 1948, that explain in a rigorous manner the origins of the chemical elements of the early universe. This light from that epoch of recombination comes to us from a spherical shell around us called the Surface of Last Scattering (SLS).

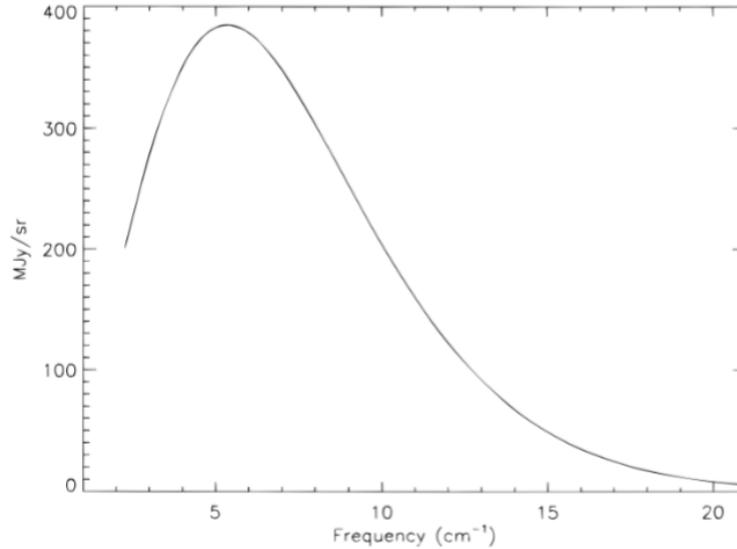


FIGURE 1.10: Uniform frequency spectrum (intensity vs frequency) fitted on the theoretical Planck Black Body curve. The errorbars are only a small fraction of the line thickness[41].
[See text for details]

Later on several experiments were deployed to characterise this emission. Fast forward in time, in 1989, the FIRAS spectrometer (Far-InfraRed Absolute Spectrophotometer) [41], placed on the COBE satellite (COsmic Background Explorer) manage to measure the frequency spectrum of the CMB emission in high precision. It was detected with a temperature of $T_{cmb,f} = 2.75277 \pm 0.002\text{K}$ at a peak-intensity located around 282.2GHz corresponding to a maximum wavelength of $\lambda_{max} = 1.067\text{mm}$. The measured spectrum revealed a Planck Black Body spectrum of CMB emission as shown in Fig. 1.10.

This observed CMB radiation is very isotropic. However, as it was first measured by a small device on the U-2 aircraft [44], it turn out to have a dipole anisotropy, $T = T_{cmb,f}(1 + \frac{v}{c}\cos\theta)$. Current estimates give a best fit value of a temperature $3.343 \pm 0.016\mu\text{K}$ on the direction $(l, b) = (264 \pm 0.3, 48.4 \pm 0.5)^\circ$. After calibration with the rotation of the Earth around the Sun which moves around the Galaxy center, which moves around the center of the Local Group of Galaxies[45], the speed of the center of the Local Group in respect of the CMB is $627 \pm 22\text{kms}^{-1}$ towards $(l, b) = (276 \pm 3, 30 \pm 3)^\circ$.

Therefore, current experiments are dedicated to study the small anisotropies of the CMB, or in other words temperature fluctuations. As discussed before CMB contains information on the early universe. This statistical information is related to the matter power spectrum projected

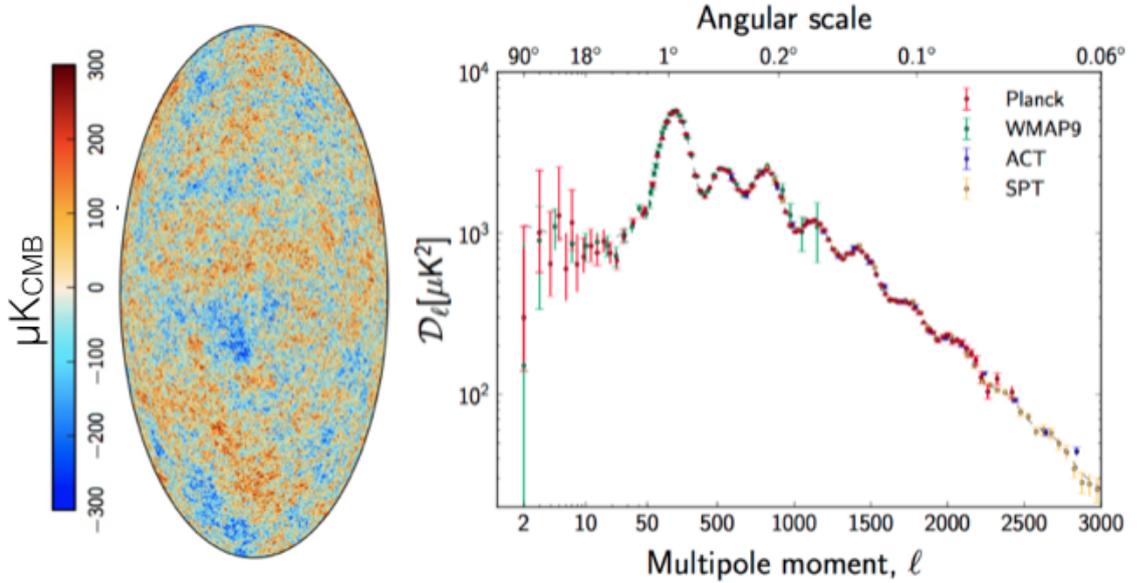


FIGURE 1.11: *Left*: CMB intensity map showing the temperature fluctuations as observed by PLANCK Satellite[42]. *Right*: Measured angular power spectra of a compilation of experiments denoted with different colours [43]. [See text for details]

in the celestial sphere. Therefore, for observational purposes, the temperature fluctuations, $\frac{\delta T}{T}(\theta, \phi)$, are decomposed in spherical harmonics, $Y_{lm}(\theta, \phi)$, where l is the multipole moment and m is the phase. The multipole moment is related to the angular size on the sky, $\Delta\theta$, through $l \simeq \pi/\Delta\theta$. From basic statistics we know that for a Gaussian random field the average and variance is enough to study its statistical properties. In the case of a primordial isotropic field $\langle a_{lm} \rangle = 0$, while the average variance is defined through $C_l = \langle |a_{lm}|^2 \rangle$, for each spherical harmonic l . This angular power spectrum, C_l , is the discrete, spherical analog of the 3D power spectrum, $P(\vec{k})$, for a given mode, \vec{k} , in Fourier space.

In the right part of Fig. 1.11, the maximum posterior CMB intensity map in Mollweide Projection[46] at $5'$ angular resolution, in $K_{\text{cmb}} = 2.725K$ physical units, as it is derived from the joint baseline analysis of the data of Planck satellite[42], WMAP (Wilkinson Microwave Anisotropy probe) [47, 48], and 408 MHz Map [49] as was compiled by Planck Team[42]. The small strip of the Galactic plane, 1.6% of the sky, is filled in by a constrained realisation that has the same statistical properties as the rest of the sky. Furthermore, this map shows that the isotropic behaviour of cosmic structures is of the order of 10^{-5} , in the early universe. In the right panel of Fig. 1.11, the re-scaled temperature angular power spectrum, $\mathcal{D}_l = l(l+1)C_l/2\pi$, is displayed as measured by a compilation of the current most precise experiments Planck Satellite, WMAP, ACT(Atacama Cosmology Telescope[50]) and SPT(South Pole Experiment [51, 52]). The former two experiments are satellites and are observing large portions of the sky nearly 60% acquiring high precision on the largest possible scales (small modes l) of the measurements, avoiding earth's atmospheric light pollution. The latter two experiments are ground base telescopes aiming the small angular scales (high l). At that time up to nine peaks are observed on the power spectrum due to baryon acoustic oscillations, allowing a precise measurement of the angular acoustic scale, at $\theta_* = 0.596724 \pm 0.00038 \text{ deg}$. This measurement was validated later with

the large scale structure surveys detecting the Baryon Acoustic Oscillation peak position[53, 54], at about $r_s \simeq 100h^{-1}\text{Mpc}$.

This observable directly show the different modes of the ripples of the matter density field at 380000 yrs after the planck time. One can observe the remarkable agreement between experimental results and best fit- ΛCDM -model.

1.3 Gravitational Waves

It is always fruitful to imagine what the outcomes could be from new observations. Since this thesis is written 101 years after the most amazing prediction of physics and its first detection, i.e the gravitational waves (GWs hereafter), it is worth devoting a small section with the possible impact to our study. Since, we live in the era of multi-messenger astronomy is hard but not to stargaze, how our current picture of our universe is going to be more revealed or modified by the new observations of GWs which are waves that *propagated through the spacetime continuum*. Although, GWs are not yet fully understood from current observations, for example there is still high inaccuracy in their location on the sky, they open new potentials on spacetime observations. Currently, the range of the observed mass of the binary black holes[55] suggested the existence of a spatial distribution of Primordial Black Holes that could trace the cosmic fluid all over the universe, since primordial black holes are generated at the center of supermassive galaxies and quasars. Additionally, some authors argued[56] for a very challenging observation of the very small modification of the large scale structure clustering due to GWs. The detection of the GWs is summarised in appendix [A.10](#).

Chapter 2

Cosmological Principle

In this chapter, we are briefly discussing the Cosmological Principle: from the observational, historical and mathematical points of view. We introduce the different aspects of distributions and discuss its possible implications of this principle. Finally, we discuss the challenges of this study and our approach.

2.1 Observational and Historical views

In cosmology, as discussed in the previous chapter, one of the basic pillars of our understanding of the universe is the *Cosmological Principle*. This principle states that the universe should be smooth, i.e. homogeneous and isotropic, on the largest scales. Recall that this basic assumption is used to derive every tool that we use in cosmology; i.e. the FLRW metric, the Friedmann equations as long as Einstein-Boltzmann equations. One may wonder, though: "*Do we have enough evidence for that?*", "*What are the observables and the tools for this study out there?*" or "*Why make this assumption to begin with?*". Well, we have evidences, but before discussing those, lets discuss some of the difficulties of the "*what*". Although, it is easy to assert a test to study how isotropic is the universe using the earth as the observational point, it is difficult to investigate the homogeneous behaviour of the universe. The reason for that is simple. We do not have the luxury to look from everywhere the universe. If only we could start travelling along three orthogonal directions in the universe we would be able to measure the differences of the positions of the galaxies in order to assert a test for homogeneity. Thus, we rely on observations from ground based or satellites telescopes that turn around the earth, known as observatories.

Now we are able to answer the why-question. As it was discussed in the previous chapter [1.2](#), having a homogeneous and isotropic universe provides us with a less simple picture of our universe. However, one can contradict the latter statement and suggest a simple universe in which the Milky Way (our galaxy) is the center and the rest of the universe is only isotropically distributed around it. By constructing such a picture for our universe, we end up in another simple model for our universe. However, this model has not many convincing evidence so far. Nonetheless, cosmologist were lead to the interest of cosmography of our universe, to distinguish

between these two possible pictures. Cosmography is basically the analogous of topography for the universe, i.e. one measures the location of the objects in order to understand the whole configuration. Large scale structure studies are mostly cosmography.

There are several observational evidences about isotropy, one of the two parts of Cosmological Principle. Currently, large spectroscopic galaxy surveys, show no evidence for anisotropies in the projected galaxy distribution in volumes of a few Gpc³ [57]. At $z \simeq 1100$, Cosmic microwave Background temperature anisotropies correspond to density fluctuations in the young universe of the order of 10^{-5} [58], as discussed in the previous chapter. Furthermore, isotropy tests, using the luminosity function of galaxy survey [59], have shown no statistical deviations from isotropy. In the more recent Universe, at $z \sim 2.6$, analysing the structures of the Lyman-alpha Forests [60], statistical isotropy was estimated for the first time in low precision. Additionally, the distribution of sources in X-ray emission of baryonic matter structures [61] at $z \sim 0.02$ and surveys of radio sources [62] at $z \sim 1$ strongly supports the isotropic behaviour of our universe.

Historically, going towards the past of cosmological principle construction, it was first Copernicus observing the night sky realised that the Earth rotate around the most luminous celestial object on our day sky, the *sun* which is the center of the solar system, during the 14th century. This led him to assume that we live in place which is not privilege. This was named as the *Copernican Principle*. Later on, in 1687, Cosmological Principle (CP hereafter) was asserted as a philosophical statement by Newton in his book with title "*Philosophiae Naturalis Principia Mathematica*" [63]. Newton was the first to realise that there is a deviation of the gravity center and the position of the sun. Therefore he generalised the notion of the center of the universe to any center which is at rest. This made him conclude that the earth goes around the sun and then sun moves within an infinite, homogeneous and isotropic universe. Going forward in time, in 1923, the CP was, mathematically, asserted into the Einstein Field Equations by Friedmann, to obtain a simplistic solution, thus making easier to construct cosmological models as we described in the previous chapter and established the current standard cosmological model, Λ CDM-model.

However, we should stress that, combining isotropy and the Copernican principle, which states that our position in the Universe is not privileged, does not strictly implies that the Universe is homogeneous [64–67]. As shown by Durrer et al. [68], this implication is not true if we only have projected isotropy. This means that the *projected* isotropy is a weaker assumption than *spatial* isotropy. In order to distinguish the two notions, i.e. spatial and projected isotropy, lets consider the following. A spatial isotropic 3D point-distribution can be described by the equality, $\rho(r, \theta_1, \phi_1) = \rho(r, \theta_2, \phi_2)$ for every $r, \theta_1, \phi_1, \theta_2, \phi_2 \in \mathbb{R}^5$. This information is different from the projected isotropy in which we integrated all the radial information, i.e. $\rho_{proj}(\theta_1, \phi_1) = \rho_{proj}(\theta_2, \phi_2)$ for every $\theta_1, \phi_1, \theta_2, \phi_2 \in \mathbb{R}^4$, where $\rho_{proj}(\theta, \phi) = \int \rho(r, \theta, \phi) dr$. Thus observational isotropy evidences, such as CMB isotropy, for instance, at a given redshift or after projection over a given range of redshifts are not enough to ensure a homogeneous universe.

Additionally, an observer, using a 3D spatial distribution through a spectroscopic survey, can only probe the statistical properties of the universe only on the observer past light-cone and not inside of it. Thus, the observed galaxies are not at the same epoch. An attempt to overcome this limitation was to use the star formation history in order to probe the homogeneity of structures inside the past light-cone [69], but this is model dependent through strong astrophysical

assumptions. A team[70] introduced another solution for probing the homogeneity by using a combination of secondary CMB probes, including integrated Sachs-Wolfe, kinetic Sunyaev-Zeldovich and Rees-Sciama effects to overcome this limitation. Nonetheless, no observational evidence of this effect exists to date.

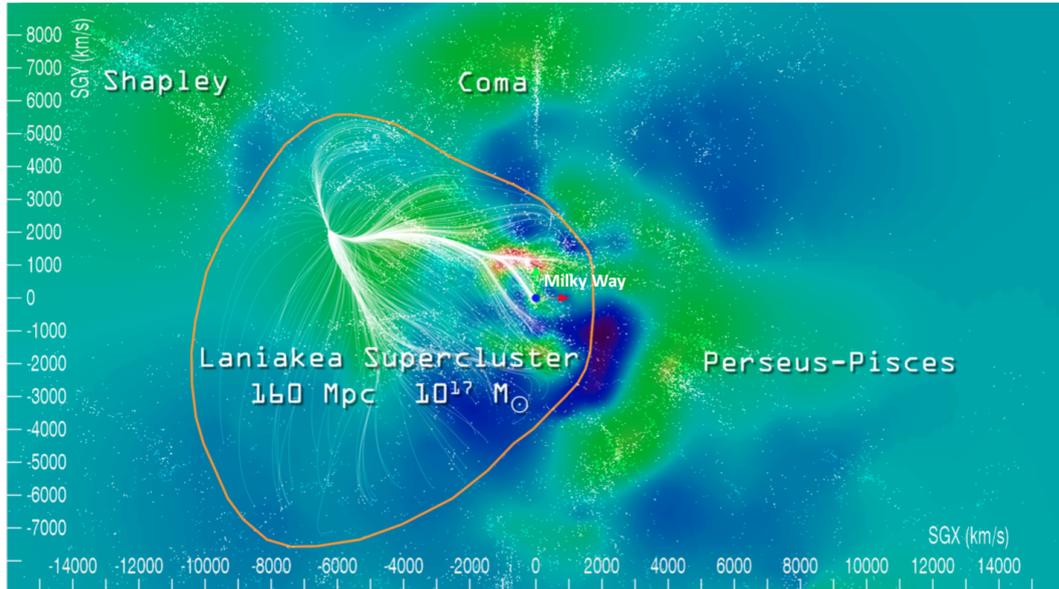


FIGURE 2.1: *Left:* The *Lianakea* structure in the equatorial plane, identified by a velocity Wiener filter algorithm with a compilation of (z, θ, ϕ) coordinates of galaxies[71]. Shaded contours represent density values within the equatorial slice with red at high densities and blue in voids. Individual galaxies from a redshift catalog are given as white dots. Velocity flow streams within the Laniakea basin of attraction are shown in white. The orange contour encloses the outer limits of these streams.

Moreover, in the galaxy distribution, there are special patterns which we need to identify before studying the homogeneity scale. There is a plethora of studies in the literature for identification of such large scale structures. In 1983, Clowes [72] performed a Power Spectrum Analysis to identify a *Large Quasar Group* of $\sim 100\text{Mpc}$ size at $z \sim 0.37$. Later on, Tully [73] used a percolation analysis identifying the 350Mpc *Pisces-Cetus Supercluster Complex* up to $z \sim 0.1$. Two years later, Geller and Huchra [74] using a spectroscopic survey, identified a $\sim 240\text{Mpc}$ structure, naming it as the *Great Wall*. In 2001, Klypin and Kopylov [75], by using the Kormendy relation as a distance indicator, i.e. taking into account galaxy luminosities, they discovered a Giant Void with size $\sim 300\text{Mpc}$ with $h=0.5$. Two years later, Gott III et al. [76] used a spectroscopic survey identified the 420Mpc *Sloan Great Wall*. Moving fast forward in time, in 2006, a large scale filamentary structure was identified around the protocluster at $z=3.1$ with a size of 65Mpc [77]. Close to the CMB Cold spot Rudnick et al. [78] identified a 140 Super Void structure with NVSS sources smoothed to a few degrees. In 2012, Clowes et al. [79] found a filamentary structure in the early universe, $z \sim 1.3$ with a characteristic size of 500 Mpc. At the same epoch, Horvath et al. [80] using a Kolmogorov-Smirnov test[81] on the angular positions of a γ -Ray Burst distribution, identified the *Hercules-Corona Borealis* with a tremendous size of 2.2Gpc size. When this thesis was started, Tully et al. [71] used a velocities wiener filter to identify $160h^{-1}\text{Mpc}$ structure that they named *Lianakea*, due to its shape as a giant fluid in space as illustrated in Fig. 2.1. Finally,

last year, a $271h^{-1}\text{Mpc}$ massive supercluster system[82] was identified in the Sloan Digital Sky Survey at $z \sim 0.47$.

Some of the aforementioned studies with structure of sizes larger than $270h^{-1}\text{Mpc}$, obviously do not contradict the measured homogeneity scales with lower values, since homogeneity is a statistical property of the universe. Some studies found a transition to homogeneity scale between 70 and $150 h^{-1}\text{Mpc}$ [83–93]. However, some others have not found such a transition. Large scale structure studies[94–98] using observables, similar to our study, obtaining evidence for a fractal distribution of galaxies. However, the galaxy redshift surveys used in the aforementioned studies are too shallow, sparse, or have survey geometries too complicated, to give conclusive results for a fractal universe.

Another difficulty arises from the fact that those measurements cannot go beyond the observable universe which is defined by the comoving shell of $d_C(z \sim 1100) \simeq 49\text{Gly}$. However, the validity of Copernican Principle or the Cosmological Principle outside the observable universe, is currently impossible to test. Therefore, we always rely in the assumption that the universe remains the same even outside the observable horizon, *particle horizon*.

2.2 Mathematical definition

Let's break down mathematically the Cosmological Principle. The Cosmological Principle is comprised by two assumptions as said before. A statistically homogeneous universe at large scales. A statistically isotropic universe at large scales. In other words, homogeneous universe relies on the translational invariance of the properties of the universe while isotropy rely on the rotational invariance of those properties. However these kind of tests are impossible to perform.

An isotropic distribution is, mathematically, defined as the distribution whose properties, defined by a mathematical function f , are rotationally invariant:

$$f(x') = f(Rx) \quad (2.1)$$

where $f(x)$ quantifies the function of the property in study, observed in a coordinate system with coordinates x , R is the rotational matrix and x' is the coordinates on the rotated system. Furthermore, a homogeneous distribution is the distribution whose properties are translationally invariant:

$$f(x') = f(x + \psi) \quad (2.2)$$

where ψ is linear displacement field. Generally, though, we can have 4 distinct different classes of distributions:

- Class 1: Homogeneous and Isotropic 3Dimensional distributed properties.
- Class 2: Non homogeneous nor isotropic 3Dimensional distributed properties.
- Class 3: Non homogeneous but isotropic 3Dimensional distributed properties.
- Class 4: A homogeneous but non isotropic 3Dimensional distributed properties.

However, in this thesis, we study point-distributions since we are dealing with galaxies. These kind of distributions are a subclass of the previous classification, since the last class 4 is not valid for infinitely large point distributions. By construction a 3 dimensional homogeneous point-distribution is uniform in space. This means that the average density is constant with translations in every direction. Thus if one selects one point in space of a 3D homogeneous distribution and observe around, he will notice that the average number density does not change with the different rotations.

The class 1 is illustrated by the left panel of Fig. 2.2, a two dimensional distribution of galaxies that are treated as point sources. One can observe that the average number of galaxies does not change under any translation or rotations of the observer. The 3rd class, isotropic but inhomogeneous, is illustrated by a two dimension spatial distribution of galaxies, treated again as point sources. The average number of galaxies do not change if we rotate the observer as one can observe in the right panel of fig Fig. 2.2. However, one notices that the inside shell has more galaxies per degree than the outer shell, which reflects the radial dependence of the point distribution.

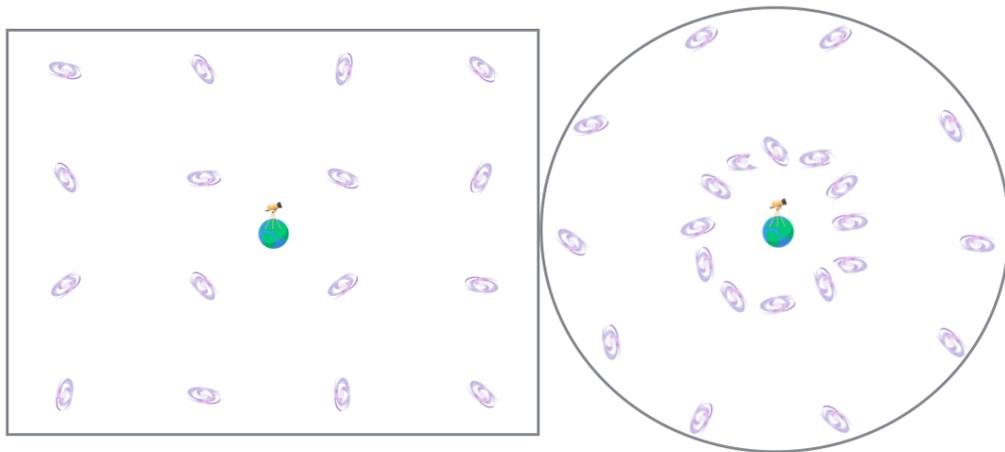


FIGURE 2.2: *Left*: 2D representation of homogeneous (and isotropic) galaxy distribution
Right: 2D representation of an isotropic (but not homogeneous) galaxy distribution [See text for explanation][Credit on [99]]

However, there are potentially various properties of a galaxy distribution that can be use to study homogeneity and isotropy of the universe. A short list is summarised below:

- spatial number density of a galaxy
- average magnitude (color) of a galaxy
- average spin of a galaxy
- average size of a galaxy
- average shape of a galaxy
- Star Formation Rate of a galaxy

- Hubble rate of a galaxy
- e.t.c.

Currently, in observational cosmology, it is most common to study the spatial number density for isotropic and homogeneity tests, as long as quantities related to this property such as the fractal correlation dimension and the power spectrum. Another familiar, approach is to test the Hubble rate[100] against the radial and angular dependence.

2.3 Implications of Smoothing

Criticism on the assumption of the Cosmological Principle has been particularly active in the recent years for alternative explanations of the accelerating expansion of the universe. Models relaxing the homogeneity assumption can get rid of the need of Dark Energy which is an exotic form of fluid. A particular example is to introduce an alternative model, in which the Milky way is at the center of the universe and beyond a large amount of void, a huge clump of matter exists that attracts the celestial objects between our Milky way and the large clump of matter, i.e. the acceleration that we observe with the supernovae[57]. This will results in an image of a universe in which Dark Energy is not needed and the apparent acceleration is due to the large void between the milky way and the large clump. However, as Caldwell and Stebbins [101] have shown those models should produce large amount of spectral distortions on the observed Black Body spectrum of CMB. But those distortions are not yet detected despite the high precision measurements, to date.

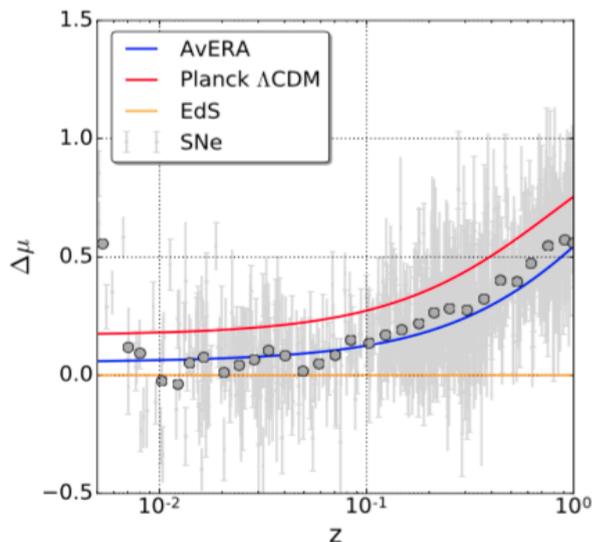


FIGURE 2.3: Relative Distance module measurement as a function of redshift of a compilation of Supernova, fitting separate isotropic universe (AvERA) better than homogeneous and isotropic universe (Planck best-fit model).[See text for details]

Furthermore, Kovács and García-Bellido [102] have found strong evidence for overdensities and underdensities in the cosmic web. In particular they have found evidence for an alignment between CMB Cold Spot, located in $(l, b) = (207.8^\circ, -56.3^\circ)$ galactic coordinates, and the Eridanus

supervoid and in the antipode observational part with the overdensity structures such as Hercules, Corona Borealis and the Great Pillar, indicated by a continuous blue line. They have also found an alignment between the Supervoid Draco and LEP structure in the cosmic web as plotted with a blue dash line in Fig. 2.4. However, these peculiar baryonic structures are only a small fraction of the observed volume to date, $\sim 26h^{-3}\text{Gpc}^3$ [103]. Therefore their existence is statistically insignificant against the statistical homogeneity that we observe on large scale structure volumes.

Another recent study by Racz et al. [104] have revealed evidences for the conjecture of separate isotropic expanding universes without dark energy. They have used the relative distance module as a function of redshift at $z \leq 1$ to discriminate between the AvERA (Average Expansion Rate approximation) model, which is in accordance with the aforementioned conjecture, and Λ CDM-model which assumes a homogeneous and isotropic universe. They have excluded Einstein de Sitter universe, $(\Omega_M, \Omega_\Lambda) = (1, 0)$. They have also found that AvERA matches better the observational data of SN than best fit Λ CDM-model of PLANCK satellite measurement as shown in Fig. 2.3. However, this apparent disagreement might be consistent with the known tension between the Planck H_0 measurement[6] and the Riess et al. [28] distance ladder study. Therefore, a more cautious and thorough investigation is needed on this study.

Thus, for the aforementioned reasons, we consider homogeneous but anisotropic models or isotropic and inhomogeneous models. A particular toy model for the latter class was described by the Lemaitre(1933)-Tolman(1934)-Bondi(1947) model. This model considers the spherically symmetric dust solution of Eq. 1.1, which is modelled by the metric:

$$ds_{LTB}^2 = -c^2 dt^2 + \frac{[\partial R(t,r)/\partial r]}{1 + E(r)} dr^2 + R^2(t,r) d\Omega \quad (2.3)$$

where $R(t,r)$ and $E(t,r)$ are functions that depend in time and radial direction of the observer to allow for an isotropic universe.

The homogeneous but anisotropic models are a class of Bianchi Cosmological Models. In the simplest case, they are modelled by a spherically anisotropic solution of Eq. 1.1:

$$ds_{Bianchi,I}^2 = -c^2 dt^2 + a_X^2(t) dx^2 + a_Y^2(t) dy^2 + a_Z^2(t) dz^2 \quad (2.4)$$

where there are 3 different time-dependent scale factors, $a_X(t), a_Y(t), a_Z(t)$. However, recently, these models were strongly disfavoured by CMB observations, where strong constrains were obtained for the anisotropic expansion of the universe parameter, $s_V/H_0 < 4.7 \times 10^{-11}$ at 95%*C.L.* [105].

Furthermore, another way of relaxing the homogeneity assumption was introduced by Schwarz [106], Buchert [107] and it is called the *backreaction* mechanism. In order to explain this mechanism consider the following. Assuming that GR holds, the motivation behind a *backreaction* mechanism comes from the non-commutation of the temporal evolution and spatial averaging in inhomogeneous space-time. To explain this consider that in the usual Λ CDM-model one solves Eq. 1.1 when the perturbed FLRW metric (Eq. 1.22) is firstly smoothed at first order, $g_{\mu\nu} \rightarrow \langle g_{\mu\nu} \rangle$ and then it evolves through Einstein Field Equations. This is formulated as,

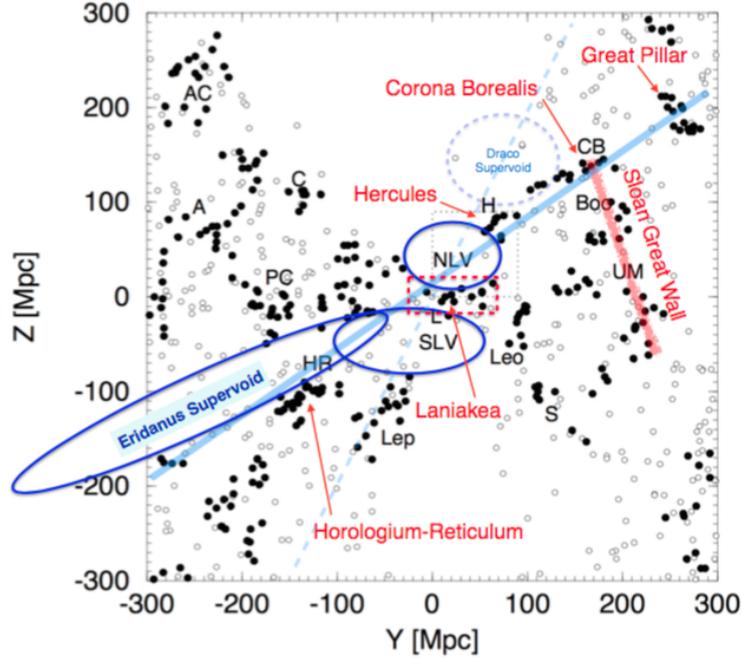


FIGURE 2.4: An alignment of the overdensity and underdensity spots in the cosmic web as discovered by a compilation of supercluster catalogues. [Image taken from [102]]

$\tilde{G}_{\mu\nu}(\langle g_{\mu\nu} \rangle)$, where $\tilde{G}_{\mu\nu} = G_{\mu\nu} - \frac{8\pi G}{c^4} T_{\mu\nu}$. Then one needs to add an additional term, $\Lambda g_{\mu\nu}$ to fit the observational data, that is come naturally from the gauge freedom of those equations. However, in the second case, *backreaction*, the perturbed metric evolves with EFE. Then the evolved EFE equations can be smoothed to first order giving rise to this additional term. This is described mathematically as, $\langle G(g_{\mu\nu}) \rangle$. Therefore, now one can observe the difference of those equation $\tilde{G}_{\mu\nu}(\langle g_{\mu\nu} \rangle) \neq \langle \tilde{G}_{\mu\nu}(g_{\mu\nu}) \rangle$. This leads to two different metric solutions, the standard homogeneous and isotropic pertrubed FLRW-metric, which we discussed in chapter 1.2.1, and the *backreaction* one. This is how backreaction mechanism could introduce a new way to understand the acceleration of the universe, by substituting the Dark Energy with an underlying mechanism of evolution of structures through an inhomogeneous background metric.

The *backreaction* mechanism can be observationally apparent in the following example. Let's consider that the ideal measurement of the Hubble constant proceeds as considered by Li [108]. Considering N perfect standard candles [109] that located within a physical volume $V(t)$ homogeneously. In this idealised case, perfect standard candles are those candles, i.e. very low redshift candles $z \ll 1$, that allow for the determination of the expansion rate today, H_0 . Then their physical distances, d_i , are measured via relative magnitude measurements and their velocities by the redshift information, $v_i = cz_i$, as shown by Ruiz-Lapuente[110]. Then to obtain the Hubble constant one takes the average:

$$H_0 \equiv \frac{1}{N} \sum_{i=1}^N \frac{v_i}{d_i} \quad (2.5)$$

In the limit of a large sample, $N \rightarrow \infty$, this turns into a volume average:

$$H_0 \equiv \frac{1}{V} \int \frac{v(\vec{r})}{d(\vec{r})} d^3\vec{r} \quad (2.6)$$

If we neglect the effect of the lightcone, which is true for $z \ll 1$, the spatial average is a good approximation for the average over the past light-cone, since the expansion rate of the Universe changes insignificantly at timescales shorter than the Hubble time, $t_H = H_0^{-1}$. However, in standard cosmology, the expansion rate is defined as $H_0^{\Lambda CDM} = \frac{\dot{a}(t)}{a(t)}$. The issue is to establish the connection between H_0 and $H_0^{\Lambda CDM}$. In the standard model, they agree by construction if the volume $V(t)$ becomes large enough. However, due to the nonlinearity of the Einstein equations at smaller scales, overdensities affect the evolution of the averaged values, such as H_0 . There is still a significant debate to the question of the backreaction mechanism, with the large fraction of the community convinced that the effect is negligible[111].

2.4 Fractal Universes?

Nature is full of fractal patterns, be it from amazing snowflakes to romanesco broccoli. Naturally due to the peculiar structure of galaxies around us, there were several claims that the universe behaving like a fractal one[97, 98]. In 1996, Rassem and Ahmed [112] have described the dynamics of a fractal universe. Introduced in 1982 by Mandelbrot, fractals are mathematical objects whose structure is invariant of scale. However, there is no strict, or unique, definition but it relies on unconventional views of the scaling and dimension notions. Traditional notions of geometry dictate that the shape scales predictable according to intuitive and familiar ideas about space they are contained within. Here, we consider that a fractal F satisfies the following properties:

- F has a finite structure, i.e. the distribution is structured to arbitrarily small scales.
- F cannot be defined with usual geometric and analytical tools due to its irregular pattern.
- F has one or more repetitive patterns, with different proportions each.
- F is usually described by a recursive process.

A famous fractal example is the Koch snowflake (Kock curve) as shown in Fig. 2.5. The construction of the Koch curve starts with the design of an equilateral triangle. Then recursive change each line segment following 3 simple rules:

1. Each line segment is divided into three segments of equal length.
2. An equilateral triangle is drawn in the middle segment from step 1 as its base and points outward.
3. The line segment that is the base of each triangle from step 2 is removed.

After one iteration of this process, the resulting shape is the outline of a hexagram. After four iterations a snowflake is created. The Koch snowflake is the limit approached as the above steps are followed over and over again.

One of the simplest mathematical definitions will make the above definition more clear. The box counting method consists of filling the space with boxes of size δ around the 3 dimensional

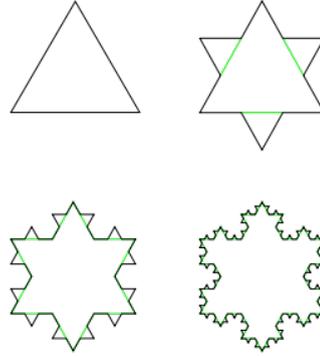


FIGURE 2.5: The first four iterations of the Koch snowflake.[Image taken from [113]]

fractal object and then count the number of only those boxes that cover the space of the fractal. For a line fractal element with size l contained in a 3 dimensional space we find $N(\delta) \simeq l/\delta$ and for surface fractal element with surface A we obtain $N(\delta) \simeq A/\delta^2$. Mathematically, the box-counting dimension is defined as:

$$D_0 = \lim_{\delta \rightarrow 0} \frac{\log [N(\delta)]}{\log \delta^{-1}} \quad (2.7)$$

where the number of boxes $N(\delta)$ containing the 3dimensional fractal object. This is also the definition of the Minkowski Dimension. If this limit does not exist, we define the superior and inferior limits instead. This definition of fractal dimension can be generalized in the Generalised Fractal Dimension, D_q , which we discuss in chapter A.9. However, in our study, we are interested in one special case of the generalized fractal dimension, the fractal correlation dimension, D_2 . This special case applies for a 3 dimensional point distribution, which is the objective of this thesis and we discussed its application in chapter 4.1.1.

2.5 Complications of Celestial 3Dimensional Distributions

The statistical tools that we have in our disposal to study the Cosmological Principle are the baryonic matter distribution, which is structured as galaxies and quasars in the late universe. However there are several subtleties when we want to make inferences for the universe since it consist of not only luminous matter but also Dark Matter. Therefore galaxies and quasars are treated only as tracers of the total matter of the universe and we investigate the possible implications and effect on the latter from Dark Matter to make global statistical inferences. Thus, we study the power spectrum and correlation function as long as number densities of traces of the total matter density of the universe. These subtleties that we will discuss are relevant for every large scale structure analysis.

2.5.1 Biased Tracers

Λ CDM-model predicts the power spectrum of the total matter distribution of the universe. However, observationally we have access to the statistical information of the distribution of the tracers such as galaxies, quasars Lyman- α forest or other. Galaxies and quasars are the most massive objects, which they formed in the densest regions of the universe, resulting in an increase of the observed density perturbations. Therefore the power spectrum is retrospectively amplified by a factor b^2 :

$$P_{tr}(k) = b^2 P_m(k) \quad (2.8)$$

where $P_m(k)$ is the matter power spectrum and $P_{tr}(k)$ is the power spectrum of the tracer we consider. The name of this factor, b is called the *cosmic bias*, or *bias* for short. The same amplification adjust also the amplitude of the 2ptCF, $\xi(r)$ which is only the Fourier transform of the power spectrum. Note that this factor, b^2 , depends on the tracer that we are considering to measure, $\xi(r)$ and $P(k)$, therefore it varies with *redshift*, the *magnitude* or the *mass* of the tracer.

2.5.2 Redshift Space Distortions

We show in chapter 1.2.4 that the *redshift* of an object allows us to identify its position along the line of sight. However, the relation between *redshift* and position is not exact for every case since the object has also a different velocity component of the one coming the hubble expansion. The *redshift* measured for an object with velocity in the line of sight of the observer has higher value than the redshift of an object without a velocity. Therefore, when we reconstruct the position along the line of sight, the object with velocity seems further than it is. Conversely, an object, whose velocity direction is towards to the observer in opposite direction of the hubble flow, seems closer. These effects are called *Redshift Space Distortions* (RSD). Due to the gravitational interaction, the velocity field of the tracers is correlated with the gradient of the density of matter. This correlation modified the power spectrum and the 2ptCF of the tracer.

Schematic representation of the effect of distortions in redshifts space. At the top, the points represent the positions of the tracers in the real space, and the arrows their velocities. The figures at the bottom show the reconstructed positions in the redshifts space, in the case where the observer is at the bottom of the image.

On large scales ($r_{ls} > 10h^{-1}\text{Mpc}$), tracers tend to move towards the overdensities of Dark Matter due to the huge Potential as shown in Fig. 2.6. Tracers which falls towards an overdense region situated between the observer and this tracer thus has lower removal speed that if it followed only the hubble flow. Therefore, when reconstructing the radial position from the tracer's redshift, the tracers appears closer to the center of the overdensity. The same reasoning can be performed if the positions of the tracer and the overdensity are reversed. Since it falls towards the overdensity, the radial velocity of the tracer appears more important to the observer that if the tracer was taken only in the Hubble flow: it also seems closer of the tracers around the overdensities along the line-of-sight, which then appear denser in the space of the redshifts than in the real space.

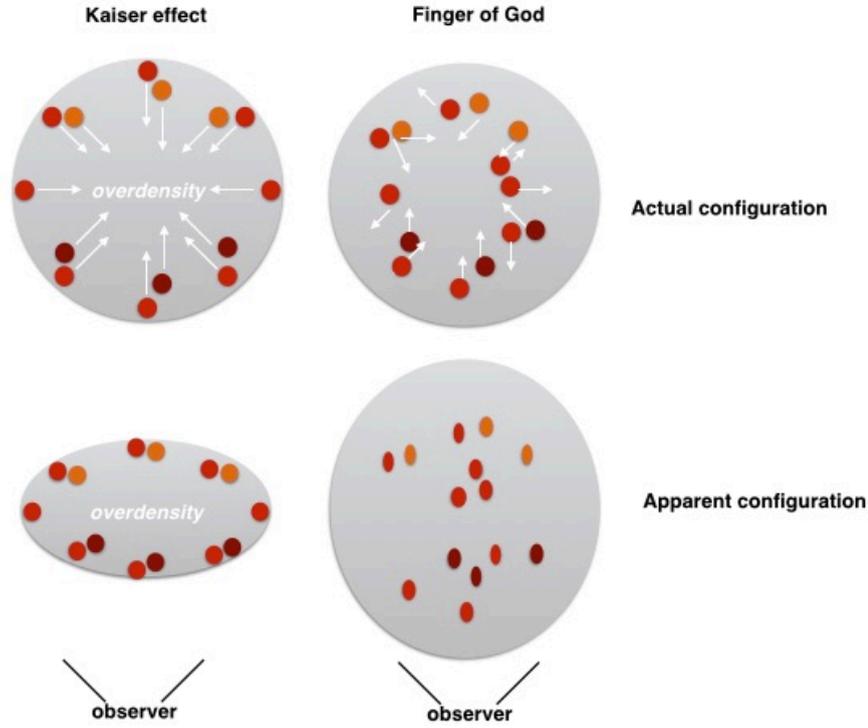


FIGURE 2.6: Schematic illustration of the *Redshift Space Distortion* effect. In the top the figures represent the actual position of the celestial objects and their velocities. On the left the motions are due to gravitational potential of the overdensity and on the right due to gravitational interaction among the objects. Bottom figures represent the apparent position of the objects and the apparent shape of the overdensity. [See text for details]

Consequently, this effect, increases the signal to noise ratio of clustering and therefore amplifies the 2ptCF and Power Spectrum.

Kaiser [114] has shown that in first order, this effect is only a multiplication factor to $P(k)$ depending only in the direction of the underlying modes in respect of the line of sight:

$$P_{tr}(k, \mu) = b^2(1 + \beta\mu_k^2)^2 P_m(k), \quad \mu_k = \cos(\vec{k}, \vec{k}_\perp), \quad (2.9)$$

with

$$\beta = \frac{f}{b}, \quad f = \frac{d \ln D(a)}{d \ln a} \simeq \Omega_m^\gamma(z) \quad (2.10)$$

where the last equality is a valid parametrization of the rate of growth of structures, f and $\gamma = 0.55$ for the unmodified General Relativity Theory[115]. The factor, β , quantifies the RSD phenomenon. Taking the fourier transform and averaging over angles[116] we end up:

$$\xi_{tr}(r) = b^2 \left(1 + \frac{2\beta}{5} + \frac{\beta^2}{5}\right)^2 \xi_m(r). \quad (2.11)$$

At smaller scales $r_{ss} \lesssim 10h^{-1}\text{Mpc}$, the non-linear effect are dominant, define the non-linear regime. The tracers are not moving coherently in the underdensities. Since, we reconstruct the positions of the tracers, their velocity dispersion in general is affected in a marginal manner by the underdensity along the line of sight as shown in Fig. 2.6. Consequently, there is an

ajustement of the amplitude of $P(k)$ and $\xi(r)$. From the point of view of the observer, this gives the impression that the distribution of the tracer "point" towards him: and therefore this effects is named "*Finger of God*" (FoG hereafter). This can be modelled empirically, by introducing a term of the dispersion of the peculiar velocities σ_p in the power spectrum. We introduce this effect with a Gaussian form as:

$$\ln D_G(k, \mu; \sigma_p) = -\frac{1}{2} \left(k \frac{\sigma_p}{H_0} \mu \right)^2 \quad (2.12)$$

We introduce the above form of the *Damping model* in the $P_{tr}(k)$ as:

$$P_{tr}(k, \mu) = b^2 (1 + \beta \mu_k^2)^2 D_G(k, \mu; \sigma_p) P_m(k), \quad \mu_k = \cos(\vec{k}, \vec{k}_\perp), \quad (2.13)$$

Thus we can transform the measured power spectrum for the tracers into the redshift space in the matter power spectrum in the real space. Notice that this is only an overall amplitude on the spherically symmetric power spectrum. The Fourier transform of this power spectrum will serve as a model for all the observations of the following chapters.

At even smaller scales, $r_{ess} \lesssim 1h^{-1}\text{Mpc}$, the universe becomes more complex. The evolution of those scales are not explained with the linear theory, and several approaches exist that try to phenomenologically treat these problem. These scales are treated with computationally heavy N-body simulations or perturbation theory. However, those scales are beyond the scope of this thesis, since they do not affect our large scale measurements.

2.6 Current Unanswered questions of ΛCDM -model

Although our current ΛCDM -model is support by many observations still has some unexplained issues. These are summarised below:

- *Non-linear regime*, i.e. at small scales $r \simeq 10 - 20h^{-1}\text{Mpc}$ where gravity is strongly non-linear and requires using complex N-body and hydrodynamical numerical simulations.
- *Nature/existence of inflation*, is it an artifact that is used to explain the smooth flat expansion of small inhomogeneities and the horizon problem or a different explanation underlies?
- *Neutrino mass measurement*. Non-linear dynamics of the universe at small scales might be affected but the neutrini budget of the universe[117].
- *MOdified Newtonian Dynamics (MOND)*. At small scale, structure formation models can be built based upon both ΛCDM and MOND paradigms, however the latter fits better the observed data of the Tully-fisher relation[118].
- *Matter/Antimatter Asymmetry*, which is expressed by the low annihilation rate of the relic density ratio between baryons and photons $n_b/n_\gamma \sim 10^{-18}$, has many possible explanations.
- *Essence of Dark Matter*, as discussed in chapter 1.2.2.3, is an inflaton particle or a super-symmetric particle or other?

- *Nature of Dark Energy*, as discussed in chapter 2.3, it is a cosmic negative pressured fluid or a Backreaction mechanism of the treatment of the evolution of the metric or a different Gravitational Theory modification underlies?

2.7 Approaching the Challenges

The basic challenges one needs to overcome to study the Cosmological Principle are the following. Which celestial objects one needs to use for this study and how to define them. Which are the possible observables. What are the possible estimators and which of them are the optimal ones. Which is the definition of the homogeneity scale in practice. Which theoretical predictions are compatible with its measurement. To which extend the objects in study are enough to trace the universe as a whole. Are there any missing matter(dark matter) that one needs to take into account. Which is the ranging of scales one needs to look for a homogeneous universe. How, we are going to define the 3D comoving position of galaxies in the celestial sphere, in practice. Are there any alternatives to the cosmological principle, which can be identified from the data point of view. What are the possible implication of this measurement on the different theoretical models of the universe. Can we use this measurement to distinguish between alternative models against the concordance one, i.e. Λ CDM-model?

As discussed before, there is a plethora of observables in the literature that are suited in study of the homogeneity scale. All those observables are all obviously related to the number density of objects on the celestial sky. Those observables, obviously need to study the 3 dimensional spatial space of the objects. The possible objects that one needs to study are the largest possible identified objects. These are the collection of stars that we call galaxies or the cluster of galaxies or so on. Therefore, one start looking for scales beyond ten-ths of kpc which is the typical sizes of individual galaxies. However, since galaxies are objects that only trace the electromagnetic, optical spectrum of matter in the universe, we need to take into account this missing matter that we cannot observe. Thus, we include in our analysis the additional clustering of the total matter, electromagnetic and non electromagnetic interacting matter.

Furthermore, in practice, to perform the above measurement, i.e. counting of the average number density, we treat galaxies as point particle in the celestial sphere, and we need their radial and angular position. We obtain the angular position by imaging photometric(or astrometry) based on ground based telescopes. The radial position is more complicated to obtain. Since we do not have a long ruler to measure galaxies, we rely on their luminosities. Therefore, we use spectroscopic instruments that acquire the total flux of the galaxies as a function of wavelength. This permit us to compare their rest frame spectrum, which is the template of the standard emission lines on different chemical elements composing the different galaxies that are studied in the laboratory. Then we compare the wavelength location of each emission line obtained by the spectroscopic instrument for each individual galaxy with the location of the emission lines that is studied in the lab. Therefore, we have assumed that the chemical components of each galaxy emit lines in the same sequence as the ones observed in the lab. However, experimentally, we observe that the locations lines are stretched, which is an effect of the expansion of the universe. Through, this comparison we identify the difference of the emission lines with the rest frame

location and we measure the redshift of each galaxy. Therefore, we have access to the velocities of the galaxies. Then, we use the Hubble law, to convert velocities, i.e. redshift into distances. However, this velocity field, is mixed with the peculiar velocities of each individual galaxies, since galaxies are subjected to the gravitational field of their environment, for example other massive galaxies. This introduces deviations on the velocity field. This phenomenon is named redshift space distortions and we take that into account as well. We correct for this effect by estimating a correction using the theoretical prediction. However, this correction is subjected to question if it treats or not correctly this effect and if its effect is significant to our measurement and Cosmological Principle test.

Once we obtained the 3 dimensional position of our sample of galaxies we use, basically, the local number density of galaxies and compare it to the total density of the sample. However, this treatment is again complicated since the sample of galaxies that we obtained are only a part of the galaxy and not of the whole universe. This is due to that we have only access to the observable part of the universe. The light has a constant speed. Thus it needs time to travel from the distant galaxy and reach the earth, and cannot reach the detectors instantaneously. Furthermore, the spectroscopic instrument can only observe only a finite part of the celestial sky since our earth is located to a specific location inside the Milky Way. The electromagnetic radiation from the distance galaxies at low galactic-latitudes is screened by the very dense electromagnetic wall of our own galaxy, the Milky Way. For all those reasons, we need to assume that this partial sample of the whole universe, is representative of the whole universe. This is consistent with the isotropy hypothesis and a more general notion, the *ergodic theorem*. The ergodic theorem suggests that a 3 dimensional spatial average is equivalent to an ensemble average. Where the ensemble, is the set of different realisations of the survey, i.e. simulated galaxy catalogues. Moreover, we are going to use a suite of simulation of the galaxy catalogue to test this hypothesis. Additionally, we will see that our instrument surveys the celestial sky but leaves some regions unobserved, therefore surveying only a peculiar geometry of the celestial sky. In order to treat this effect there is again a plethora of methods, but we chose to use conventionally the comparison of galaxy counts with the counts of a random distribution which has the same properties as the galaxy sample. This way we are going to quantify a scale at which the universe transit to a homogeneous state.

The theoretical predictions that describe our observable universe are related to the dynamics and evolution of the universe. This is quantified by the power spectrum of the matter of the universe at different epoch of the universe. Therefore we use the power spectrum in order to predict the underlying visible and non visible matter field of the universe and via the definition of the homogeneity scale at the epoch we perform our observations and finally compare it to the measurement from our galaxy data sample.

If we are convinced enough by this first study that the homogeneity hypothesis holds, then the quantification of the homogeneity scale is going to define standard spheres on the universe that correspond to this transition. Therefore, we can use this scale to probe different scenarios of Λ CDM-model by measuring different set of cosmological parameters with it. To explore its sensitivity first we are going to confront it against a previous standard ruler, the BAO scale, by measuring it on thousands of simulated galaxy catalogues. Then, we are going to use the

standard Fisher Analysis Formalism, which allows us to determine the best sensitivity of a probe, i.e. a relation between small changes of the cosmological parameters with small changes on the measurements of this specific ruler, the homogeneity scale.

Another investigation we are going to follow is to determine the sensitivity of this probe using different experiments. Therefore we are going to simulate the future galaxy surveys, i.e. simulate the number density, angular coverage and redshift precision of each experiment and we are going to mitigate their effect by defining scaling laws with the precision of the homogeneity scale.

Finally, to probe further the assumption of homogeneity we are going to relax the homogeneity hypothesis on our reconstruction of our observables. We are going to achieve that by using two different ways to measure distances. The first one is the cosmic chronometers that are used to estimate the hubble expansion rate against different redshifts. To perform this measurement one relies only on astrophysical models that only assume general theory of relativity for gravity and standard structure formation history models that assume no homogeneity or isotropy. This is an alternative way of defining the hubble rate without the assumption of homogeneity. The second one is the Type Ia Supernovae distance module data. Those data are going to provide us with the information of the transverse comoving distance between the positions of the galaxies that are located at the same epoch as the one of the Supernovae. We perform the Supernovae distance module so as not to assume any curvature about the spatial curvature of our universe. Therefore by using those two methods, we can reconstruct a 3 dimensional comoving distribution (x, y, z) from the regular (z, θ, ϕ) measurement of each galaxy. Then we perform the same analysis discussed above and quantify the transition to homogeneity.

Chapter 3

Large Scale Structure Surveys

In this chapter, we discuss existing and upcoming large scale structure surveys. We introduce the difference between spectroscopy and photometry. Furthermore, we present the main goals of the Sloan Digital Sky Survey and the data reduction pipeline for the Baryon Oscillation Spectroscopy Survey which produced the data used in this thesis. Finally, we discuss the galaxy sample selection that was performed for the analysis developed in this thesis.

3.1 Introduction

Comprehensive wide-field imaging and spectroscopic surveys of the sky have played a key role in astronomy, leading to new breakthroughs in our understanding of the solar system, our Milky Way Galaxy and our Universe ¹. In the era of the upcoming Stage IV experiments, various projects are going to provide us with a clearer and wider image of the nature of our Universe. Those projects are a class of large scale structure surveys which are one of the main interest of modern observational cosmology. While the angular positions of galaxies are in general easy to measure, the radial distances are not so easy. Indeed in an expanding universe, the radial distance is a function of the history of expansion and therefore is model dependent. In order to estimate those distances, large surveys typically use the spectrum of light emitted by galaxies, including absorption and emission lines and more general features, all resulting from the integrated stellar light. This results in a measurement of the redshift. Then we infer their distances, via the established hubble law, mapping the 3 dimensional cosmic distribution. It is important to stress that these measurements are model dependent, be it from from the astrophysical models to acquire the redshift to the expanding history of the universe models.

Various projects are aiming at this three dimensional mapping of the universe. Ground based telescope and satellite telescopes. There is a plethora of existing ground based telescopes. The 3.9m ground based Anglo-Australian Telescope, located at the Siding Spring Observatory near

¹ remarkable breakthrough is the detection of the Gunn & Peterson trough, which is a feature of the presence of neutral hydrogen in the intergalactic medium[119] which led to the discovery of the reionisation epoch. Another remarkable breakthrough is the imprint of the Baryon Acoustic Oscillations in the large scale structures of our universe [53, 54].

Coonabarabran, New South Wales, Australia, was used for the completion of 2 cosmological surveys. The 2dF Galaxy Redshift Survey which survey the large scale structures in a window of 1500 deg^2 of the sky at $z_m \sim 0.2$ and completed in April 2002. In January 2011, the WiggleZ Dark Energy Survey[120] was completed, reaching a cosmic volume of about 1Gpc^3 up to redshift $z = 1$ covering 1000 deg^2 of the equatorial sky in 7 regions. It obtained accurately 239000 spectra for UV-selected emission-line galaxies at a median redshift of $z_{med} = 0.6$. Currently, the Sloan Digital Sky Survey, hosting many projects, operates the 2.5m telescope in New Mexico, mainly mapping millions of galaxies and thousand of quasars in a total volume, for BOSS $V_{boss} \sim 17h^{-3}\text{Gpc}^3$ and for its successor eBOSS $V_{eboss} \sim 98h^{-3}\text{Gpc}^3$, at $z \sim 0.6$ and 3.5 , respectively.

Upcoming ground based surveys such as, the Large Synoptic Survey Telescope (LSST hereafter) and Dark Energy Spectroscopy Instrument (DESI hereafter) are going to increase the survey volume allowing for a better understanding of the cosmos especially regarding Dark Energy. In particular

- The DESI instrument[121] will conduct a five-year spectroscopic survey, covering $14,000 \text{ deg}^2$ of the sky, using the 4-m Mayal telescope located in Kitt Peak, Arizona. It is designed to obtain 5000 spectra simultaneously over the optical wavelength range, $360 < \lambda < 980\text{nm}$, with a wavelength-dependent resolution, $R = \lambda/\Delta\lambda \sim 2000 - 5500$. The main DESI scientific goals are focusing on Dark Energy Science)[121] with several Dark Matter tracers spanning from a variety of different type galaxies at $z \sim 1.7$ to Ly- α absorption features at higher redshifts, $2.1 < z < 3.5$.
- The LSST telescope, having a large spectrum of scientific goals, from Dark energy science to Asteroid science, will survey 18000 deg^2 of the sky using photometry. Using a 8.4m telescope with a special three-mirror design, which create a wide field of view, LSST has the ability to survey the entire observable sky only in three nights. The LSST Summit Facility is located on the Cerro Pachon ridge in north-central Chile. The LSST camera will produce extremely high quality images by adopting a new technology of a 3.2 Gigapixel flat focal plane array, obtaining 15TB of raw imaging data per night[122].

The Euclid Mission[123], is the most promising upcoming satellite, which will be launched in 2020. It will spend 6 years on the main scientific goal, i.e. enriching the cosmic geometry measurements by mapping 2 billions galaxies spread over more than one third of the sky, with a wide survey $\sim 18000 \text{ deg}^2$ and a deep survey of 40 deg^2 . The telescope design is a 1.2 m Korsch, with a three mirror anastigmat. A three-mirror anastigmat telescope consist of at least three types of mirrors. The typical order is a concave first mirror, a convex second mirror and a concave third mirror. These mirrors are arranged such that the first mirror and the second mirror form, from an object at infinity, an intermediate image situated between the second mirror and the third mirror. The third mirror forms, from this intermediate image, a final image in the focal plane of the telescope. In this kind of architecture of telescope, at least the surface of the concave third mirror is an n-polynomial surface[124]. This allow in a compact manner a very large field of view $\sim 0.54\text{deg}^2$. This telescope will perform Near Infrared Spectroscopy (NISF hereafter), $1100 < \lambda < 2000 \text{ nm}$, for galaxy clustering science, and Near Infrared Imaging Photometry for

Weak and Strong Gravitational Lensing science. It will gather cosmological information reaching high redshifts up to $z \simeq 2.0$.

3.2 Photometry vs Spectroscopy

Photometry is the technology of simply recording astronomical images of sources in the sky. The light from the sources passes through coloured filters where the wavelength information is only restricted to the intensity of the light admitted through the filter and modified by the detector transfer function. The main advantage of this method is that one creates a fast image on the sky. Another advantage is that the light from a source over the filter band is concentrated onto a spot on the detector allowing better S/N. Furthermore, photometry allows data collection for thousands of sources per image. However, the main disadvantage is that the effective wavelength resolution is only as good as the width of filter bandpass, degrading the ability to distinguish between individual spectral features.

In spectroscopy, the light follows a path through a sequence of dispersive elements. It is reflected from those dispersive elements, which usually comprise of a prism or diffraction gratings. As a result, different wavelengths focus on different pixels so that the actual spectrum of the source is measured. The main advantage is that one is allowed to record the light of one or more objects over a large wavelength range having excellent resolution in wavelength space, resulting in clear individual spectral features of the sources. However, one of the main disadvantages is that the light of the sources is spread over many detectors, rendering a small S/N ratio on each detector. This results in a difficult acquisition of the spectra in one exposure, usually $t_{exp} \simeq 10 - 100$ seconds on multi-slit or multifiber spectrographs. Therefore, one needs a large amount of exposure time.

These days, both techniques usually use a charged coupled device (CCD hereafter) as a detector. What differs is the instrument prior to the detector. For spectroscopy one uses a dispersive element while for photometry one uses a coloured filter. While spectroscopy results in high resolution in wavelength providing a quality distinction of the spectral feature, it requires a prior photometry to identify each targets. In spectroscopy only a small number of objects are actually measured at each exposure. Therefore, it requires efficient target selection at the prior stage, which can be only achieved through photometric observations and target identifications.

3.3 The Sloan Digital Sky Survey

One of the most successful astronomical surveys to date, the Sloan Digital Sky Survey, SDSS[125], started observations in 1998 and has completed three different phases. The data collected include optical imaging of one fourth of the night sky, $\sim 10,000\text{deg}^2$, as well as focusing on optical and near-infrared spectroscopy of over 3.5 million stars, galaxies, and quasars. These observations were all performed by the 2.5-meter Sloan Foundation Telescope at Apache Point Observatory, APO; [126]. The fourth phase, SDSS-IV, builds upon and extends both the infrastructure and scientific legacy of the previous generations of surveys.

The two first versions of the survey, SDSS-I and II, measured, during eight years of operation, more than 900,000 galaxies and 100,000 quasars, allowing the construction of the largest three dimensional map of large scale structures of the Universe. The project had a huge scientific impact, with more than 8000 publications related to SDSS data.

The third generation of the SDSS[127], SDSS-III, started in autumn 2008 based on the same telescope as SDSS-II with improved instruments in order to cover three main themes: dark energy and cosmological parameters, the history and structure of the Milky Way, and the population of giant planets around other stars. Three major programs cover those themes. One of the main surveys, the Baryon Oscillation Spectroscopic Survey, hereafter (BOSS) measured redshifts of 1.5 million massive galaxies and the Lyman-alpha forest spectra of 150,000 quasars, to study the first theme at redshifts $z < 0.7$ and at $z < 2.5$, respectively. After producing regular public Data Releases for more than 8 years, SDSS was updated in summer 2015. In this update update, there was the successor of BOSS, extended BOSS (eBOSS hereafter). Furthermore, there are also two subprograms that are concurrently operated with eBOSS. The next generation SDSS-IV has new goals that build upon the scientific results of previous SDSS surveys in the areas of galactic archaeology, galaxy evolution, and cosmology. Those programs are listed below:

- *APO Galactic Evolution Experiment 2* (APOGEE-2) expands the APOGEE-1, probing the Galactic history through mapping the chemical and dynamical patterns of its stars via high resolution ($R \sim 30,000$), high signal-to-noise-ratio ($S/N \geq 100$ for most targets), with a near-infrared spectroscopy, $1.51\mu\text{m} \leq \lambda \leq 1.70\mu\text{m}$. The second generation program has northern and southern components, APOGEE-2N and APOGEE-2S, respectively. APOGEE-2N continues at APO, with primary use of the bright time. APOGEE-2S utilises new infrastructure and a new spectrograph now installed at the 2.5-meter du Pont Telescope at Las Campanas Observatory (LCO). The pair of spectrographs at APO and LCO together aiming to target 400,000 spectra of stars. APOGEE-2's near-infrared observations yield access to key regions of the unobserved parts of the Milky Way from other surveys, which are predominantly conducted at optical wavelengths as was reported by Majewski et al. [128].
- *Mapping Nearby Galaxies at APO* (MaNGA) provides a high resolved mapping of the internal structure of nearby galaxies, at an early epoch, $z \sim 0.03$, establishing a new observing mode of SDSS, integral field spectroscopy. This new mode includes the spatial distribution of both gas and stars, enabling access to the dynamics, stellar populations, and chemical abundance characterisation as a function of the environment within each galaxy. Using about half of the dark time at APO, MaNGA relies on novel fiber bundle technology to observe 17 galaxies simultaneously by feeding the fiber output of independent integral field units into the optical BOSS spectrographs. MaNGA plans to observe 10,000 nearby galaxies with various environments and the stellar mass range $10^9 - 10^{11} M_{\odot}$. The MaNGA observations cover a spectrum of $0.35 < \lambda < 1 \mu\text{m}$ at $\sim 65 \text{ km s}^{-1}$ velocity resolution and $\sim 2 \text{ kpc}$ spatial resolution[129].
- *Spectroscopic Identification of ERosita Sources* (SPIDERS) is executed concurrently with eBOSS, using $\sim 5\%$ of the eBOSS fibers on sources related to X-ray emission. In late 2018,

cotargeting with the eROSITA instrument which it was installed at the Spectrum Roentgen Gamma satellite[130], SPIDERS will provide the most successful characterisation of 80,000 X-ray identified sources ($\sim 57,000$ X-ray galaxy-clusters and 22,000 Active Galactic Nuclei).

- *Time Domain Spectroscopic Survey* (TDSS) aims to spectroscopically characterize the variable sources identified by Pan-STARRS 1 data release[131] or between SDSS and PS1 wide field imaging, using $\sim 5\%$ of the eBOSS fibers. This will provide a statistical complete selection of 200,000 variable objects on the sky down to faint objects, $i = 21$, after the TDSS targeting of 20,000 objects not otherwise included by eBOSS target strategy.

3.3.1 (e)BOSS: Main Scientific Goals

Since the first measurements of the baryon acoustic oscillation peak in the large-scale structures independently by two teams using the 2dF Galaxy Redshift Survey[54] and by SDSS-II [53], a new window was opened for more accurate BAO measurements of using larger and deeper survey, which is the main goal of BOSS. The interest of mapping the large scale structures and observe the BAO feature is that the latter can be used as a standard ruler to give us information for the Dark Universe and its evolution, i.e. the nature of Dark energy and Dark matter.

The Baryon Oscillation Spectroscopy Survey (BOSS) was divided into two main spectroscopic surveys, executed concurrently over $10,000 \text{ deg}^2$ of the sky. The first one was dedicated to the survey of more than 1.5 million of massive galaxies in the redshift range $0.15 < z < 0.75$, with an eight times larger effective volume, traced using BOSS galaxies, than for SDSS-II. The second survey targets 150,000 quasars and their Ly α forest absorption over $2.1 < z < 3.5$. The galaxy survey was accurately analysed by independent teams of the BOSS collaboration resulting in consensus values[21] on transverse distance measurements² and hubble expansion rate at late epoch, $d_M(z \simeq 0.32) = 1294 \pm 21[r_d/r_{d,fid}] \text{ Mpc}$ and $H(z \simeq 0.32) = 78.4 \pm 2.3 \text{ kms}^{-1}\text{Mpc}^{-1}$ respectively, as well as at earlier epochs, $d_M(z \simeq 0.57) = 2179 \pm 35[r_d/r_{d,fid}]\text{Mpc}$ and $H(z \simeq 0.57) = 96.6 \pm 2.4 \text{ kms}^{-1}\text{Mpc}^{-1}$. While the quasar sample performed on higher redshifts.

The extended version of BOSS (eBOSS) completes the gap of the BOSS galaxy targeting into the new redshift regime, $0.6 < z < 2.2$. Using single fiber spectroscopy, it targets $\sim 250,000$ new galaxies in the range $0.6 < z < 1.1$ and 60,000 new quasars at redshifts $z > 0.9$. These samples improve the investigation of the expansion of the universe using the Baryon Acoustic Oscillation (BAO) and the growth of structure using large-scale redshift space distortions, as was forecasted by Dawson et al. [132]. In particular, a 1.2%(2.8%) accuracy is expected on angular distance measurements, $d_A(z)$, and 2.1%(4.2%) precision on $H(z)$, combined with billion BOSS galaxies at $z < 0.6$ (60,000 BOSS quasars at $z > 2.1$). However, systematic effects, such as metal absorption, damped Lyman- α systems or broad absorption lines listed by Bautista et al. [33], were not taken into account into this prediction. Furthermore, another science goal is the improvement of the accuracy on the Growth rate measurements[133] which will reach $\sim 3.3\%$. Finally, improvements are expected with eBOSS Lyman- α Forest analysis on neutrino

²Transverse (motion) distance is defined via the Eq. A.4, as $d_M(z) = \frac{d_A(z)}{1+z}$

mass constrains, currently giving values, $\sum m_\nu < 0.12eV$ (95%*C.L.*) when combined with CMB Planck Data [117].

3.3.2 The Telescope

The Sloan Digital Sky Survey utilises a telescope located at Apache Point Observatory, in New Mexico, United States of America. The telescope was designed by a modified version of the two-corrector Ritchey-Chretien which has a 2.5-m, f/2.25 primary, an 1.08m secondary, a Gascoigne astigmatism corrector, and one of a pair of interchangeable highly aspheric correctors near the focal plane, one for imaging and the other for spectroscopy. This results in a focal ratio of type f/5³. The telescope is instrumented by a wide-area, multiband CCD camera and a pair of fiber-fed double spectrographs. The telescope includes the following:

- A 3 deg diameter (0.65 m) focal plane with an excellent image quality and small geometric distortions over a wide wavelength range (300-1,060 nm) operating in imaging mode. It is developed with good image quality combined with very small lateral and longitudinal color errors in the spectroscopic mode. The unusual requirement of very low distortion is set by the demands of time-delay-and-integrate (TDI) imaging.
- A very high precision motion supports these TDI observations.
- A unique wind baffle/enclosure construction that maximizes image quality and minimizes construction costs.

Time-Delay-and-Integration is a technique dating back to the age of filming. It was developed to solve the problem of image smearing and improve the signal-to-noise ratio. The essence of the technique is to allow the camera shutter open and roll the film at the same relative velocity as the target image. This results in collecting a long strip of film of an apparently static scene. The main advantage is the elimination of the need for a mechanical shutter. Secondly it allows long exposure times without introducing smear. Finally, it enables a capture of tremendous amounts of high-resolution from a relatively small camera when compared to an equivalent framing camera. Instead of rolling the film, the registers of the CCD are clocked in synchronization with the velocity of the targeted object. Therefore, the charge is "coupled" (hence charge-coupled device) from row to row in the sensor. Image signal integrates higher and higher as it traverses the sensor such that the final image becomes brighter compared to an equivalent static image. Typical TDI CCDs have 64 rows allowing for a 64 times improvement in the signal and an 8 times improvement in the S/N .

The first light of the telescope was obtained in May 1998 and began regular survey operations in 2000. More details about the telescope design can be found in Gunn et al. [134]. The telescope is shown in the Fig. 3.1.

³Focal ratio is the ratio between focal length and the entrance pupil.

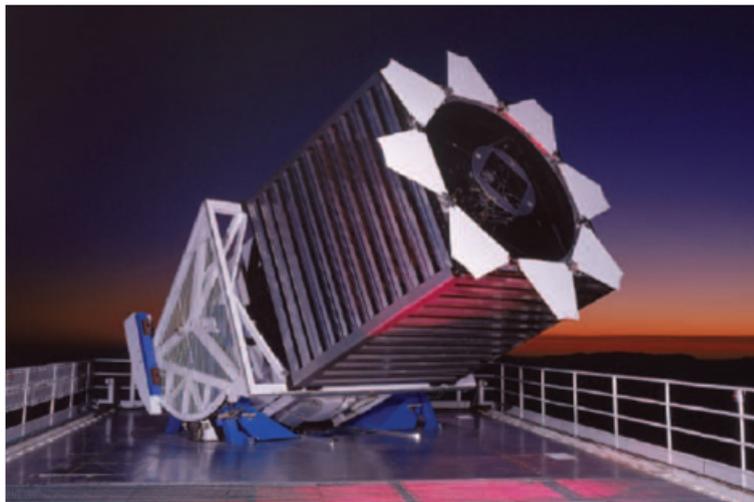


FIGURE 3.1: SDSS Telescope located in the Apache Point Observatory, in New Mexico, U.S.A.

3.3.3 The Camera

The camera operated by Sloan Digital Sky Survey (SDSS-I and SDSS-II) is a large-format mosaic CCD camera constructed by Gunn et al. [135] in 1998. It was optimized to perform in TDI imaging mode and better cover the 3 deg^2 field of view of the telescope.

The camera consist of six columns of five CCD, one for each photometric band, of $2k \times 2k$ ($=2048 \times 2048$) pixels each. The effective imaging area of 720 cm^2 as shown in Fig. 3.2. An astrometric array of 24 smaller ($400 \times 2k$)-CCDs with the same pixel size, is placed around the main CCDs and is used for astrometry and calibration. The TDI drift scan direction is shown by the Y axis.

The main photometric CCDs is equipped to carry out photometry simultaneously in 5 color bands (u,g,r,i,z), i.e. spanning the range from UV to near infra-red, which is the limit of silicon detectors. In Fig. 3.3, we see the quantum efficiency of each filter taking into account the optic, CCDs response and atmospheric extinction. The low relative efficiency of the u and z filters justify their positioning near the center of the focal plane.

The photometric detectors are arrayed in the focal plane in six columns of 5 chips each such that two scans cover a filled stripe 2.5° wide. For more technical and engineering details of the camera review Gunn et al. [135]. It is worth mentioning that this camera was used for the SDSS-I and SDSS-II photometric operations but not for the BOSS/eBOSS spectroscopic surveys.

3.3.4 The spectrographs

BOSS is constituted by a system of two identical spectrographs. The employed technology is similar to the SDSS-II versions. However, spectrograph quality improvement was acquired to reach BOSS scientific requirements. The system is executed in a single-fiber configuration by (e)BOSS, TDSS and SPIDERS. The twin multi-object fiber spectrographs are consist of a

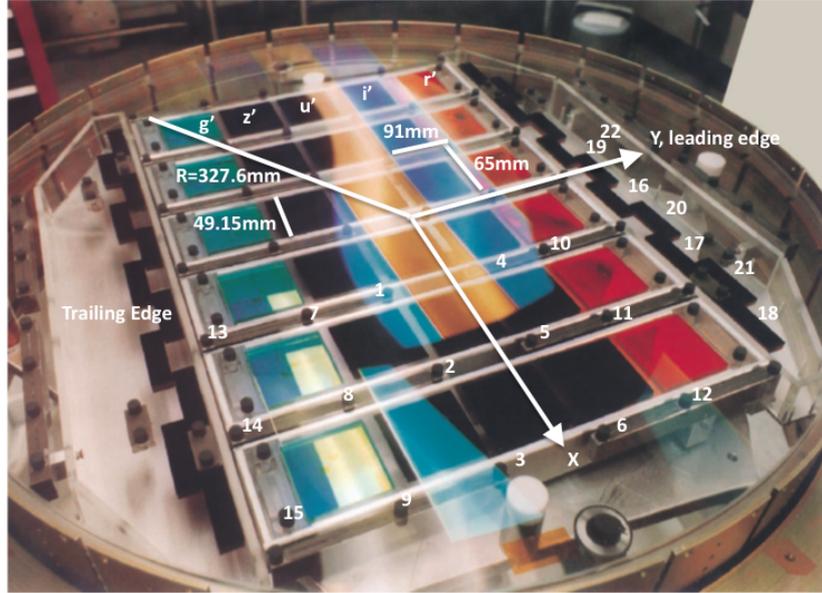


FIGURE 3.2: Photometric Camera Layout and its design overplotted. Optical layout of the focal plane of the SDSS camera. Field 22 (top and bottom) are focus CCDs; Fields 16-21 are astrometric chips, and 1-15 are the photometric array. The TDI drift scan direction is coming from the nose of Y-arrow to its tail, so a star-light traverses this array from right to left. [Image taken from [135] and modified to include the design]

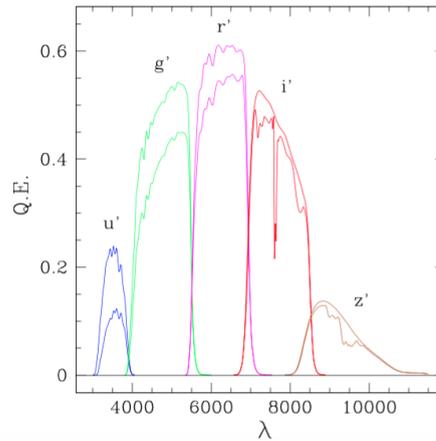


FIGURE 3.3: Quantum Efficiency of the telescope camera for the u,g,r,i,z photometric bands. The expected throughput of the optics is included; the lower of each pair of curves includes the expected atmospheric extinction. [Image taken from [135]]

simple optical layout with reflective collimators, gratings, all-refractive cameras, and state-of-the-art CCD detectors to produce hundreds of spectra concurrently in two channels over a bandpass covering the near-ultraviolet to the near-infrared, with a resolving power $R = \lambda/FWHM \sim 2000$ as shown in Fig. 3.4. The spectrographs were upgraded for BOSS with volume-phase holographic gratings and modern CCD detectors, improving the peak throughput by nearly a factor of two, extending the bandpass to cover $360 < \lambda < 1000$ nm, and increasing the number of fibers from 640 to 1000 per exposure. The throughput for BOSS was measured by averaging the throughput of 24 stars over three plates, split evenly between the two spectrographs. Full technical details about BOSS spectrographs can be found in Smee et al. [136].

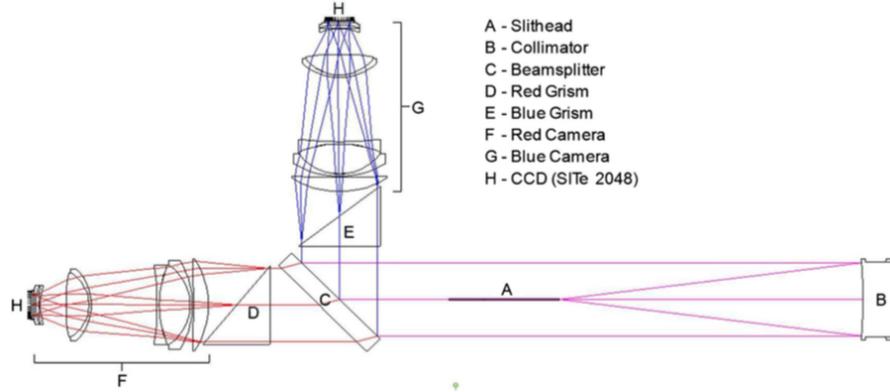


FIGURE 3.4: Optical layout of the SDSS spectrographs. Light enters each spectrograph through 320 fibers ($180\ \mu\text{m}$ diameter) terminating at a curved slit plate mounted inside the slithead. The slit plate positions the fiber ends on a radius concentric with the spherical collimating mirror (B). The 45° dichroic beamsplitter reflects the blue portion of the bandpass ($\lambda < 605\ \text{nm}$) and transmits the red wavelengths ($\lambda > 605\ \text{nm}$). Immediately after the beamsplitter in each channel is a grism (D,E), dispersing the light which exits the grisms and enters all-refractive, eight-element (F,G). Each camera contains a single $2^{11} \times 2^{11}$ CCD with $24\ \mu\text{m}$ pixels (H). [Image taken from Smee et al. [136]]

3.3.5 Wide-field Imaging

During the first stage of the survey (2000 to 2005) wide-field imaging was deployed covering the full $\sim 11,600\ \text{deg}^2$ of the Northern Galactic Cap (NGC) using the five photometric filters (u,g,r,i,z)-bands. The falls of the years 2008 and 2009, Southern Galactic Cap (SGC) was also imaged with an effective area of $\sim 3,100\ \text{deg}^2$. The final image covers $\sim 14,500\ \text{deg}^2$, more than one fourth of the sky, and contains almost a billion objects detected individually up to a magnitude of $r \sim 22.5$.

Spectroscopy is much more time consuming process than imaging, therefore a careful selection of the targets for spectroscopy needs to be performed prior to observation. Therefore, the next step consists in the spectroscopy of all interesting objects, galaxies, quasars, Lyman- α forest, (AGN) for (e)BOSS. For eBOSS, our main targets are galaxies, $z \sim 0.01 - 1.6$ and $z \sim 2 - 4$ quasars and Lyman- α forest. The BOSS collaboration developed algorithms for selecting these kind of targets. The next section briefly describes the process of galaxy target selection, which was used in our analysis.

3.3.6 Galaxy target selection

For galaxies, the main goal is to select galaxies composing a homogeneous distributed high density sample over the covered volume. Instead of selecting luminous red galaxies (LRG) as in SDSS I and II, BOSS galaxies of a same mass and no intrinsic color restrictions are selected, resulting in a denser sample ($\bar{n} \sim 3 \times 10^4 h^{-1} \text{Mpc}^3$ at $z = 0.57$). For BOSS, two classes of galaxies were selected based on their redshift, called *LOWZ* and *CMASS* samples. Following

Eisenstein et al. [137], we define the combination of color magnitudes as:

$$c_{\parallel} = 0.7(g_m - r_m) + 1.2(r_m - i_m - 0.18) \quad (3.1)$$

$$c_{\perp} = (r_m - i_m) - 0.25(g_m - r_m) - 0.18 \quad (3.2)$$

$$d_{\perp} = (r_m - i_m) - 0.25(g_m - r_m)/2 \quad (3.3)$$

where c_{\parallel} and c_{\perp} describe the low redshift locus while d_{\perp} the high redshift locus. It has been shown that d_{\perp} is highly correlated with redshift, z . Those new colors, $\vec{c} = (c_{\parallel}, c_{\perp}, d_{\perp})$, are appropriate to separate the galaxies based on their evolution. Targets are selected according to cuts on the new colors, \vec{c} .

The LOWZ algorithm is designed to select luminous red galaxies (LRG) at $0.15 < z < 0.45$. This the color cuts $16 < r_m < 19.5$, $r_m < 13.6 + c_{\parallel}/0.3$, and $|c_{\perp}| < 0.2$ correspond to the LOWZ galaxy sample. The CMASS galaxy sample, $0.43 < z < 0.7$, is based on the color-cuts: $r_i - i_m < 2$, $17.5 < i_m < 19.9$, $d_{\perp} > 0.55$, and $i_m < \min[19.86 + 1.6(d_{\perp} - 0.8)]$. The last condition is designed by Maraston et al. [138], ensuring a sample with all galaxies having the same constant stellar mass. The distribution of their masses is narrow and centered at $\log(M/M_{\odot}) \sim 11.3$ for both galaxy samples. Full details about galaxy target selection are discussed in Reid et al. [139].

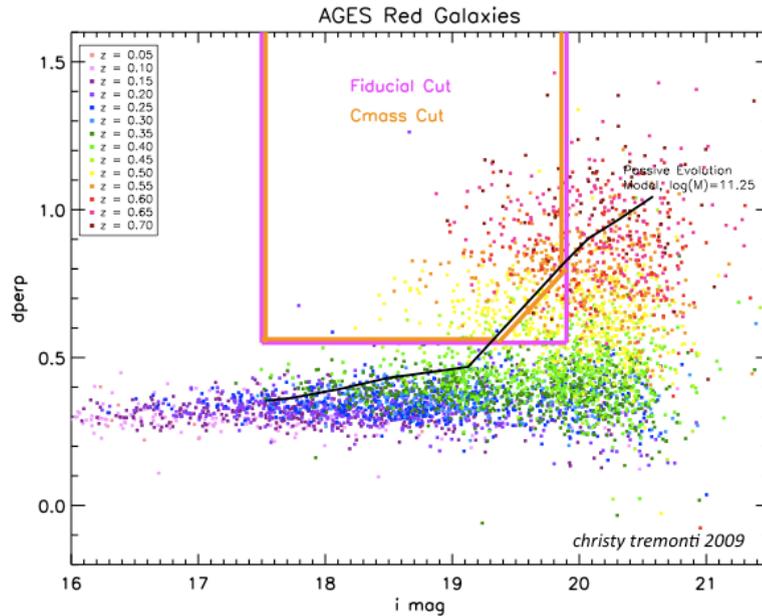


FIGURE 3.5: Target Selection algorithm, (see details in the text)

The diagram of Fig. 3.5 nicely describes in a pedagogical manner the description of the algorithm of various sample-cuts that were performed. In this figure it is plotted the magnitude in the i -band versus the d_{perp} variable of a sample of targets. The d_{perp} variable is a combination of color-magnitudes that are highly correlated with redshift as can be understood from the plot. Different colors represent different redshift. There is a color-cut at $d_{perp} \simeq 0.55$ which corresponds to a redshift $z \simeq 0.43$ which defines the high redshift galaxy catalogue. The $i = 17.5$ cut is applied to ensure the removal of very luminous objects while the $i_m = 20$ cut is applied

removes of very faint objects. The diagonal cut is dedicated to ensure a constant stellar mass limit on the output galaxy sample according to a passively evolving model [138].

3.3.7 Observations

As in the first two stages, SDSS-III and IV uses the multi-object spectroscopy technology, allowing the observation of hundreds of spectra simultaneously. The focal plane of the telescope is equipped with a 1.5 deg radius spectroscopic aluminum plate, accommodating 1000 optical fibers of 2" diameter as shown in Fig. 3.6.

Science targets are assigned to fibers in plates in a process referred as *tiling*. The *tiling* process is intended to maximize number of targets observed with the minimum number of tiles, *tiling efficiency*. The fraction of objects observed in a given class that were assigned by fibers is called *tiling completeness*. Approximately, 160-200 fibers per plate are dedicated to the main quasar targets, 560-630 to galaxy targets, and 20-90 to ancillary science targets. The minimum distance between two fibers on the same plate corresponds to 62" on the sky, which results in some collisions between targets. It may, however, be possible to observe both colliding targets if they are in the overlap region between two or more plates. The optical fibers are positioned on the drilled aluminium plate to guide the photons to the spectrographs. Measurements of the sky background and calibration with F-stars are assigned to ~ 100 fibers, distributed uniformly over the plate to ensure consistent data quality for all spectra, regardless of their position in the focal plane.

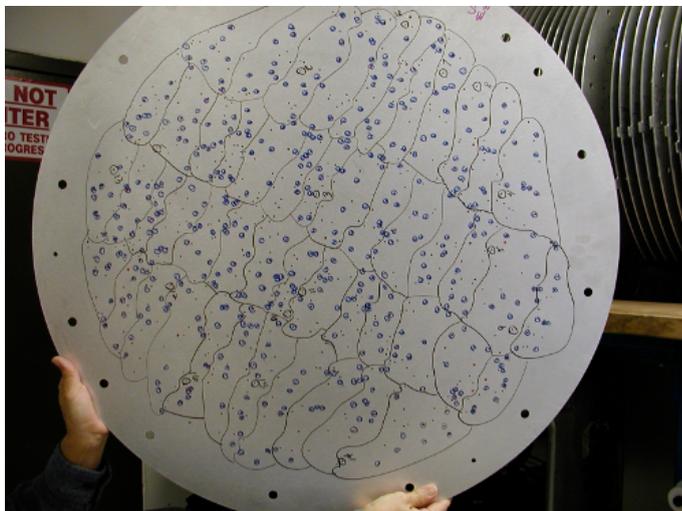


FIGURE 3.6: BOSS Aluminum plate. Bundles are separated by black bounded edges, and holes are marked in blue to reduce contamination between nearby emission line galaxies or quasars. Holes for guide star fibers are marked in black and denoted by the corresponding number ranging from 1 to 16.

Atmospheric differential refraction and time of the observation are taken into account to obtain the (x,y)-positions of the fibers in the plate. Galaxies, stars and sky fibers are placed so as the throughput is maximized near $\sim 540\text{nm}$, while quasar fibers at $\sim 400\text{nm}$ in order to improve the

signal in the Ly α forest. The z direction is corrected by the use of washers⁴. This allow to gather 1000 spectra per exposure [140], achieving a typical redshift uncertainty of a few tens on kms⁻¹ [141].

The observation starts, once a given plate is installed in the focal plane with all fiber connected. The procedure of observations consist of the following three basic steps:

- Flat-field calibration exposures and positioning of the telescope.
- A 15' sequence of science exposures, with a simplified version of data reduction pipeline being run for data-quality at each exposure.
- Re-observation of the science exposures which do not satisfy quality requirements.

The data reduction pipeline is basically a comparison between the magnitude of each object with the mean signal-to-noise of pixels in the g band for the blue cameras and in the i band for the red ones. For a low signal-to-noise output a new observation is scheduled.

3.3.8 Data reduction pipeline

eBOSS utilises $\sim 50\%$ of dark time at APO. The eBOSS pipeline is a slightly modified version of the BOSS pipeline[103]. A quicklook pipeline is used to pre-process eBOSS data during each observation. Estimation of S/N in real time result in informed decisions about continuing to subsequent exposures. Quality tests are examined each day for unexpected failures identification of the observing system or pipelines. Each morning, following a night of eBOSS observations, the data are transferred from the APO to the Lawrence Berkeley National Laboratory to run the automatic data reduction pipeline (pipeline hereafter). The eBOSS pipeline adopts the existing aspects of the BOSS pipeline. We have used the data from BOSS pipeline, since there was the most uniform sample of galaxies when this thesis started. We are going to describe briefly the main points of the BOSS pipeline.

In particular, the BOSS pipeline extracts, calibrates, coadds, classifies, and fits the redshift of all spectra using all exposures as described by Bolton et al. [141]. Firstly, the data are collapsed from two-dimensional images into one-dimensional spectra. Secondly, an algorithm of redshift characterisation and classification of objects is executed. This algorithms performs several tasks. The main tasks are described below.

An *Optimal Extraction* method is used during pipeline which extracts the individual spectra, building a spatially-dependent model of the sky spectrum from the 80 sky fibers and subtracts that model from each object fiber. The optimal extraction algorithm is able to sort out the cross-talk between closed fibers on the CCD.

Pre-processing of raw CCD frames is used for subtraction of a bias model, which is overscanned by subtracting a dark current model, and dividing by a pixel flat-field model for each CCD.

⁴Washer is a thin plate, typically disked shape, used to distribute the load of a threaded fastener, such as screw or nut.

Sky subtraction is performed by using a background model derived from the sky-fibers that were assigned during plate design. The background model varies with fiber position to account for small differences across the focal plane. The spectral response, i.e. the flux calibration, is determined using template spectra of the standard stars obtained during in plate design.

Co-addition of spectra from individual exposures of each fiber are combined into a single frame with a resampled grid that is linear in $\log(\lambda)$. Blue and red data are combined into a single spectrum covering the full wavelength range of the instrument, $361 < \lambda < 1014\text{nm}$.

Finally, a *Classification* of each object is defined through a fit of a star, galaxy, or quasar template into their combined spectrum. Due to some low S/N spectra the identification is sometimes inaccurate, therefore those spectra are flagged with warning tags, *ZWARNING* and special treatment is given to them as we describe in chapter 3.4.2. A set of templates is deployed to determine redshifts, z , with separate sets for stars, galaxies, and quasars. For stars, the templates consist of individual archetypes; for galaxies and quasars, the templates consist of Principal Component Analysis (PCA) sets that are linearly combined to fit the data at each potential redshift. The best redshift and classification (star, galaxy, or quasar) is determined based on the $\Delta\chi^2$ between the data and the templates. For galaxies, the pipeline performs an additional fit on the velocity dispersion parameter of the galaxy, by comparing the spectra with linear combinations of a set of high-resolution stellar templates. The pipeline conducts emission line flux and width measurements for a fraction of the major emission lines.

The BOSS pipeline confirms that for the color-selected CMASS sample of massive galaxies targeted at redshift $0.4 < z < 0.8$ the pipeline achieves an automated classification success rate of 98.7% and confirms 95.4% of unique CMASS targets as galaxies (most of them being M-stars⁵). A subset of BOSS galaxies was visually inspected, retrieving about only 0.2% of confidently reported CMASS sample classifications and redshifts are incorrect, and approximately 0.4% of all CMASS spectra are objects unclassified.

The pipeline undergoes continuous improvement as problems are identified and repaired. Future versions will benefit from ongoing efforts to improve sky subtraction and spectrophotometric calibrations. Specifically, from 2016 the SDSS operate the eBOSS pipeline. In particular, one of the new things implemented on eBOSS is the determination of galaxy and quasar redshifts with a large number of fixed archetypes using the *REDMONSTER* suite[142] rather than a PCA basis set.

The eBOSS pipeline has been applied to all SDSS-III BOSS data as well and further compared to *spec2d* suite for calibration purposes. The first SDSS-IV data release DR13 contains a rereduction of BOSS data, including plates from SDSS-IV completing the SEQUELS sample (Sloan Extended QUasar, ELG and LRG Survey). In DR14, the first two years of eBOSS data will be released and become available before the fall of 2017.

A sample of results of both pipelines (BOSS/eBOSS) are displayed in Fig. 3.7 by a sample of six spectra from the first year of eBOSS, processed through a preliminary version of the eBOSS pipeline and two additional galaxy spectra obtained by the BOSS pipeline (top row). One can observe the different elements identified for each target.

⁵Class M-stars are by far the most common living in the main sequence.

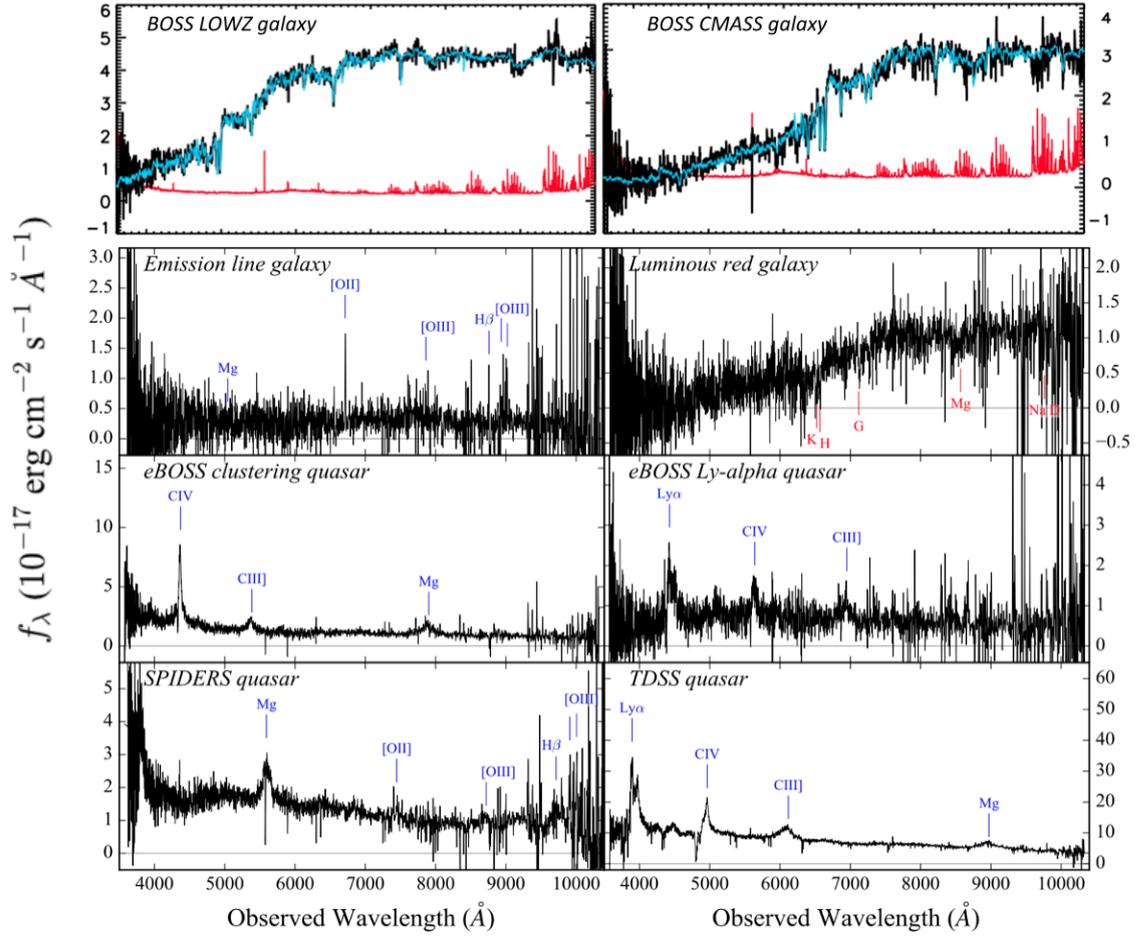


FIGURE 3.7: Sample of SDSS spectra. (Top) 2 SDSS-III targeted galaxies spectra (black) and the noise at each pixel (red) with the best redshift measurement (blue). (2nd row to Bottom) are 6 SDSS-IV targeted objects (black) with their identified emission lines (blue). [Image extracted by [103] and modified]

3.4 Sample Selection Strategy

In this thesis, we used the CMASS galaxy sample that has a number of properties that are useful in the homogeneity study:

- Volume-limited galaxy sample.
- Time-invariant galaxy sample properties⁶
- Mass Galaxy invariance

Furthermore to fulfill statistical test of this sample we used mock galaxy catalogues dedicated to mimic the aforementioned sample and the survey characteristic, which we discuss in chapter 3.4.3

⁶ Selecting only passively evolving objects, we ensure to obtain a sample were no different type of galaxies contaminate the sample.

3.4.1 Homogeneity study Sample

A 3D representation of the CMASS galaxy sample is given in Fig. 3.8 around the Milky Way (M.W.). In this figure, one can also distinguish the redshift-slicing of the sample that we have selected to probe the cosmic evolution of the homogeneity scale. The large masses of the obtained objects ensures this sample to have a strong bias with respect to the underlying Dark Matter field, allowing a high signal to noise ratio on the two-point correlation function as described by previous observations [143]. This allows us to measure the cosmic bias which is an important quantity as we explain in chapter 1.2.6.

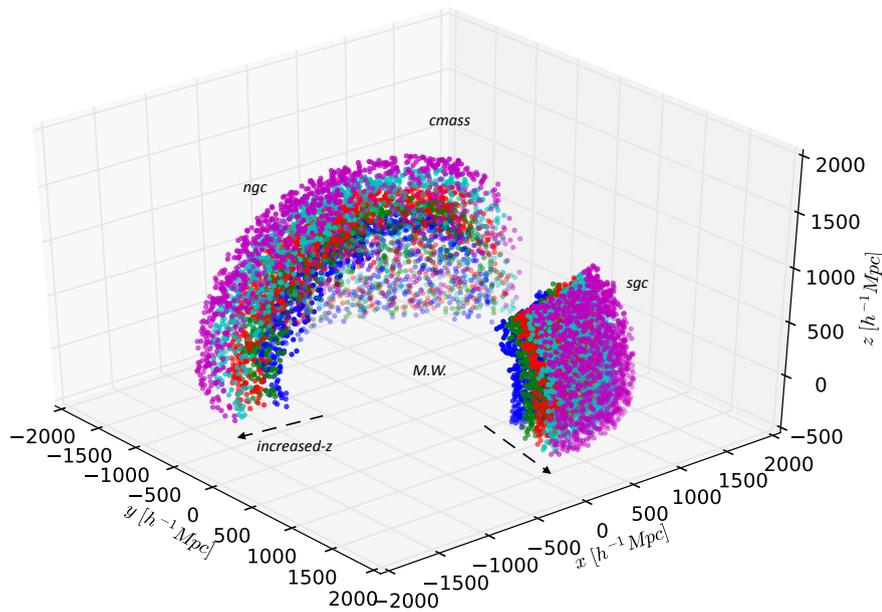


FIGURE 3.8: 3D representation of the CMASS sample. Different colors represent different redshift slices applied in homogeneity galaxy analysis[144].

The CMASS galaxy sample consists of $\sim 10^6$ galaxies with accurate spectroscopic redshifts. The CMASS sample has been used to measure the Baryonic Acoustic Oscillation feature (BAO) with unprecedented accuracy allowing to put strong constraints on the cosmological parameters [145].

z	NGC	SGC
0.430-0.484	101,383	40,170
0.484-0.538	174,468	63,518
0.538-0.592	151,084	56,805
0.592-0.646	97,155	37,179
0.646-0.700	47,289	17,899
0.430-0.700	571,379	215,571

TABLE 3.1: DR12 data sample in 5 redshift intervals in the NGC and SGC

We used the Data Release 12 (DR12) and we divide our galaxy sample according to SDSS specifications on the North Galactic Cap that span approximately 6934deg^2 and South Galactic Cap of effective area $2,560 \text{deg}^2$ (NGC and SGC respectively) [139].

Furthermore, we divide our data sample into 5 redshift regions as shown in table 3.1 to study the evolution of the clustering of our galaxy sample. The redshift profile of data sample is shown in the bottom panel of Fig. 3.9 which spans between $0.43 \leq z \leq 0.7$ that define the CMASS galaxy sample as explained in [139] and references there in for a volume limited sample. The R.A. takes values in the region between $(109, 264) \text{ deg}$ for NGC, while for the SGC takes values in the region $(0, 45) \text{ deg}$ and $(316, 360) \text{ deg}$. The Dec variable takes values in the region $(-6, 69) \text{ deg}$ for the NGC while for the SGC in the region $(-10, 36) \text{ deg}$. These profiles are shown in the upper part of Fig. 3.9. We use the prediction of flat ΛCDM -model to convert the redshift measurements to comoving distances using Eq. 1.2. The cosmological parameters used for calculating distances are given by:

$$\mathbf{p}_{\text{cosmo}} = (\omega_{\text{cdm}}, \omega_b, h, n_s, \ln [10^{10} A_s]) = (0.1198, 0.02225, 0.6727, 0.9645, 3.094), \quad (3.4)$$

and therefore we assume flatness. The numerical values in Eq. 3.4 are from Planck 2015 T T, T E, EE + lowP analysis [6]

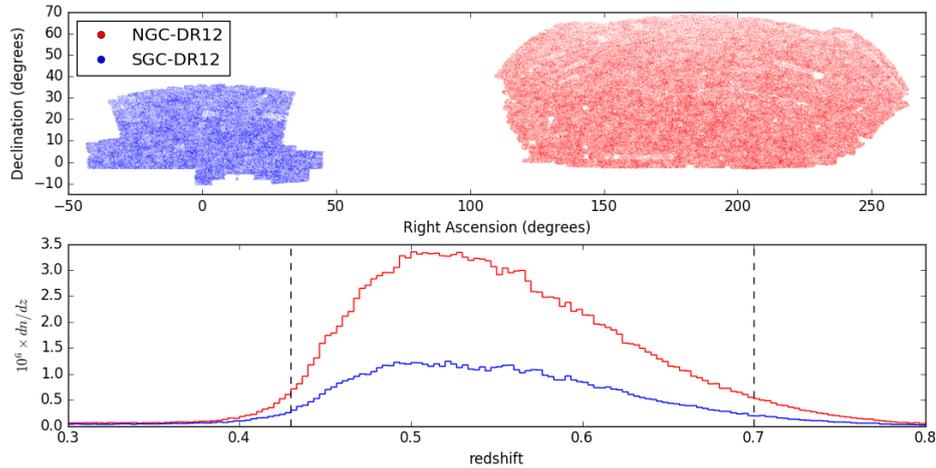


FIGURE 3.9: R.A. and Dec profiles (top) and redshift distribution (bottom) of the DR12 sample for the NGC and SGC. The vertical dashed lines define the redshift-cut used in the analysis [144].

3.4.2 Galaxy weighting scheme

There are a number of effects occurring when one measures the position of the galaxies in practice. The effects that were currently identified are the ones following Reid et al. [139] specifications. Due to fiber coating, one cannot assign optical fibers on the same plate to two targets that are closer than $62''$. Therefore, one needs to weight with the close-pair weight, w_{cp} . This means that for two galaxies that are within $62''$ angular distance we keep one of them count it as two galaxies to account for the unobserved one. Targets for which the pipeline failed to measure the redshift are identified resulting to the w_{noz} weight. This weight therefore corrects for the number of galaxies that we had a redshift but was poorly identified. Furthermore, there is a dependance of the observed galaxy number density with the stellar surface density, especially in the equatorial plane where the stellar density is more dominant. The observed local galaxy

density is also correlated with the *seeing*. So one weights each galaxy with a w_{star} and w_{see} weight to account for the two aforementioned effects, respectively. Finally, in order to reduce the variance of the two-point correlation function estimator, we use the FKP weight, w_{FKP} , [146]. This weight is defined as:

$$w_{fkp} = \frac{1}{1 + \bar{n}(z_i)P_0} \quad (3.5)$$

where $\bar{n}(z_i)$ is the mean density of the sampled galaxy distribution at redshift z_i . Furthermore, P_0 defines the value of the power spectrum amplitude at which the error is minimized. In this analysis we have used $P_0 = 20,000h^{-3}\text{Mpc}^3$ at $k = 0.10 = h/\text{Mpc}$. However, the latter approximation ignores the scale dependence of the power spectrum, but it has been shown that do not affect the final results and errors. Therefore, our galaxy weighting scheme is weighting each galaxy according to a total weight given by:

$$w_{gal} = (w_{cp} + w_{noz} - 1) \times w_{star} \times w_{see} \times w_{fkp} . \quad (3.6)$$

3.4.3 Mock Galaxy Catalogues

Mock Galaxy Catalogues are an important tool for determining uncertainties in galaxy surveys and tuning data analysis pipelines. With those mock galaxy catalogues, we are able to perform statistical and null tests to falsify our observations. In practice, mock catalogues are used to replace the standard sophisticated methods of jackknife, or bootstrap techniques[147] on the estimation of the covariance matrix of the galaxy sample. The substitution of the standard methods is done because they poorly mimic the Cosmic Variance that is an inherent source of error on our observations on large scales structures. Thus, we use the mock galaxy catalogue generation techniques that use both our knowledge of the cosmic variance and the survey errors for the construction of the catalogue.

There is a plethora of methods for constructing mock catalogues in the literature, each having their own weaknesses and strengths. The basic method for constructing such catalogues is done in the following summarising steps:

- predicting the evolution of the underlying mass field
- locating and characterizing the properties of dark matter halos
- populating the halos with mock galaxies
- applying survey characteristics to the *box*⁷ of galaxies

The survey characteristics are summarized on the selection function which is constructed through the information of:

- telescope efficiency (*seeing*)
- survey geometry

⁷The mass field and the mock galaxy catalogue both occupy a volume larger than the one of the survey in study, namely *box*. This is done so that the box will contain the survey.

- luminosity function
- selection function

All the above criteria are essential ingredients to mimic for the construction of the catalogue. The procedure of predicting the galaxy field is usually relying on modelling the non-linear dynamics on small scales, $r < 20 h^{-1}\text{Mpc}$, such as Halo-Occupation Algorithms [148], which often provides an accurate and reliable method to mimic the small scales. Finally, one need to compare the generation of those catalogues with N-body simulations⁸ through different statistic measures, such as the 2pt Correlation function. The latter simulations are the most realistic simulations of the universe, but they are computationally heavy.

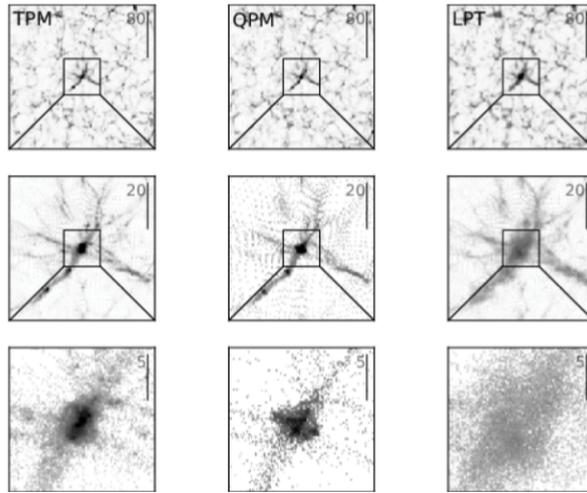


FIGURE 3.10: Qualitative comparison mock catalogues QPM and LPT with N-body simulations (TPM). The advantage of QPM against LPT is clear on the non linear regime.[image taken from White et al. [150]]

In the literature, several semi-analytical methods were developed in 2nd-order Lagrangian Perturbation Theory (LPT)[151]⁹ to model the growth of structure on the nonlinear regimes. As a result, several set of codes exists to construct the mock catalogues such as PTHalos [153], COLA[154], L-PICOLA[155], EZmocks [156].

In this study, we used the state of the art 1000 mock catalogues constructed with the quick-particle-mesh (QPM) algorithm [150] dedicated to mimic the BOSS survey mask. These are expected to be more realistic than mocks based on LPT methods as shown in Fig. 3.10. These mock catalogues use a different flat ΛCDM -cosmology than the one we use in this analysis¹⁰:

$$\mathbf{p}_{\text{qpm}} = (\Omega_{\text{cdm}}, \Omega_b, h, n_s, \sigma_8) = (0.274, 0.046, 0.7, 0.95, 0.8) \quad (3.7)$$

where σ_8 is the variance of the matter power spectrum computed in spheres of $8h^{-1}\text{Mpc}$ radius. However, we take that into account on our analysis pipeline, as we explain in Fig. 4.1.5.2

⁸The most powerful N-body simulation called Millenium developed in MPA [149]

⁹The first order of Perturbation Theory on the Galaxy Power Spectrum impressively agrees with the Zel'dovich Approximation [152]

¹⁰This difference is accounted for in our analysis, see chapter 4.1.5.2

Chapter 4

Homogeneity Measurement

In this chapter, we discuss how to measure the homogeneity scale of the universe using galaxies as tracers of the matter density field. The development of an alternative estimator of the characteristic scale of transition to homogeneity is presented which is part of my original contribution and published in Ntelis et al. [144] using an Landy Szalay inspired estimator. The test with the Fractal Distribution was done with the help of P.Laurent. While the statistical tests were done with the help of J.M. Le Goff, J.Rich and Nicolas Busca. On the Computational side, I had a lot of help from M.Stolpovskiy, R.Banerji and D.Poletti.

4.1 Homogeneity Observables

As discussed before in order to study the homogeneity of the universe, we use the galaxy distribution in which we treat the galaxies as point particles. Currently, three are the most common observables that can be used for studying the statistics of the clustering of point distribution. The counts-in-spheres, $N(< r)$, the fractal correlation dimension, $D_2(r)$ and the correlation function, $\xi(r)$. We will see in this chapter that all those observables are related to each other, although giving different estimation of the statistical clustering of the galaxy distribution. Furthermore, in this chapter, we will discuss the different selection criteria we used in our analysis and the treatment of the Redshift Space Distortion (RSD hereafter) models.

4.1.1 Counts-in-spheres and Fractal Correlation Dimension

The “counts-in-spheres”, $N(< r)$, is defined as the average number of objects in a sphere of radius r . It should obviously scales as r^3 for a homogeneous distribution since such a distribution has a spatial independent density. However, for inhomogeneous distribution the measured counts-in-spheres variable does not scale the same. Instead, we relax the power of the scaling with a parameter, D_2 , to observe departures from a homogeneous behaviour. Therefore, this power can be redefined, as:

$$D_2(r) \equiv \frac{d \ln N(< r)}{d \ln r} . \quad (4.1)$$

This is the definition of fractal correlation dimension. For a homogeneous distribution one gets $D_2 = 3$ at any scale, r , while for a fractal distribution $D_2 < 3$ is a measure of the fractal dimension, due to clustering of the distribution, since fractals have not radial-independent mean density.

However, these simple expressions are no longer valid in the case of the observable survey that has a peculiar geometry and a non-uniform completeness. The number of incomplete spheres is more important as r increases. This drastically reduces the accuracy of our measurement at large scales. However, to correct for this geometrical effect we use a random catalog to take into account all spheres, including those that are not completely included in the survey. This way we correct for these geometrical effects. The correction is simply the following procedure. We uniformly fill the survey volume with a random point distribution and we compare the counts-in-sphere on the data galaxy catalog with those of the random point catalog. We therefore define the “scaled counts-in-spheres” as:

$$\mathcal{N}(< r) = \frac{N_{gal}(< r)}{N_{rand}(< r)} \quad (4.2)$$

which is the ratio between the counts-in-spheres of the galaxy distribution, $N_{gal}(< r)$, and the counts-in-spheres of the random point distribution inside the same survey geometry, $N_{rand}(< r)$. This quantity is now expected to be independent of r , since the ratio of two quantities, a homogeneous galaxy (data) distribution and a random point distribution, does not scale with r^3 . Therefore, for a homogeneous distribution the fractal dimension is redefined as:

$$\mathcal{D}_2(r) \equiv \frac{d \ln \mathcal{N}(< r)}{d \ln r} + 3, \quad (4.3)$$

in order to be equal to 3 for a homogeneous distribution.

Thus, $\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$ are the two observables we will consider in this analysis, and we show in chapter 4.1.4 that there are different ways to obtain them from the galaxy data.

4.1.2 $\xi(r)$ - $\mathcal{N}(< r)$ relation

In chapter 1.2.6, we have defined the correlation function, $\xi(r)$, as the excess probability, with respect to a homogeneous distribution, of finding two objects separated by a distance r . When the characteristic scale of homogeneity is reached, one expects to obtain a vanishing $\xi(r)$. However, this definition implies that estimators of this observable requires the prior knowledge of the total average matter density n_{univ} . The existence of an average density of matter requires that the distribution of matter become homogeneous on large enough scales. Therefore it is not strictly rigorous to use $\xi(r)$ to establish the homogeneity of the universe. On the other hand, measuring simply the number of objects in counting spheres $\mathcal{N}(< r)$ and the fractal correlation dimension $\mathcal{D}_2(r)$, defined in chapter 4.1.1, do not, a priori, require the knowledge of the average density of tracers.

However, for the purposes of our analysis, we need to consider the relationship between the 2-pt correlation function and the scaled counts-in-spheres. As shown by Peebles [157], the counts-in-spheres of the distribution of galaxies is related to the correlation function by:

$$dN(\vec{r}) = \bar{\rho} [1 + \xi(\vec{r})] d^3\vec{r}. \quad (4.4)$$

Assuming isotropy $\xi(\vec{r}) = \xi(r)$ and integrating we get:

$$N(< r) = 4\pi\bar{\rho} \int_0^r [1 + \xi(r')] r'^2 dr'. \quad (4.5)$$

For the random homogeneous distribution we have, $N_{rand}(< r) = \bar{\rho} \frac{4\pi}{3} r^3$, so

$$\mathcal{N}(< r) = \frac{N(< r)}{N_{rand}(< r)} = \frac{3}{r^3} \int_0^r [1 + \xi(r')] r'^2 dr' = 1 + \frac{3}{r^3} \int_0^r \xi(r') r'^2 dr'. \quad (4.6)$$

This simple formula is used in two cases. We use it to predict the counts-in-spheres and fractal correlation dimension from our theoretical model. We use it as an alternative estimator for the scaled count in spheres. Notice that Eq. 4.6 implies that Λ CDM-model predicts different counts-in-spheres for a tracer of the total matter of the universe and quite different for the total matter of the universe, since it is described by $\xi_{tr}(r) = b^2 \xi_{mat}(r)$.

4.1.3 Homogeneity scale definitions

In our study we explore two different classes of homogeneity scale definitions. Following Scrimgeour et al. [93], we define a homogeneity scale R_H for both \mathcal{N} and \mathcal{D}_2 observables, as the scale for which the observable approaches its nominal (homogeneous) value within a percentage value, p_H . We then have two a-priori different homogeneity scales defined by:

$$\mathcal{D}_2(\mathcal{R}_H^{\mathcal{D}_2}) = \mathcal{D}_2^{nominal} (1 - p_H) \quad (4.7)$$

and

$$\mathcal{N}(\mathcal{R}_H^{\mathcal{N}}) = \mathcal{N}^{nominal} (1 + p_H). \quad (4.8)$$

where obviously $\mathcal{N}^{nominal} = 1$ and $\mathcal{D}_2^{nominal} = 3$. The "-" and "+" signs is because we expect increasing and decreasing functions. We can arbitrarily choose either $p_H = 1\%$ or $p_H = 0.1\%$. We examine thoroughly the former possibility and we give only as simple estimate for the latter in the chapter 4.5. The aforementioned homogeneity scale definitions are arbitrary but they do not depend on the survey since they do not depend on the estimation of the error on the observable, \mathcal{N} or \mathcal{D}_2 but only in their absolute values. Therefore they can be used to test cosmological models and compare different survey measurements as long as the same definitions are used in all cases. However its drawback is that it depends on the value of the bias as we are going to see.

Other authors have proposed different ways of defining the homogeneity scale. In particular, the homogeneity scale can be defined as suggested by Yadav et al. [158], i.e. *the scale beyond which it is impossible to distinguish the given distribution from a homogeneous distribution,*

within the errors of estimation. Thus, an alternative way to define the homogeneity scale, using any observable $\mathcal{O} = \{D_q, N, \dots\}$ and its dispersion as a function of galaxy separations. This is obtained formally by:

$$\mathcal{R}_H^E := \min \{r \in \mathbb{R} : (\mathcal{O} + \sigma_{\mathcal{O}})(r) > \mathcal{O}_{rand}\} \quad (4.9)$$

where \mathcal{O}_{rand} is the nominal value of the observable applied to a random distributed point sample, and $\sigma_{\mathcal{O}}$ is the 1σ standard deviation of the observable.

This is obviously another proper definition of the homogeneity scale because it is independent of the bias and similar conclusions can be drawn as we demonstrate at the end of section chapter 4.6. However, we use the Eq. 4.7 as the Homogeneity scale estimator, for our basis analysis, since its estimation has less covariances as we show in chapter chapter 4.3.2. Furthermore, this estimator is independent of the survey, and therefore it can be used as an index of comparing different surveys.

4.1.4 Homogeneity Estimators

In order to select the proper estimator for the homogeneity study, we need to discriminate from all the possible ones. As long as we have the 3 point distributions, i.e. galaxy catalog (data) and random catalog (needed for correcting the geometric effect), we perform the following computation. We select each object of the catalog as a center and we draw a sphere of radius, r , around it. We compute the number of galaxies within this sphere. Then, we measure this number for each galaxy of the catalog. Then we perform the same computation for every galaxy of the survey as a center and compute for a higher separation distances, r . This computation is defined as

$$GG(r) = \frac{1}{G} \sum_{i=1}^G n_i^{(G)}(r), \quad (4.10)$$

where G is the total number of galaxies and $n_i^{(G)}(r)$ is the number of pairs around each galaxy center i which is separated by a distance r from the center. However in order to correct for geometric effects of the survey that we take into consideration, we perform the same computation on a random catalog that lies into the same survey region, i.e. same $R.A.$, Dec and z ranges as the galaxy catalog. Therefore we need to define,

$$RR(r) = \frac{1}{R} \sum_{i=1}^R n_i^{(R)}(r), \quad (4.11)$$

where R is the total number of random points and $n_i^{(R)}(r)$ is the number of pairs around each random-point center i . Furthermore, it is needed to compute the cross-reference between the catalogs. Therefore, we throw in our survey region both catalogs and by selecting each galaxy as a center, we compute the average number of random points for different distance separations,

$$GR(r) = \frac{1}{G} \sum_{i=1}^G n_i^{(R)}(r) \quad (4.12)$$

where $n_i^{(R)}(r)$ is the number of random-points around each galaxy center i . However, in order to compare apples with apples we need to normalize each quantity since we count a pair of galaxies twice. Therefore we define:

- $gg(r) = \frac{GG(r)}{(G-1)/2}$: the normalized galaxy pairs distant by r ,
- $rr(r) = \frac{RR(r)}{(R-1)/2}$: the normalized random-point pairs distant by r ,
- $gr(r) = \frac{GR(r)}{R}$: the normalized average galaxy-random-point pairs distant by r ,

where G and R are the total number of galaxies and random points, respectively. From these quantities, we can construct different estimators for the different observables, ξ , \mathcal{N} and \mathcal{D}_2 .

4.1.4.1 2ptCF estimators

Several authors proposed different estimators for the two-point-correlation function, (2ptCF hereafter) by combining differently the above quantities. We list them in table 4.1. In a recent study, Vargas-Magaña et al. [159] have compared, the Landy-Szalay estimator [160], $\widehat{\xi}_{ls}(r)$, with a sophisticated estimator which is constructed by a linear combination of all possible pair counts combination, $\widehat{\xi}_{vb}$. They have found that when the latter estimator is used to galaxies samples for cosmological inferences, i.e. $(\Omega_m, \Omega_\Lambda)$ measurements, gives better performance than $\widehat{\xi}_{ls}(r)$ with a gain of 10 – 15%, on old data releases DR7 and DR9. However, we use the Landy-Szalay estimator, $\widehat{\xi}_{ls}(r)$, since it is the simplest to implement and lighter computationally [160].

Author	2ptCF-Estimator	Computational Speed	$r^2\sigma_\xi(r)$ [$h^{-2}\text{Mpc}^2$]
Peebles and Hauser [161]	$\widehat{\xi}_{ph}(r) = \frac{gg(r)}{rr(r)} - 1$	high	20-40
Hewett [162]	$\widehat{\xi}_{he}(r) = \frac{gg(r) - gr(r)}{rr(r)}$	high	20-40
Davis and Peebles [163]	$\widehat{\xi}_{dp}(r) = \frac{gg(r)}{rd(r)} - 1$	high	20-40
Hamilton [164]	$\widehat{\xi}_{ha}(r) = \frac{gg(r) \times rr(r)}{rr(r)} - 1$	high	10-20
Landy and Szalay [160]	$\widehat{\xi}_{ls}(r) = \frac{gg(r) - 2gr(r) + rr(r)}{rr(r)}$	high	10-20
Vargas-Magaña et al. [159]	$\widehat{\xi}_{vb}(r) = \sum_i c_i f_i [gg(r), gr(r), rr(r)]$	low	9-19

TABLE 4.1: The current list of 2pt correlation function estimators, $\xi(r)$, and their rescaled RMS, $r^2\sigma_\xi(r)$, in the region $r \in [80 - 120] h^{-1}\text{Mpc}$ as was estimated by Vargas-Magaña et al. [159]. Clearly the best option are between the three last estimators.

4.1.4.2 $\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$ estimators

We define the estimator of the scaled counts-in-spheres, $\mathcal{N}(< r)$, using the number density estimators defined in the previous chapter 4.1.4 as:

$$\mathcal{N}_{str}(< r) = \frac{\int_0^r gg(s) ds}{\int_0^r rr(s) ds} = \frac{\frac{1}{G(G-1)/2} \sum_{r_j=0}^r \sum_{i=1}^G n_i^{(G)}(r_j)}{\frac{1}{R(R-1)/2} \sum_{r_j=0}^r \sum_{i=1}^R n_i^{(R)}(r_j)} \quad (4.13)$$

Notice that this definition of $\mathcal{N}(< r)$ in Eq. 4.13 recalls the Peebles-Hauser estimator, $\widehat{\xi}_{ph}(r) = gg(r)/rr(r) - 1$ for the two-point correlation function [165]. Therefore, Laurent et al. [166] developed a Landy-Szalay inspired estimator for the counts-in-spheres:

$$\widehat{\mathcal{N}}_{lau}(< r) = 1 + \frac{\int_0^r [gg(s) - 2gr(s) + rr(s)] ds}{\int_0^r rr(s) ds} . \quad (4.14)$$

They have shown that the aforementioned estimator has minimal variance on scales where $\xi(r) \ll 1$ in analogue of the $\sigma_{\xi_{ls}}(r) \ll \sigma_{\xi_{ph}}(r)$, as shown in table 4.1.

Alternatively, we can directly estimate the counts-in-spheres from the Landy-Szalay estimator of the two-point correlation function itself, as explained in chapter 4.1.2, by only using an estimator for the 2ptCF, i.e. using Eq. 4.6. Thus we developed an alternative estimator of the scaled-count-in spheres defined by:

$$\widehat{\mathcal{N}}_{cor}(< r) = 1 + \frac{3}{r^3} \int_0^r \widehat{\xi}_{ls}(s) s^2 ds . \quad (4.15)$$

We expect that these two estimators \mathcal{N}_{lau} and $\mathcal{N}_{cor}(< r)$ are expected to be closer to optimal than the estimator $\mathcal{N}_{str}(< r)$.

The aforementioned estimator result in three different estimators for the fractal dimension using Eq. 4.3. We name them accordingly $\widehat{\mathcal{D}}_{2,str}(r)$, $\widehat{\mathcal{D}}_{2,lau}(r)$ and $\widehat{\mathcal{D}}_{2,cor}(r)$. In section 4.3.9, we show that in our final analysis, unless stated otherwise for comparison, we use the *cor*-estimators and we drop the *lau* indices and the hat for sake of simplicity. We summarise all different estimators in table 4.2.

Name	\mathcal{N} -Estimator	\mathcal{D}_2 -Estimator
str	$\mathcal{N}_{str}(< r) = \frac{\int_0^r gg(s) ds}{\int_0^r rr(s) ds}$	$\mathcal{D}_{2,str}(r) = 3 + \frac{d \ln}{d \ln r} [\mathcal{N}_{str}(< r)]$
lau	$\mathcal{N}_{lau}(< r) = 1 + \frac{\int_0^r [gg(s) - 2gr(s) + rr(s)] ds}{\int_0^r rr(s) ds}$	$\mathcal{D}_{2,lau}(r) = 3 + \frac{d \ln}{d \ln r} [\mathcal{N}_{lau}(< r)]$
cor	$\mathcal{N}_{cor}(< r) = 1 + \frac{3}{r^3} \int_0^r \xi_{ls}(s) s^2 ds$	$\mathcal{D}_{2,cor}(r) = 3 + \frac{d \ln}{d \ln r} [\mathcal{N}_{cor}(< r)]$

TABLE 4.2: The three different estimators for the *scaled count-in-spheres*, $\mathcal{N}(< r)$ and *fractal correlation dimension*, $\mathcal{D}_2(r)$, studied in this thesis.

To implement our estimators we used:

- The trapezoidal rule[167] for the integration of the estimator $\mathcal{N}(< r)$.
- The second order central difference rule[168] for the derivative of the estimator $\mathcal{D}_2(r)$.

Finally for all those estimators we apply the definition of Homogeneity scale chapter 4.2.3, using equation Eq. 4.7 and Eq. 4.8, respectively.

4.1.5 Theoretical modelling

Our theoretical model is obtained by solving the Einstein-Boltzmann equation derived to describe the dynamic behaviour of the energy fluid. We use the latest software dedicated for those studies,

Cosmic Linear Anisotropy Solving System (CLASS) [169]. CLASS software like CAMB (Code for Anisotropies in the Microwave Background [170]), is a software dedicated to provide the predictions for observables from the standard Λ CDM or any other flavour of Λ CDM-models, such as sCDM (standard Cold Dark Matter without dark energy), w0wa Λ CDM(Λ CDM-model with a time-dependent equation of state), etc.. The predicted observables are the auto and cross power spectra of 2nd and 3rd order decompositions in 2 and 3 dimensions.

4.1.5.1 Theory for $\xi(r)$, $\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$ observables

We use CLASS to compute the theoretical matter power spectrum, $P_{mat}^{(r)}(k)$, where the (r) exponent indicates that it is calculated in real space, by opposition to redshift space, for any fixed fiducial cosmology. Here the fiducial cosmology we use is defined via Eq. 3.4. In order to get a prediction for the redshift-space galaxy power-spectrum, $P_{tr}^{(s)}(k) = P_{gal}^{(s)}(k)$ where the (s) exponent indicates redshift space, we model RSD as discussed in chapter 2.5.2.

We integrate over all orientations relative to the line-of-sight, μ variable, the redshift-space galaxy power spectrum reads:

$$P_{gal}^{(s)}(k; b, \sigma_p) = b^2 \int_0^1 (1 + \beta\mu^2)^2 D_G(k, \mu; \sigma_p) d\mu P_{mat}^{(r)}(k), \quad (4.16)$$

We can analytically integrated over μ the above function as:

$$P_{gal}^{(s)}(k; b, \sigma_p) = P_{mat}^{(r)}(k) \times \frac{b^2 H_0}{2k^5 \sigma_p^5} \times \left[-2 \exp\left(-\frac{k^2 \sigma_p^2}{2H_0^2}\right) k \beta \sigma_p H_0 (k^2 (2 + \beta) \sigma_p^2 + 3\beta H_0^2) + \sqrt{2\pi} \text{Erf}\left(\frac{k \sigma_p}{\sqrt{2} H_0}\right) (k^4 \sigma_p^4 + 2k^2 \beta \sigma_p^2 H_0^2 + 3\beta^2 H_0^4) \right] \quad (4.17)$$

and $\beta = \beta(b; \gamma = 0.55)$ defined by Eq. 2.10. Applying a Fourier Transform to equation (4.17) results in the two-point correlation function with two parameters b and σ_p :

$$\xi^{(s)}(r; b, \sigma_p) = FT \left[P_{gal}^{(s)}(k; b, \sigma_p) \right]. \quad (4.18)$$

where notice that the free parameters are (b, σ_p) . Then, using Eq. 4.15 and Eq. 4.3 we compute the Λ CDM prediction for $\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$ for the fiducial cosmology Eq. 3.4.

4.1.5.2 Bias and RSD corrections

One may, however, question the strict necessity of including these redshift space distortion corrections in our large scale homogeneity analysis. To address this question, we have made a simple test: we calculate the ratio between the theoretical homogeneity scale (as defined in Eq. 4.7) with or without accounting for redshift space distortions. We use a range of values for the RSD parameters around the typical ($b \approx 1.95$, $\sigma_p \approx 300 h \text{ km/s}$), [171]. This percentage ratio is shown in Fig. 4.1 for our second redshift bin (for which the effect is expected to be the largest) and turns out to be ranging between 15 and 25% depending on the value of the galaxy bias. We observe no significant effect on the ratio regarding the galaxy velocity dispersion. Although, we

need to take it into account because it is degenerated with bias parameter. The redshift space distortion effects are therefore not negligible when comparing the homogeneity scale for data and Λ CDM predictions, which justifies the use of this model for our analysis.

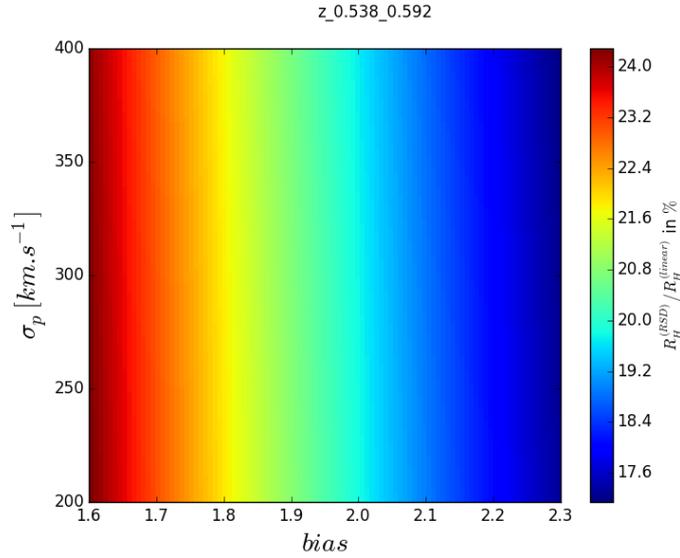


FIGURE 4.1: Test of necessity of RSD modelling: We show the percentage ratio of the predicted homogeneity scale defined in Eq. 4.7 with and without accounting redshift space distortions according to Eq.4.17. The effect ranges from 15 to 25% and varies according to the bias parameter and is therefore not negligible.

The estimators presented in chapter 4.1.4 measure the clustering properties of the galaxy distribution, we need to convert them into estimators that describe the clustering properties of the total matter distribution. We fit $\xi(r)$ in the range $1 h^{-1} < r < 40 h^{-1}$ Mpc with our model of equation 4.18 to obtain b and σ_p . If we take into account only the bias and the Kaiser effect, the angle-averaged correlation function is just multiplied by a constant factor, a squared effective bias. Since this factor is independent of r , the same factor applies to $\mathcal{N}(< r) - 1$, which is an integral over $\xi(r')$ (see equation 4.15). Taking into account also finger-of-God effect, the multiplicative factor is no longer independent of r . To transform the estimated $\mathcal{N}(< r) - 1$ for galaxies into an estimation for matter, we multiply it by the ratio of our model for $\mathcal{N}(< r) - 1$ for matter (i.e. $b = 1$ and $\sigma_p = 0$) to our model for the best fit value of b and σ_p , as

$$\widehat{\mathcal{N}}(< r) = \frac{\mathcal{N}^{\text{model}}(< r; b = 1, \sigma_p = 0) - 1}{\mathcal{N}^{\text{model}}(< r; b, \sigma_p) - 1} \times [\widehat{\mathcal{N}}_{gal}(< r) - 1] + 1. \quad (4.19)$$

Then the fractal correlation dimension $\widehat{\mathcal{D}}_2(r)$ is obtained from $\widehat{\mathcal{N}}(< r)$ using equation 4.3.

Section 4.3.6 shows that taking into account the finger-of-God effect only contributes in a small change in the measurement of the homogeneity scale by typically 1%. So the error due to the imperfection in the modelling of this effect is negligible.

Furthermore, the QPM mock catalogs were generated with a different fiducial cosmology than the CMASS galaxy distribution which are in the "true" cosmology, i.e. the recently measured from Planck 2015. Therefore, we account for the difference in cosmology in two ways. Firstly,

to account for the effect on the y-axis of Fig. 4.3 we use:

$$\widehat{\mathcal{N}}_{qpm}^{corrected}(< r) = \frac{\mathcal{N}^{\text{model}}(< r; b_{qpm}, \sigma_p^{qpm}, p_{\text{cosmo}}) - 1}{\mathcal{N}^{\text{model}}(< r; b_{qpm}, \sigma_p^{qpm}, p_{qpm}) - 1} \times [\widehat{\mathcal{N}}_{qpm}(< r) - 1] + 1. \quad (4.20)$$

where the b_{qpm} and σ_p^{qpm} are the mean values of the 1000 QPM mock catalogs estimated for each individual mock with the procedure explained in chapter 4.2.2. Secondly, to account for the effect on the x-axis, we multiply the scale r with:

$$a_{qpm,corr}(z_m) = \frac{d_V(z_m; p_{\text{cosmo}})}{d_V(z_m; p_{qpm})} \quad (4.21)$$

where z_m is the mean redshift of each slice bin described in table 3.1 and $d_V(z; \mathbf{p})$ is the volume distance defined via:

$$d_V(z; \mathbf{p}) = [czH^{-1}(z; \mathbf{p})(1+z)^2 d_A^2(z; \mathbf{p})]^{1/3} \quad (4.22)$$

where $H(z; \mathbf{p})$ is the hubble expansion rate as a function of redshift defined via Eq. 1.27 and $d_A(z; \mathbf{p})$ is the angular diameter distance defined via Eq. A.4.

4.1.6 Core Analysis Algorithm

In summary, the scheme in Fig. 4.2 describes briefly the overall procedure to measure the homogeneity scale. Firstly, we perform a mapping between (z, θ, ϕ) -coordinates to (r, θ, ϕ) -coordinates for every object on the three different catalogs we use in our analysis, using a Λ CDM-model, i.e. using Eq. 1.31 in the fiducial cosmology. This introduces a mapping between (z, θ, ϕ) -coordinates to cartesian comoving coordinates, (x, y, z) . This allows us to reconstruct the 3Dimensional catalogs of the three different types of catalogs, i.e. galaxies, randoms and mocks. Then, we measure the mean pairs of objects at different separations between the same catalog and between different catalogs, i.e. galaxy-galaxy catalog, $GG(r)$ the galaxy-random catalog, $GR(r)$, and random-random catalog, $RR(r)$. By selecting an estimator for our observables, i.e. a mathematical combination of the pair-counts, we estimate the 2pt Correlation function, ξ , the scaled count-in-spheres, \mathcal{N} , fractal correlation dimension, D_2 , as a function of separation of objects, r, for Galaxy-Randoms combination and Mock-Randoms combination. The first combination, i.e. Galaxy-Random estimators, allows us to the model dependent data estimators as a function of separations r. The second combination, i.e. the Mock-Random estimators, allows us to build the covariance matrices for the different estimators, C_{ij}^X where $X = \{\xi, \mathcal{N}, D_2\}$. Then performing a spline fitting in the separation around the homogeneity scale, using a definition that depends on our observables as discussed in chapter 4.2.3, we estimate the Homogeneity scale of the galaxy distribution. The error of the homogeneity scale is transferred by error propagation from the abscissa to the ordinate. By using a RSD model, we determine the bias and peculiar velocities parameter, (b, σ_p) to obtain an estimate of the clustering bias between luminous and total matter of the universe. This allows us to determine the scaled count-in-spheres, \mathcal{N} and fractal correlation dimension, D_2 for the matter distribution and therefore from a spline fitting acquire the Homogeneity scale for the matter distribution. We note that we can use either count in spheres or fractal dimension to determine the homogeneity scale, using Eq. 4.8 or Eq. 4.7, respectively. Note, that we do not use the 2ptCF, since we cannot define mathematically a %

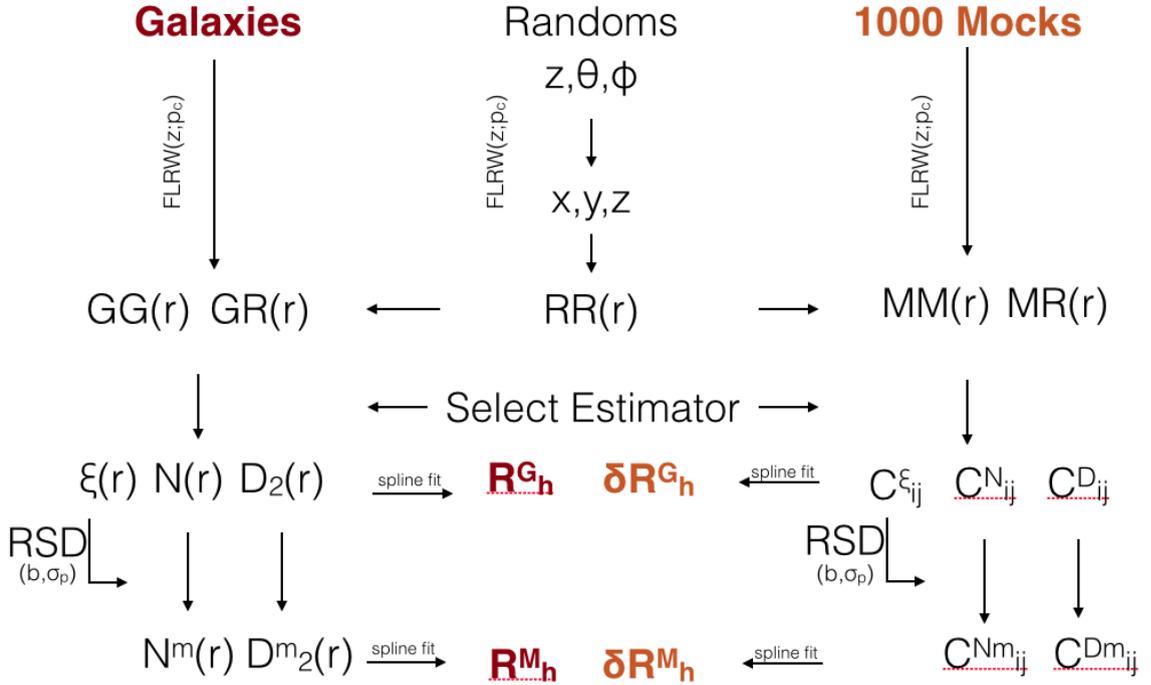


FIGURE 4.2: Core analysis pipeline procedure. See text for detailed explanation

difference from the nominal value of the 2ptCF, i.e. the value of a homogeneous distribution which is 0.

4.2 Results

In this section we discuss the basic results with the choices that we have made for measuring the homogeneity scale for the galaxy distribution and the matter distribution, as long as estimates of this observable against the cosmic depth of our survey, redshift z on the North Galactic Cap and South Galactic Cap.

4.2.1 Homogeneity scale for the galaxy distribution

We first compute $\mathcal{N}^{gal}(< r)$ and $\mathcal{D}_2^{gal}(r)$ for the CMASS galaxy distribution using equation 4.15 and 4.3, without correcting for bias and redshift space distortions. This provides a measurement that does not rely on Λ CDM model to determine the bias. Results are shown in figure 4.3 for the $0.430 \leq z \leq 0.484$ redshift interval in the North Galactic Cap. As described in chapter 4.3.2, the error are obtained from the 1000 QPM mock catalogs. The estimators of QPM mock catalogs are treated with Eq. 4.20 as discussed at the end of chapter 4.1.5.2.

The horizontal black-dashed lines in figure 4.3 indicate the value of the corresponding observable for an homogeneous distribution. The two observables reach homogeneity on large scales. The intersection of the data with the red-dashed lines (at $\mathcal{N} = 1.01$ and $\mathcal{D}_2 = 2.97$, nominal values) defines the homogeneity scale. In order to determine it, we perform a spline fit over 6 data

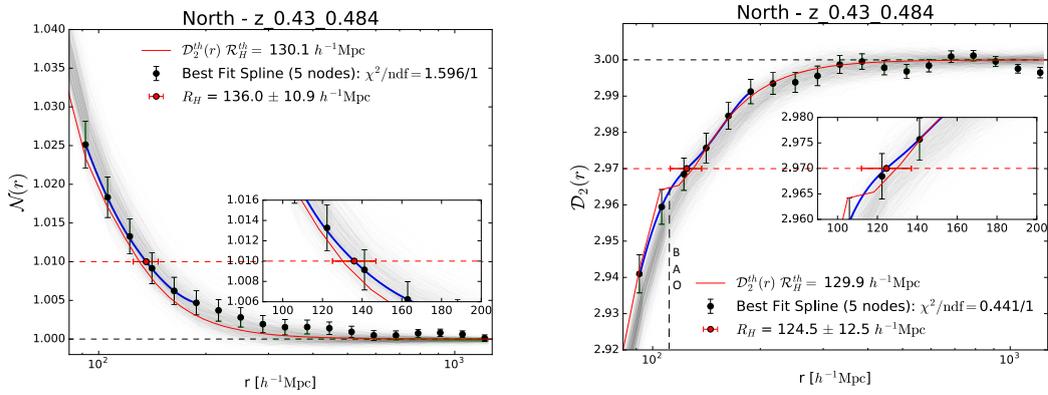


FIGURE 4.3: Scaled counts-in-spheres, $\mathcal{N}(<r)$, (left) and fractal correlation dimension, $\mathcal{D}_2(r)$, (right). The black points with error bars are the result for the galaxy distribution in the NGC in the $0.430 \leq z \leq 0.484$ redshift bin. The blue lines are the best spline fit of the data. The green lines are the results for the 1000 QPM mock catalogs. The red continuous line is the Λ CDM prediction for $b = 1.95$ and $\sigma_p = 252$ km/s. The vertical dash line shows the BAO feature on the fractal correlation dimension computation.

point around the intersection. The number of nodes of the spline fit is set to 5 in order to get an average χ^2 equal to the number of degrees of freedom for the 1000 QPM mocks. The error on R_H was obtained from the errors in the spline factors by error propagation, as described in chapter A.12 where we consider perturbations of 1%. The error propagation is the ratio between an estimation on the errors of the coefficient of the spline with the slope of the function $\mathcal{N}(<r)$ or $\mathcal{D}_2(r)$ at the nominal values. The results, presented in Table 4.4 for the case of \mathcal{D}_2 , are consistent with our Λ CDM model predictions. We finally stress the fact that the behaviour of the data is consistent with that of the 1000 QPM mock catalogs, which are denoted with green color lines on the figure.

4.2.2 Determining the bias and the velocity dispersion

As illustrated in figure 4.4, we fit our model (Eq. 4.18) to the measured correlation function for the CMASS galaxy sample, in order to determine the galaxy bias, b , and velocity dispersion σ_p for the redshift slice $0.538 \leq z \leq 0.592$. As explained in chapter 4.3.7, we account for theoretical uncertainties in the RSD modelling by boosting the covariance matrix on the relevant small scales with the eff-parameter. We thus achieve a good $\chi^2/\text{n.d.f.}$ in the range $[1, 40]$ h⁻¹ Mpc. In chapter 4.3.7 it is also shown that this change in the covariance matrix at small scales has a negligible effect on the measurement of the homogeneity scale.

Results are given in table 4.3. The mean precision is 2.6% (NGC) and 4.1% (SGC) for the bias and 12% (NGC) and 23% (SGC) for the velocity dispersion. The bias is in agreement with values obtained by several authors [172–177] and the velocity dispersion is consistent with the typical CMASS-galaxy-sample velocity-dispersion, $\sigma_p \approx 240 \pm 50$ km/s [171].

Notice in Fig. 4.4 that at large scale of the 2ptCF, we also detect the comoving BAO peak position. One can observe the bao peak position at about $110h^{-1}\text{Mpc}$.

z	NGC			SGC		
	b	σ_p [km/s]	χ^2_{red}	b	σ_p [km/s]	χ^2_{red}
0.430-0.484	1.879 ± 0.023	243.0 ± 9.7	0.44	1.872 ± 0.035	238.7 ± 15.2	0.85
0.484-0.538	1.846 ± 0.021	247.1 ± 7.4	0.67	1.840 ± 0.032	234.6 ± 12.1	0.66
0.538-0.592	1.944 ± 0.021	252.0 ± 8.9	0.91	1.943 ± 0.032	236.2 ± 13.7	0.59
0.592-0.646	1.995 ± 0.022	234.2 ± 10.9	1.47	2.001 ± 0.035	237.9 ± 19.6	0.79
0.646-0.700	2.153 ± 0.028	235.3 ± 21.7	0.79	2.081 ± 0.045	210.3 ± 40.1	0.50

TABLE 4.3: Fitted values of bias, b , and velocity dispersion, σ_p , in the different redshift bins, together with the corresponding reduced χ^2 for 24 degrees of freedom.

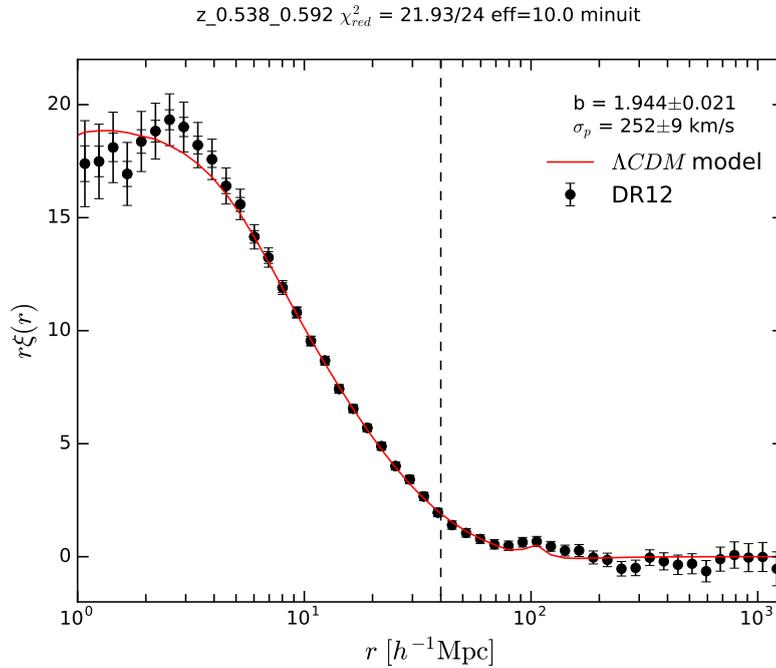


FIGURE 4.4: The correlation function of the CMASS galaxies in the $[0.538 - 0.592]$ redshift range. Data points have both their original error bars and the error bars enlarged to take into account the uncertainty of the RSD model on small scales (see section 4.2.2 and appendix 4.3.7). The red line is the result of the fit performed over the range $[1, 40] h^{-1}\text{Mpc}$ (up to the vertical black-dashed line).

4.2.3 Homogeneity scale for the matter distribution

We use Eq. 4.19 with parameter b and σ_p obtained in section 4.2.2 to transform $\mathcal{N}^{gal}(< r)$ into $\mathcal{N}(< r)$ for matter. We then use Eq 4.3 to get $\mathcal{D}_2(r)$ for matter distribution. Results are shown in Fig. 4.5 for the redshift interval $0.538 \leq z \leq 0.592$. The two observables indicate homogeneity on large scales in this redshift interval, and in the four other intervals as well. As in section 4.2.1, we fit the data points to determine the homogeneity scales. We stress that the fit range, $40 < r < 100 h^{-1}\text{Mpc}$, does not overlap with the fit range for determining the bias, $1 < r < 40 h^{-1}\text{Mpc}$.

The results, presented in Table 4.4 for the case of \mathcal{D}_2 , are consistent with our ΛCDM model predictions. In the redshift interval $0.538 \leq z < 0.592$ we get a precision of 1.6%, which is a factor 5 better than Scrimgeour et al. [93] in spite of their wider redshift range, $0.5 < z < 0.7$. Note that the more recent analysis by Sarkar et al. [178] does not give R_H for matter distribution.

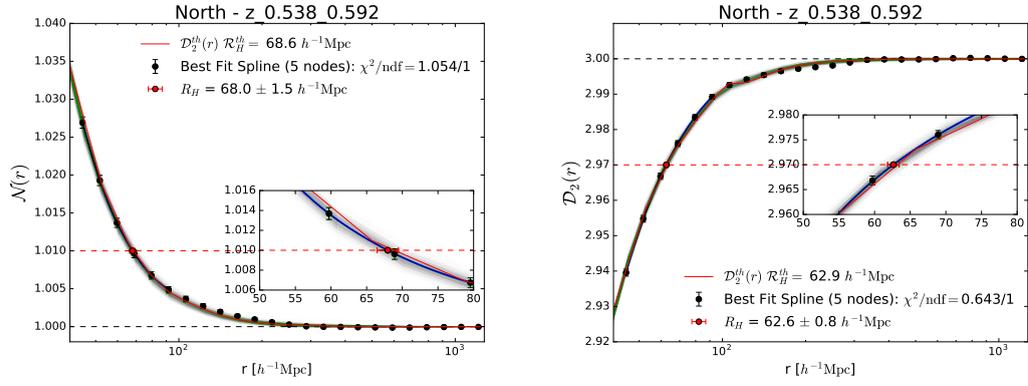


FIGURE 4.5: Same as figure 4.3 for matter distribution in the redshift interval $0.538 \leq z \leq 0.592$.

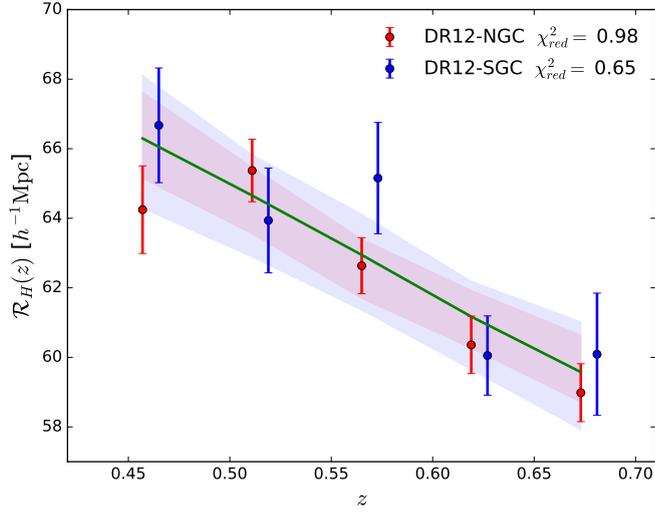


FIGURE 4.6: The homogeneity scale $R_H^{\mathcal{D}_2=2.97}(z)$ measured in the NGC (red) and in the SGC (blue) as a function of redshift. The green line is the Λ CDM model prediction. The shaded areas indicate the 1 σ range for the 1000 QPM mock catalogs.

Furthermore, we perform the measurement in the 5 redshift linear bins as are defined table 3.1, so that to study the time evolution of the homogeneity scale. Figure 4.6 shows that the measured homogeneity scale is compatible with Λ CDM, with reduced χ^2 smaller than unity both in the NGC and the SGC. The scale is increasing with time, from $z = 0.7$ to $z = 0.43$, as expected when clustering is increasing with time. The reduced χ^2 remains smaller than unity when adding the result from Laurent et al. [166], which results in homogeneity scale, $\mathcal{R}_H(z = 2.4) = 26.2 \pm 0.9$ h⁻¹Mpc. The reduced χ^2 is $\chi_{red}^2 = 0.98$ for the North galactic cap and $\chi_{red}^2 = 0.65$ for the South galactic cap.

z	$\mathcal{R}_H[h^{-1}\text{Mpc}]$ Galaxy		$\mathcal{R}_H[h^{-1}\text{Mpc}]$ Matter	
	NGC	SGC	NGC	SGC
0.430 – 0.484	124.5 ± 12.5	121.1 ± 9.8	64.2 ± 1.3	66.7 ± 1.6
0.484 – 0.538	111.9 ± 4.9	119.8 ± 8.8	65.4 ± 0.9	63.9 ± 1.5
0.538 – 0.592	116.4 ± 7.8	110.5 ± 5.1	62.6 ± 0.8	65.2 ± 1.6
0.592 – 0.646	108.8 ± 3.9	120.1 ± 11.7	60.4 ± 0.8	60.1 ± 1.1
0.646 – 0.700	125.8 ± 7.3	147.4 ± 8.4	59.0 ± 0.8	60.1 ± 1.8

TABLE 4.4: Homogeneity scale, $R_H^{\mathcal{D}_2=2.97}(z)$, for the galaxy and matter distributions in the north and south galactic caps.

4.2.4 Consistency with Λ CDM

In order to characterize more precisely the agreement with the Λ CDM model, we fit $\mathcal{D}_2(r)$ from the data and the 1000 mock catalogs in the range $40 < r < 1300 h^{-1}\text{Mpc}$ with

$$\mathcal{D}_2^{\Lambda\text{CDM}} = \mathcal{D}_2(ar; \mathbf{p}_{\text{cosmo}}). \quad (4.23)$$

We fix the cosmological parameters $\mathbf{p}_{\text{cosmo}}$ and leave only free the *dilatation* parameter a .

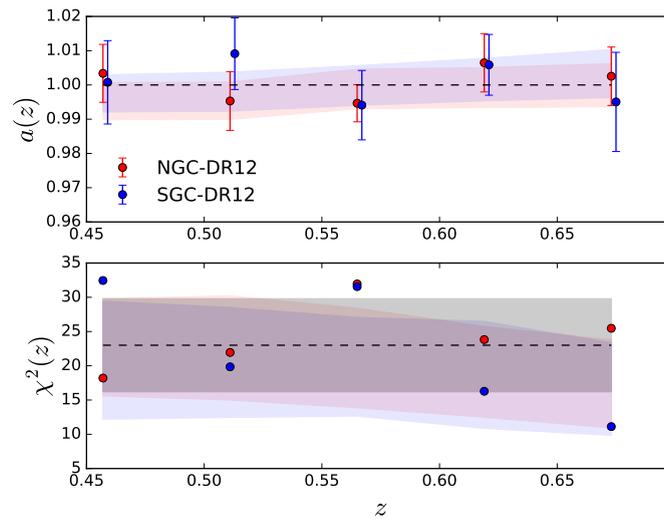


FIGURE 4.7: Top: value of the *dilatation* parameter a resulting from the fit of the NGC (red) and SGC (blue) $\mathcal{D}_2(r)$ data with the model of equation 4.23, in different redshift bins. The shaded areas correspond to the 1σ region for the 1000 QPM mock catalogs. Bottom: The corresponding χ^2 .

Figure 4.7 shows the results of the fits in different redshift bins for both NGC and SGC. The values of a , all consistent with 1, demonstrate a good agreement with Λ CDM model. Furthermore, the bottom plot of the figure shows that, for the data, the χ^2 are consistent with the number of degrees of freedom (23), as illustrated by the grey shaded area that indicates the expected 1σ extension, $\chi^2 = 23 \pm \sqrt{2 \times 23}$. The mock catalogs, however, are a bit at variance with this grey area at large z . We summarise the results in table 4.5

z	NGC		SGC	
	a	χ^2	a	χ^2
0.430 – 0.484	1.003 ± 0.008	18.2	1.001 ± 0.012	32.4
0.484 – 0.538	0.995 ± 0.009	21.9	1.009 ± 0.010	19.8
0.538 – 0.592	0.995 ± 0.005	31.9	0.994 ± 0.010	31.6
0.592 – 0.646	1.006 ± 0.009	23.8	1.006 ± 0.009	16.3
0.646 – 0.700	1.003 ± 0.009	25.5	0.995 ± 0.015	11.1

TABLE 4.5: Mean, error and χ^2 of the consistency test fit given by Eq. 4.23 with 23 degrees of freedom, in the NGC and SGC for the five redshift bins.

4.2.5 Constraints on fractal correlation dimension

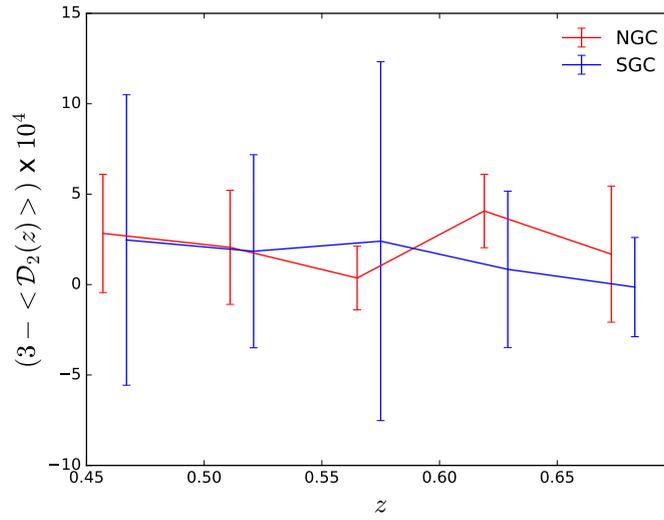


FIGURE 4.8: Fractal correlation dimension for the matter distribution, $3 - \langle \mathcal{D}_2 \rangle$, averaged over $300 h^{-1} < r < 1300 h^{-1}$ Mpc, in the different redshift bins.

z	Matter		Galaxy	
	NGC ($\times 10^{-4}$)	SGC ($\times 10^{-4}$)	NGC ($\times 10^{-3}$)	SGC ($\times 10^{-3}$)
0.430 – 0.484	2.8 ± 3.3	2.5 ± 8.1	1.3 ± 1.5	1.1 ± 3.7
0.484 – 0.538	2.0 ± 3.1	1.9 ± 5.4	0.9 ± 1.4	0.8 ± 2.4
0.538 – 0.592	3.6 ± 1.7	2.4 ± 9.9	0.2 ± 0.9	1.2 ± 4.9
0.592 – 0.646	4.1 ± 2.0	0.9 ± 4.5	2.1 ± 1.1	0.5 ± 2.3
0.646 – 0.700	1.7 ± 3.8	-0.2 ± 2.8	1.0 ± 2.3	0.1 ± 1.6

TABLE 4.6: Fractal correlation dimension, $3 - \langle \mathcal{D}_2 \rangle_{r,z}$, averaged over $300 < r < 1300 h^{-1}$ Mpc, with 1σ errors, in the NGC and SGC.

In order to assess the level of homogeneity of the CMASS DR12 galaxy sample we compute the average of $\mathcal{D}_2(r)$ over the range $300 < r < 1300 h^{-1}$ Mpc, accounting for the covariance matrix, i.e. computing the weighted average. The results are presented in table 4.6 for the different redshift bins in the NGC and SGC. All results are compatible and we average them to get $3 - \langle \mathcal{D}_2 \rangle_{r,z} = (0.9 \pm 1.2) \times 10^{-3}$ (1σ).

Strictly speaking, it is not possible to transform the results for the galaxy distribution into results for matter distribution without using Λ CDM prediction to compute the galaxy bias. However, Laurent et al. [166] show that, with a set of reasonable assumptions independent of Λ CDM, it is possible to obtain a lower limit for the tracer bias. In particular, these set of assumptions are as it follows. Firstly, consider that weak lensing is a direct measurement of the total matter in the universe. Bacon et al. [179] have measured the clustering of matter only through weak lensing data and compare it to the Λ CDM prediction and they found an amplitude of the clustering, $\xi_{WL}(z = 0.2) \leq 1.6\xi_{matter}(z = 0.2)$ at redshift $z \sim 0.2$. But we measure a model dependent bias for the cmass sample $b_{CMASS}^2 = \xi_{CMASS}(z = 0.5)/\xi_{matter}(z = 0.5) \simeq 2^2$. Assuming now that the clustering does not decrease with time, $\xi_{matter}(z = 0.2) \simeq \xi_{matter}(z = 0.5)$. Comparing the clustering of CMASS galaxy sample to the weak lensing sample we have an estimate of the bias of the cmass sample in respect to an estimate of the clustering of the total matter, $(b'_{cmass})^2 = \xi_{CMASS}(z = 0.5)/\xi_{WL}(z = 0.2) \geq 4/1.6$. Notice that we can do this measurement for any assumed A_s amplitude of the scalar perturbations parameter, as described in section 1.2.7, rendering this new bias measurement "model independent". In our case this results in $b'_{CMASS} > 1.6$ and therefore, we have an upper limit for the fractal correlation dimension for the total matter distribution, $3 - \langle \mathcal{D}_2 \rangle < 7.12 \pm 9.48 \times 10^{-4}$.

Alternatively, we can assume Λ CDM, use it to get the galaxy bias and compute $\mathcal{D}_2(r)$. We average $\mathcal{D}_2(r)$ over $300 < r < 1300 h^{-1}\text{Mpc}$ and give the results in table 4.6 and in figure 4.8. Since the galaxy bias is significantly larger than one, the constraints are tighter for matter distribution than for galaxy distribution. Averaging over redshift and caps we get $3 - \langle \mathcal{D}_2 \rangle_{r,z} = (1.7 \pm 1.0) \times 10^{-4} (1\sigma)$ for the matter distribution. This is a strong consistency check of Λ CDM.

4.3 Selection Criteria and Systematics

In this part, we are going to describe all the choices we have made to measure the homogeneity scale. In particular we are going to describe the range of scales that we are considering for this analysis, the estimation of the covariance matrix and the precision matrix, the test we performed for the selection of the different estimators for our observables. We discuss the selection of estimator against the others through a qualitative test. Finally, we are going to discuss the systematics studied for the homogeneity measurement.

4.3.1 Scale-Range cuts

Figure 4.9 shows the number density of random pairs, dN/dr , divided by r^2 and normalized such that at small scales it is unity. This scaled number density is constant at small scales and then decreases with r due to the finite size of the survey. We analyse data up to a maximum $r = 1300 h^{-1}\text{Mpc}$ with a logarithmic binning of 50 bins. We chose this cut since the scaled density goes down to 1%.

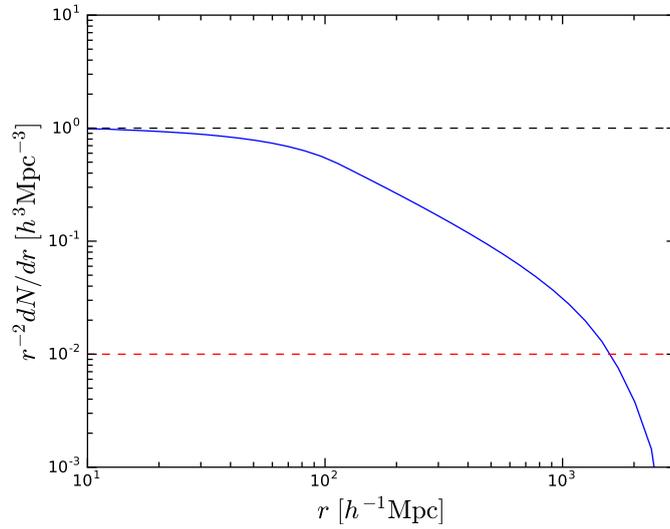


FIGURE 4.9: Scaled number density of random pairs, $r^{-2}dN/dr$, versus the comoving radius of spheres, r , normalized to be unity at small r .

4.3.2 Covariance and Precision Matrices

We use a set of $N_{mock} = 1000$ QPM mock catalogs¹ to estimate the covariance matrices for all possible estimators. We compute our observables for each of the catalogs and derive the bin-to-bin covariance matrices for each of the relevant observables, $\xi(r)$, $\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$. The covariance matrices are given simply by:

$$C_{ij} = \frac{1}{N_{mock} - 1} \sum_{k=1}^{N_{mock}-1} (\mathcal{O}_k(r_i) - \langle \mathcal{O}(r_i) \rangle) \cdot (\mathcal{O}_k(r_j) - \langle \mathcal{O}(r_j) \rangle) \quad (4.24)$$

where $\mathcal{O}_k = \{\mathcal{D}_2^{mat}, \mathcal{D}_2^{gal}, \mathcal{N}^{mat}, \mathcal{N}^{gal}, \xi\}$ and we used the notation for the average all over the mocks at different scales as: $\langle \mathcal{O}(r) \rangle = \frac{1}{N_{mock}} \sum_{k=1}^{N_{mock}} \mathcal{O}_k(r)$. However, Taylor et al. [180] have shown that this estimator of covariance shifts the estimator of the precision matrix, $\psi_{ij} = (C^{-1})_{ij}$ that we use on our fitting methods on our analysis. The corrected precision matrix has been shown to be given by :

$$\psi_{ij}^c = F \cdot \psi_{ij} \quad (4.25)$$

where $F = \frac{N_{mock} - N_{bins} - 2}{N_{mock} - 1}$. In the case where $(N_{bins}, N_{mock}) = (50, 1000)$ the correction is $F \simeq 0.949$ which is an important factor in precision cosmology. In this analysis, we present the correlation matrices since they have less dynamics than the covariance matrices. The correlation matrix is given by:

$$r_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} * C_{jj}}} \quad (4.26)$$

Fig. 4.10 displays the resulting correlation matrices, in the redshift bin $0.538 \leq z \leq 0.592$. The correlations matrices are similar in the other redshift bins.

¹The generation of the mock catalogs was discussed in Fig. 3.4.3.

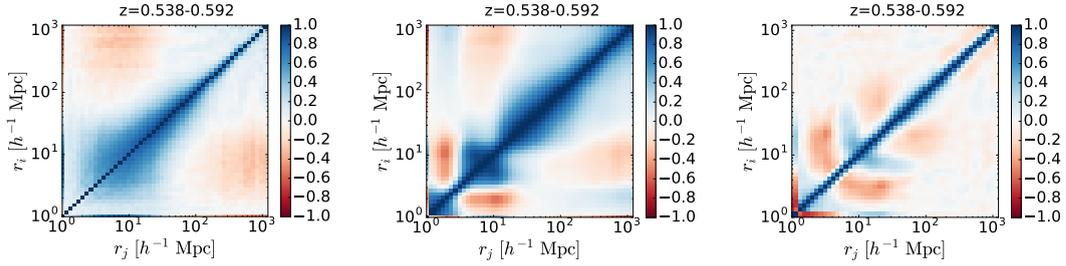


FIGURE 4.10: Correlation matrices for our estimators $\xi(r)$, $\mathcal{N}(<r)$ and $\mathcal{D}_2(r)$ for the total matter distribution estimators in the north galactic cap at the redshift slice $0.538 \leq z \leq 0.592$.

$\mathcal{N}(r)$ is more correlated on large scales than $\xi(r)$ because it is an integral over $\xi(r)$. On the other hand $\mathcal{D}_2(r)$, which is a derivative of $\mathcal{N}(r)$, is not very correlated. Therefore, when studying the redshift evolution in chapter 4.2.3, we consider only the homogeneity scale obtained with $\mathcal{D}_2(r)$, and not with $\mathcal{N}(r)$.

4.3.3 Spline error robustness

In order to make the estimation of the error using the spline error propagation on the homogeneity scale described in chapter A.12 we perform the following test. We measure the homogeneity scale of the matter distribution on each individual QPM mock catalog as described in chapter 4.2.1, i.e. with the spline method. Furthermore, we measure the homogeneity scale of the matter distribution on each individual QPM mock with a simple interpolation method between the data points of fractal correlation dimension, $\mathcal{D}_2(r)$. Additionally, we measure the homogeneity scale on the matter distribution using an MCMC² of a 5 degree polynomial. Then, we compare the standard deviation homogeneity scales with the error estimated with the spline error propagation for the galaxy catalog for all redshift bins for both the North (NGC) and the South (SGC) galactic caps with each different method, e.i. the spline and interpolation methods on 1000 QPM mock catalogues and the MCMC method on the galaxy catalogue. The result are shown in Fig. 4.11. We see that the ratio between the two different estimations is always less than 20% for all redshift bins and the two different galactic caps when we compare the basis measurement, i.e. the spline method, with the spline or interpolation methods on the mock catalogues. However, the MCMC method on the galaxy catalogue gives a ratio of the standard deviations of up to 40%. These results show that the spline method is a conservative method and therefore suitable to use in our analysis.

4.3.4 Normalization Issue between *str cor* estimators

Consider that using the definitions Eq. 4.10, Eq. 4.11 and Eq. 4.12, we construct the estimators for the scaled count-in-sphere and fractal correlation dimension summarised in table 4.2. Notice that:

$$\mathcal{N}^{str}(<r) = \frac{\int_0^r gg(s)ds}{\int_0^r rr(s)ds} = A(R, G) \frac{\int_0^r GG(s)ds}{\int_0^r RR(s)ds} \quad (4.27)$$

²MCMC: Monte Carlo Markov Chain

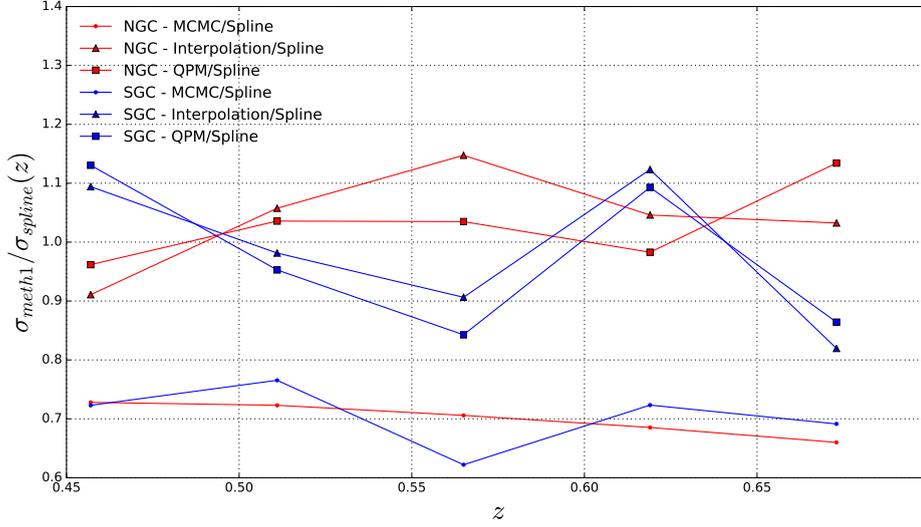


FIGURE 4.11: Ratio of the standard deviations of the homogeneity scale of the matter distribution using 3 methods over the standard deviation of the homogeneity scale using the spline method fit of our basis analysis for the different redshift bins for the NGC (red) and SGC (blue). The dot ('.') denotes the method of MCMC on the data, the ('□') square and ('^') wedge denote the standard deviation of the homogeneity scale on 1000 QPM mock catalogues with the interpolation method and the spline fit respectively.

f where $A(R, G) = \frac{R(R-1)/2}{G(G-1)/2}$. This introduces a shift on the measurement of the scaled counts-in-spheres which depends on R and G . However, on \mathcal{D}_2^{str} -estimator this shift cancels out completely since:

$$\mathcal{N}^{str}(< r) = A(R, G)r^{\mathcal{D}_2^{str}} \rightarrow \ln[\mathcal{N}^{str}(< r)(r)] = \ln[A(R, G)] + \ln[r^{\mathcal{D}_2^{str}}]$$

Then we have:

$$\mathcal{D}_2^{str}(r) = \frac{d \ln[\mathcal{N}^{str}(< r)]}{d \ln r} - \frac{d \ln[A(R, G)]}{d \ln r} \rightarrow \mathcal{D}_2^{str}(r) = \frac{d \ln[\mathcal{N}^{str}(< r)]}{d \ln r}$$

While on $\mathcal{N}^{cor}(< r)$ there is no unique normalisation factor for all pair-counts, and therefore \mathcal{D}_2^{cor} estimator is shifted according to the normalisation factor, which is not the case for the \mathcal{D}_2^{str} . Therefore we use, as a comparison, the str-estimators for scaled-counts-in-spheres and fractal correlation dimension to compute an unbiased estimation of the homogeneity scale. In Fig. 4.12, we observe that indeed the homogeneity is, asymptotically, reached within error bars for the str-estimator for both observable $\mathcal{N}^{gal, str}$ and $\mathcal{D}_2^{gal, str}$ for the galaxy distribution. For the $\mathcal{N}^{gal, str}$ observable there is a small shift on the mean against the nominal value 1.00 at large scales $r \geq 300h^{-1}\text{Mpc}$ but negligible within the 1σ dispersion. While for the $\mathcal{N}^{gal, cor}$ displayed in the left panel of Fig. 4.3, we see a shift from the nominal value which is not negligible. However, for the $\mathcal{D}_2^{gal, cor}$ of the same figure in the right panel, we have no such shift on average from $r \geq 300h^{-1}\text{Mpc}$. In Fig. 4.5 the shift of $\mathcal{N}^{mat, cor}$ vanishes but this is due to the lucky combination of R and G . Furthermore, there is no shift against the value 3 of the $\mathcal{D}_2^{mat, cor}$ for all redshift bins, as we show in Fig. 4.8. We observe also that indeed by performing the same spline fitting method described in chapter 4.1.6, the str-estimator is less precise than the cor-estimators, e.g. $\partial \log \mathcal{R}_{H, cor}^{\mathcal{D}_2} = 1.2\%$ against $\partial \log \mathcal{R}_{H, str}^{\mathcal{D}_2} = 10.3\%$ on the North Galactic Cap

for the redshift slice $0.538 \leq z \leq 0.592$. Notice also that this estimator is also compatible with the Λ CDM-model. So, finally, we conclude that the cor-estimator shift is negligible and that the accuracy gain of this estimator justifies its use in our analysis.

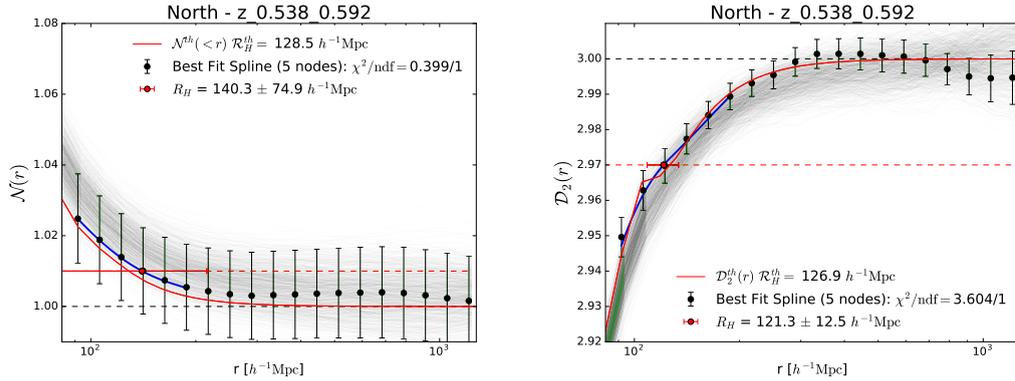


FIGURE 4.12: Same as figure 4.3 for the str-estimators in the redshift interval $0.538 \leq z \leq 0.592$.

4.3.5 Fractal correlation dimension bias

Recall that in this analysis, we have used random, i.e. homogeneous, catalogs to take into account the inhomogeneity of the survey. Thus, we may wonder if this could bias the resulting $\mathcal{N}(<r)$ or $\mathcal{D}_2(r)$ towards homogeneity. To search for such a possible bias, we generate 500 fractal realizations with a given value of the fractal correlation dimension, we pass them through our pipeline analysis and study the resulting \mathcal{D}_2 , as done by Laurent et al. [166].

Following Castagnoli and Provenzale [181], we create a cubic box of $L \approx 4 h^{-1}$ Gpc side, containing the whole survey at $z = 0.538 - 0.592$. We divide this box in $M = n^3$ sub-boxes of size L/n , where $n = 2$ and we give to each sub-box a survival probability p . We then repeat the procedure for each surviving sub-box. An infinite number of iterations would give a fractal distribution with:

$$D_2 = \frac{\log(pM)}{\log(n)} = 3 + \frac{\log p}{\log 2} \quad (4.28)$$

We perform 9 iterations. After the last iteration, we populate each sub-box with random points that follow a Poisson law of mean $\lambda < 1$. The value of λ is chosen so that after the survey cuts the number of objects in the fractal distribution is approximately the same as in our survey. Then we convert the cartesian coordinates to z, RA, DEC with the same FLRW metric as used in our analysis. Finally, we apply cuts to simulate the selection function of our galaxy survey³. Then we reconstruct the fractal distribution in cartesian comoving coordinates to measure the fractal correlation dimension as a function of scales. At the last iteration the size of the sub-box is $15 h^{-1}$ Mpc.

In Fig. 4.13, we calculate the reconstructed fractal correlation dimension for 500 fractal simulated distributions with an input fractal correlation dimension $D_2^{input} = 2.995$. Note that most of the reconstructed values tend to $D_2 = 3$ and some have way lower than the input value. Our simulation method create

³We use the `mangle` software to apply the masks of our survey that is given in `*.ply` format. <http://space.mit.edu/~molly/mangle/>

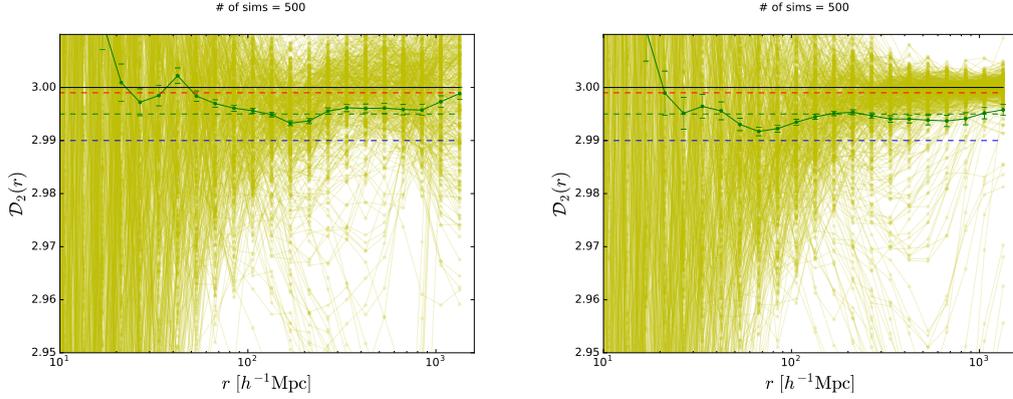


FIGURE 4.13: Reconstructed fractal correlation dimension, \mathcal{D}_2 , at the nominal value 2.995 (green dash line) in the redshift interval $z = 0.538 - 0.592$. Yellow line represent each fractal simulation. Green line is the average and error over 500 simulations for str-estimator (left) and cor-estimator (right).

a fractal distribution all over the cube, while we are reconstructing the observable only on the window of the survey. When the survey window is located to a large cube that did not survive the iteration procedure, we reconstruct a lower value of D_2 than the simulated value. On the other hand, if the sample is restricted to a zone where a small cube has been discarded, a value of D_2 is obtained which is close to 3.

In Fig. 4.14, we plot the average reconstructed fractal correlation dimension for simulated input fractal distribution of constant values $\mathcal{D}_2 = \{2.990, 2.995, 2.999\}$ for the 2 different estimators, $\mathcal{D}_2^{str}(r)$ and $\mathcal{D}_2^{cor}(r)$. For the *str*-estimator the reconstructed values are slightly larger than the input one, while for the *cor*-estimator are slightly lower, but compatible with the input values.

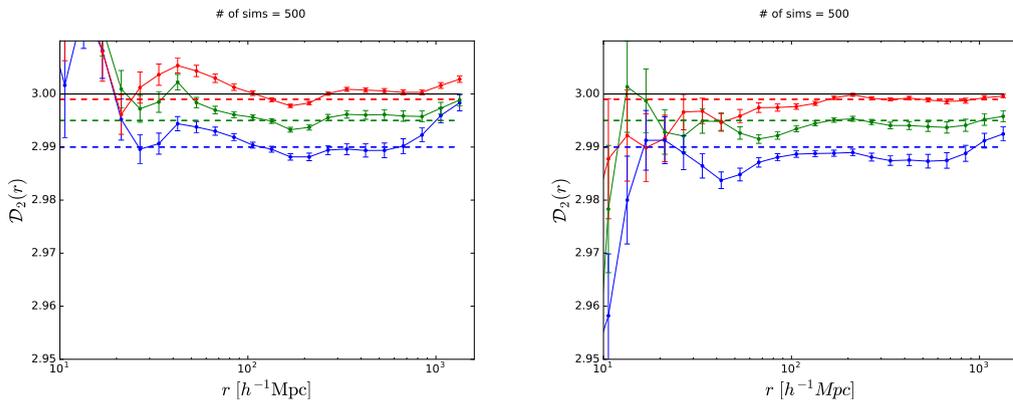


FIGURE 4.14: Reconstructed average value of \mathcal{D}_2 as a function of scales, r , with the estimators $\mathcal{D}_2^{str}(r)$ (left) and $\mathcal{D}_2^{cor}(r)$ (right), for 500 simulations. The errors are obtained with the standard deviation over 500 simulations. The dash line corresponds to the initial input value of \mathcal{D}_2

In order to summarise the above result, we average the reconstructed \mathcal{D}_2 over $r \in [15, 1300] h^{-1}$ Mpc. We obviously select the lower value to be $r = 15h^{-1}$ Mpc, since this is the smallest value of the iterating procedure for the generation of the mono-fractal distribution. Figure 4.15 presents this average value for the reconstructed fractal correlation dimension relative to the different input values $\mathcal{D}_2 = \{2.950, 2.970, 2.990, 2.995, 3.000\}$. Fitting these points with a straight line, $y = \alpha x + \beta$, results in $\alpha = 1.00 \pm 0.03$ and $\beta = (-0.4 \pm 7.9) \times 10^{-2}$ with $\chi_{\text{red}}^2 = 0.9/3$. So, in contrast with Scrimgeour et al. [93], we do not observe any bias in the reconstructed fractal correlation dimension.

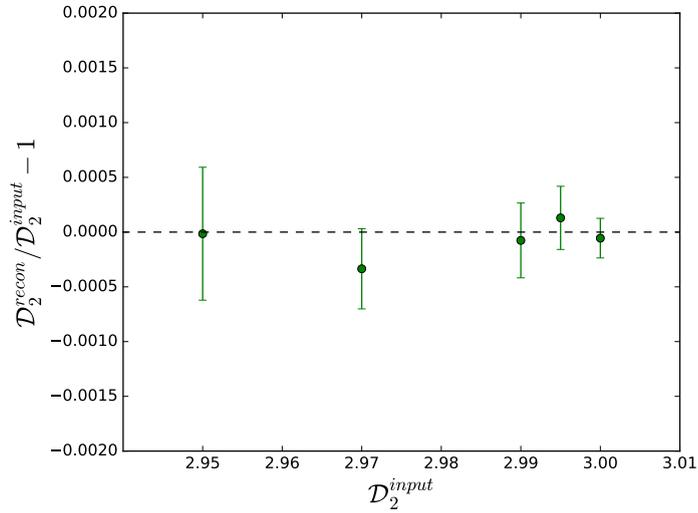


FIGURE 4.15: Reconstructed fractal correlation dimension, \mathcal{D}_2 , in the redshift interval $z = 0.538 - 0.592$, averaged over $r \in [15, 1300] h^{-1}$ Mpc.

4.3.6 RSD model Systematic

We compare the homogeneity scale obtained with two different modelling of redshift space distortions. Namely, the model with a scale-independent bias and the Kaiser effect, see equation (2.11) :

$$\xi(r; b) = FFT [P_{gal}(k, \mu; b)] , \quad (4.29)$$

This model was fitted only in the range $r \in [10 - 40]h^{-1}$ Mpc with only free parameter the cosmic bias, b . Our full model, which includes in addition a modeling of the finger-of-God effect (Eq. 4.16):

$$\xi(r; b, \sigma_p) = FFT [P_{gal}(k, \mu; b, \sigma_p)] \quad (4.30)$$

was fitted for (b, σ_p) parameters in the whole range $r \in [1 - 40]h^{-1}$ Mpc as was explained in chapter 4.2.2.

Figure 4.16 presents the homogeneity scale for these two models in the five redshift bins for the north and south galactic caps. We note that using the full RSD model lowers the χ^2 , confirming that this is a better model than the purely linear Kaiser model. The full model tends to lower R_H , by up to slightly more than 1σ , or typically 1%. The exception is only for the north galactic cap at redshift slice $0.592 \leq z \leq 0.646$ where the mean difference is higher than 1σ . This means that we should not use the purely linear Kaiser model. The remaining systematic error due to the imperfection of our full RSD model is a fraction of the difference between Kaiser and full models, and therefore negligible with respect to our statistical error.

4.3.7 Boosting error Systematics

The theoretical model for redshift space distortions (Eq. 4.16) is not perfectly accurate at the smallest scales due to the nonlinear behaviour of gravity at these scales. In order to ensure satisfying $\chi^2/n.d.f.$

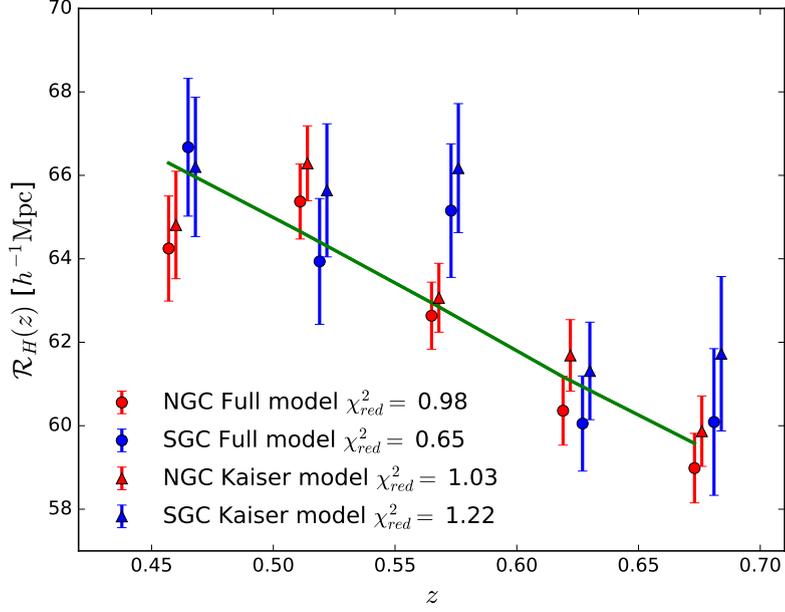


FIGURE 4.16: Measured homogeneity scale in different redshift bins, in the NGC and the SGC, with Kaiser (triangle) and full RSD model (circles).

for the RSD fitting, we boost the error on $\xi(r)$ at the relevant scales in an empirical way as

$$C'_{ij} = C_{ij}(1 + \delta_{ij}\Delta_i\Delta_j e^2). \quad (4.31)$$

Here δ_{ij} is the usual Kronecker symbol, e is a parameter that measures the amount of error boosting we apply, Δ_i is the theoretical inaccuracy, estimated as the relative difference between our model and the average of the 1000 QPM mock catalogs.

In particular, after playing with several functions to achieve good $\chi^2 - ndf$ ratio, we define the correction shape function, f_{cs} as:

$$f_{cs}(r; \vec{p}_{cs}) = 1 + \exp[p_2 r^2 + p_1 r + p_0] \quad (4.32)$$

where $\vec{p}_{cs} = \{p_2, p_1, p_0\}$ are the correction shape parameters. Then, we introduce this function to the theoretical model for the 2ptCF, (Eq. 4.18) as:

$$\xi_{fit}(r; b, \sigma_p, \vec{p}_{cs}) = f_{cs}(r, \vec{p}_{cs}) \cdot \xi_{gal}(r; b, \sigma_p) \quad (4.33)$$

which we call, empirical 2ptCF. We fit this model in the mean of 1000 QPM mock catalogs 2ptCF as shown in Fig. 4.17, in order to determine the correction shape function, i.e. determine the $\vec{p}_{cs} = \vec{p}_{cs}^{fit}$, parameters for the best fit values of (b, σ_p) - parameters. The resulting function parameters defines us the $\Delta_i = \Delta(r_i)$ function as:

$$\Delta(r) = f_{cs}(r; \vec{p}_{cs} = \vec{p}_{cs}^{fit}) - 1 \quad (4.34)$$

In this plot one can see the empirically 2ptCF, $\xi_{fit}(r; b, \sigma_p, \vec{p}_{cs})$ fitting the data with a conservative $\chi^2 = 8.1$ over 22 degrees of freedom.

In Fig. 4.18, we present the χ^2 -test for different boosting covariance matrices parametrized by parameter e described in Eq. 4.31. From left to right we can see the same test for different fitting ranges with minimum scale $r_{minfit} = [1, 2, 3]h^{-1}\text{Mpc}$. One can see that that only the value for $e = 10$ can give

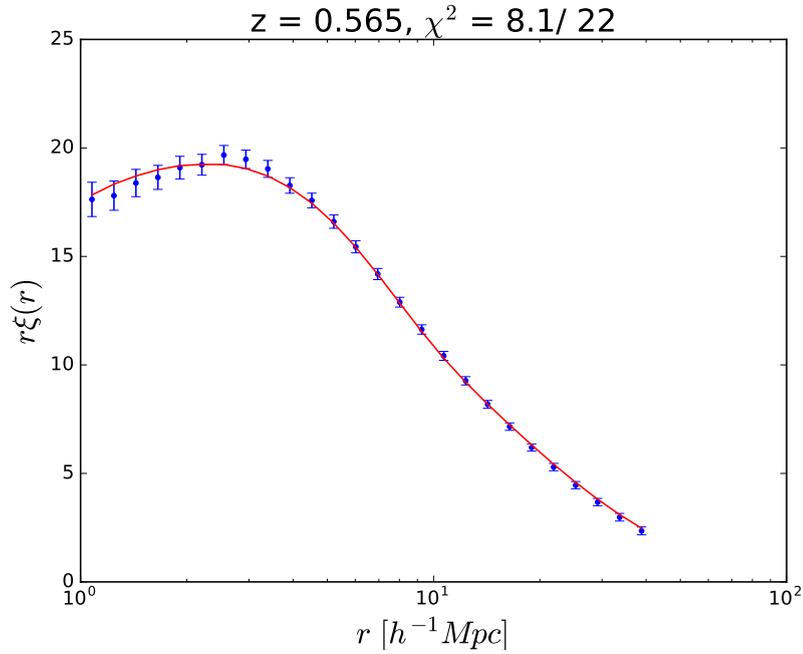


FIGURE 4.17: Empirical 2ptCF, $\xi_{fit}(r; b, \sigma_p, \vec{p}_{cs})$, as a function of scales, r , for the best fit parameters, $(b, \sigma_p, \vec{p}_{cs}) = (b^{fit}, \sigma_p^{fit}, \vec{p}_{cs}^{fit})$ for the mean of the $(0.538 \leq z \leq 0.592)$ -slice. [See text for details]

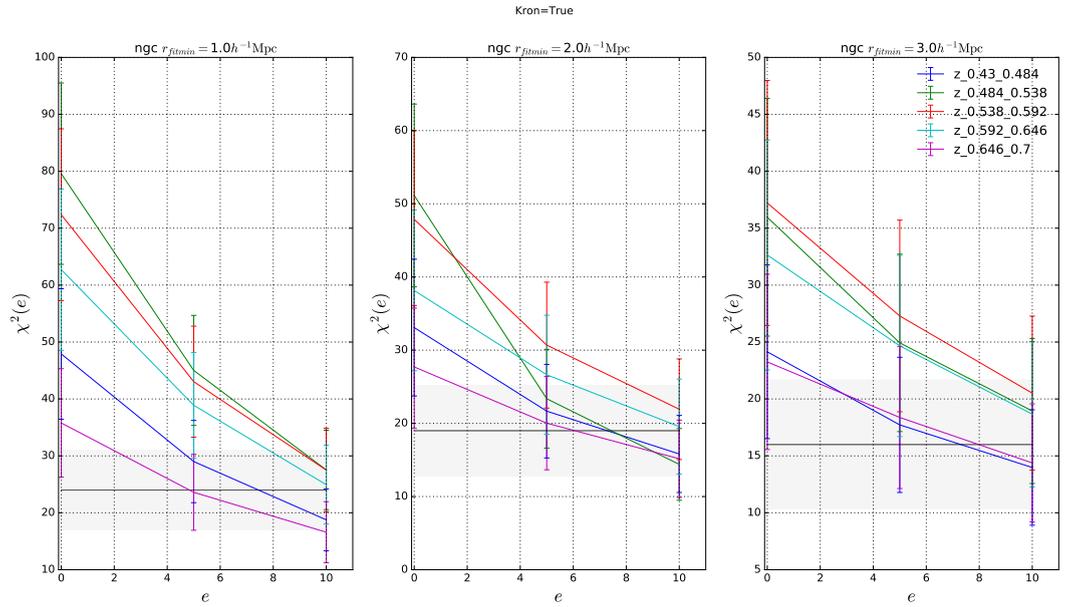


FIGURE 4.18: χ^2 test for selection of best boosting error parametrized by e , for different fitting ranges from left to right and different redshift slices (different colors). See text for details]

satisfactory χ^2 -test, i.e. the measured $|\chi^2(e) - ndf| \leq \sqrt{2ndf}$, where ndf are the number of degrees of freedom for 2 free parameters (b, σ_p) . The ndf parameter varies, 22, 19, 16 for the different values of the minimum scale bins. This is true for every minimum fitting scale, r_{minfit} and every redshift slice described in table 3.1. One can see the effect of different values of e parameter onto the estimated auxiliary parameters b, σ_p , in Fig. 4.19 and Fig. 4.20 respectively. One can see that there is no significant

bias on the measurement for each of those parameters for different values of e against the measurement with $e = 0$.

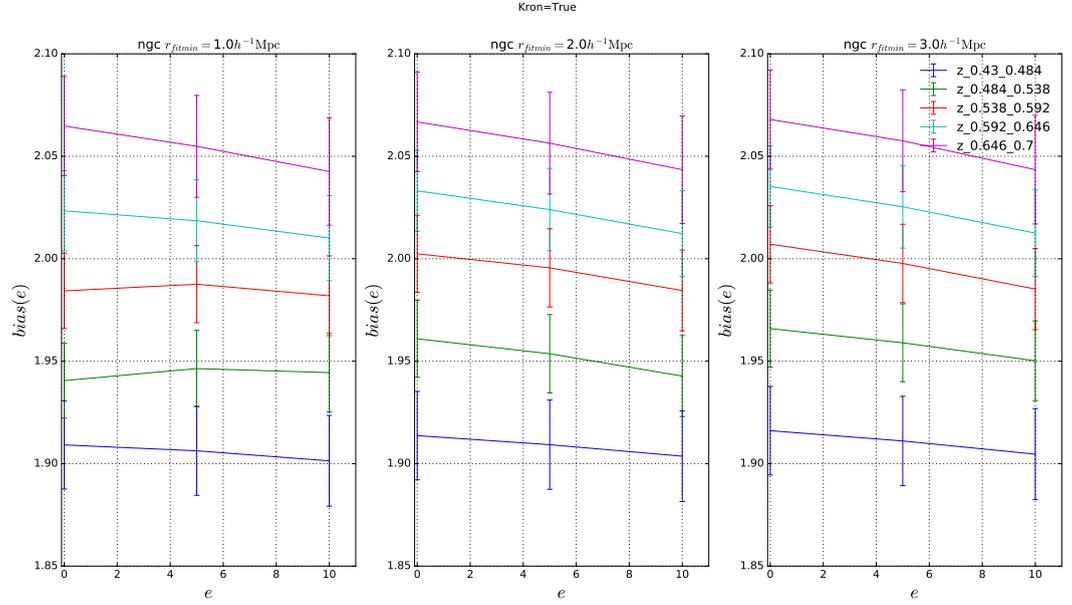


FIGURE 4.19: Cosmic bias parameter, b , as a function of boosting parameter e , for different fitting ranges from left to right and different redshift slices (different colors). See text for details]

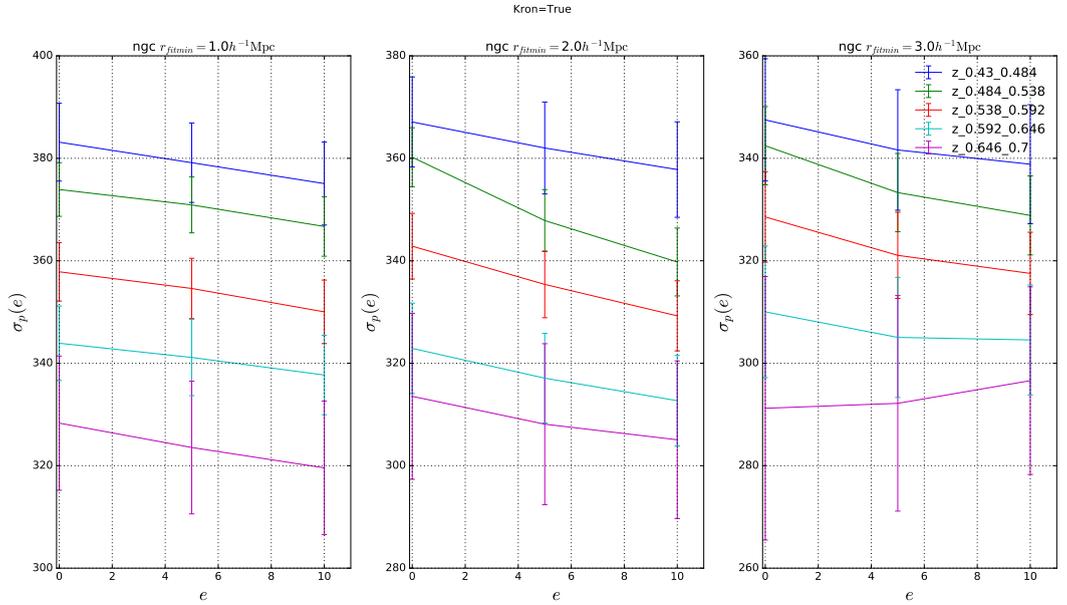


FIGURE 4.20: Same as Fig. 4.19 but for the peculiar velocities, σ_p . [See text for details]

Therefore, in Fig. 4.21 (left) we display the resulting boosting factor on the error in the correlation function for $e=10$. It appears to be significant only on scales smaller than $10\ h^{-1}\ Mpc$, i.e. up to 3 times the original error, and degrading at scales larger than $10h^{-1}Mpc$ at zero contribution to the error (i.e. down to 1% to the original one).

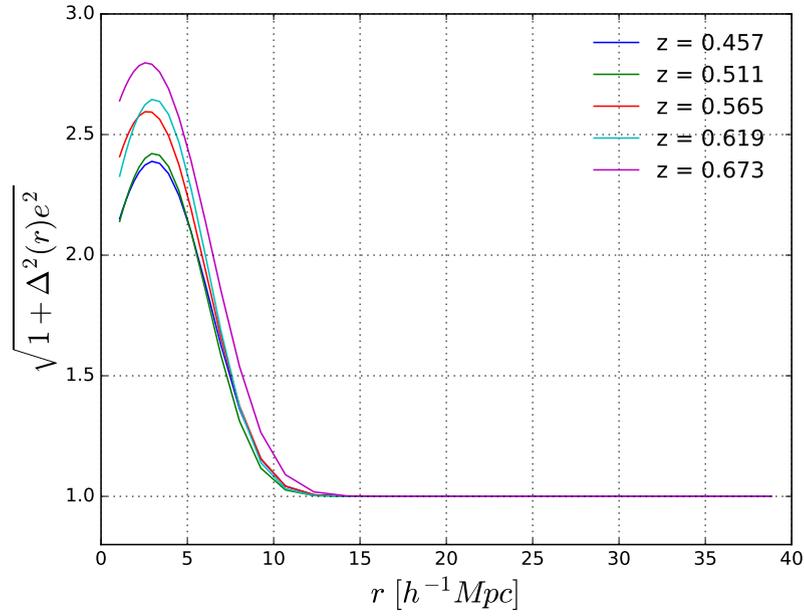


FIGURE 4.21: Boosting factor on the error in the correlation function as a function of scales, for all redshift bins (different colors).

Finally, from Fig. 4.22 we can understand that the reconstructed homogeneity scales measured with \mathcal{D}_2 is not significantly modified by the error boosting, i.e. $e=10$, and matches the measured one, i.e. $e=0$.

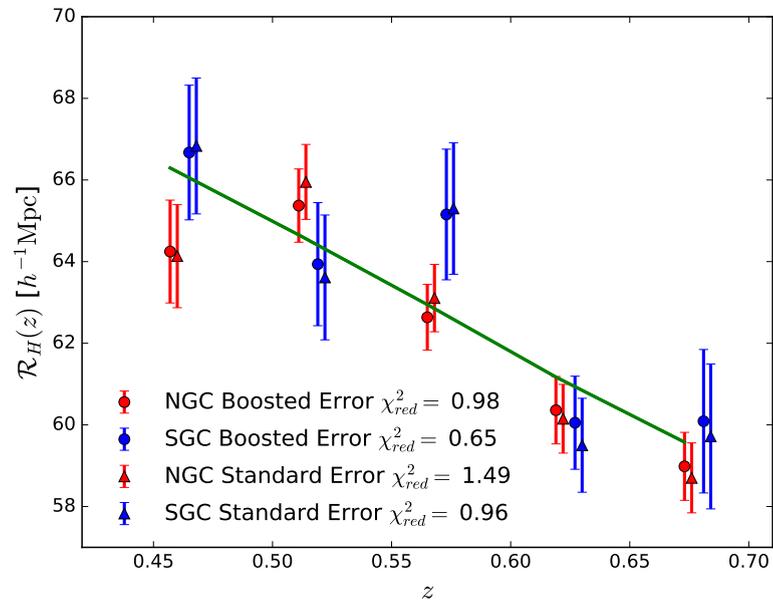


FIGURE 4.22: $\mathcal{R}_H(z)$ with (circle) and without (triangle) boosting the errors, for NGC (red) and SGC (blue).

4.3.8 Test of spline fit on QPM mock catalogs

We perform a spline fit of \mathcal{D}_2 for the 1000 QPM mock catalogs in order to obtain the homogeneity scale at 1% (Eq. 4.7). The fit is performed in the range $r \in [40, 100] h^{-1}$ Mpc with 6 data points and 1 degree of freedom. The distribution of the χ^2 of the mock should therefore follow a χ^2 -law for 1 degree of freedom. In table (4.7) we show the mean and the error on the mean of the distribution of the corresponding χ^2 for the 1000 QPM mock catalogs. The test is successful in both NGC and SGC.

z	χ^2_{NGC}	χ^2_{SGC}
0.430-0.484	1.00 ± 0.05	0.99 ± 0.04
0.484-0.538	0.99 ± 0.05	1.00 ± 0.04
0.538-0.592	1.02 ± 0.05	1.00 ± 0.04
0.592-0.646	0.99 ± 0.04	1.00 ± 0.05
0.646-0.700	1.02 ± 0.05	1.00 ± 0.04

TABLE 4.7: Mean and error over the 1000 QPM mock catalogs for the χ^2 of the spline fit with 1 degree of freedom, in the NGC and SGC for the five redshift bins.

4.3.9 Qualitative tests between estimators

We consider three estimators for $\mathcal{N}(< r)$, defined by equations (4.13), (4.14) and (4.15). We compute $\mathcal{N}(< r)$ and the resulting $\mathcal{D}_2(r)$ with the three estimators for the 1000 QPM mock catalogs. Figure 4.23 compares the mean of the 1000 mocks to the Λ CDM model. For the calculation of the estimators we have used the reconstruction method described in chapter 4.2.2 the bias and peculiar velocities that were estimated in section chapter 4.2.2. In Fig. 4.23 we show the computation for the different estimators with different colours. The result with the cor estimator are more accurate than *lau* or *str* estimators in respect of the Λ CDM model.

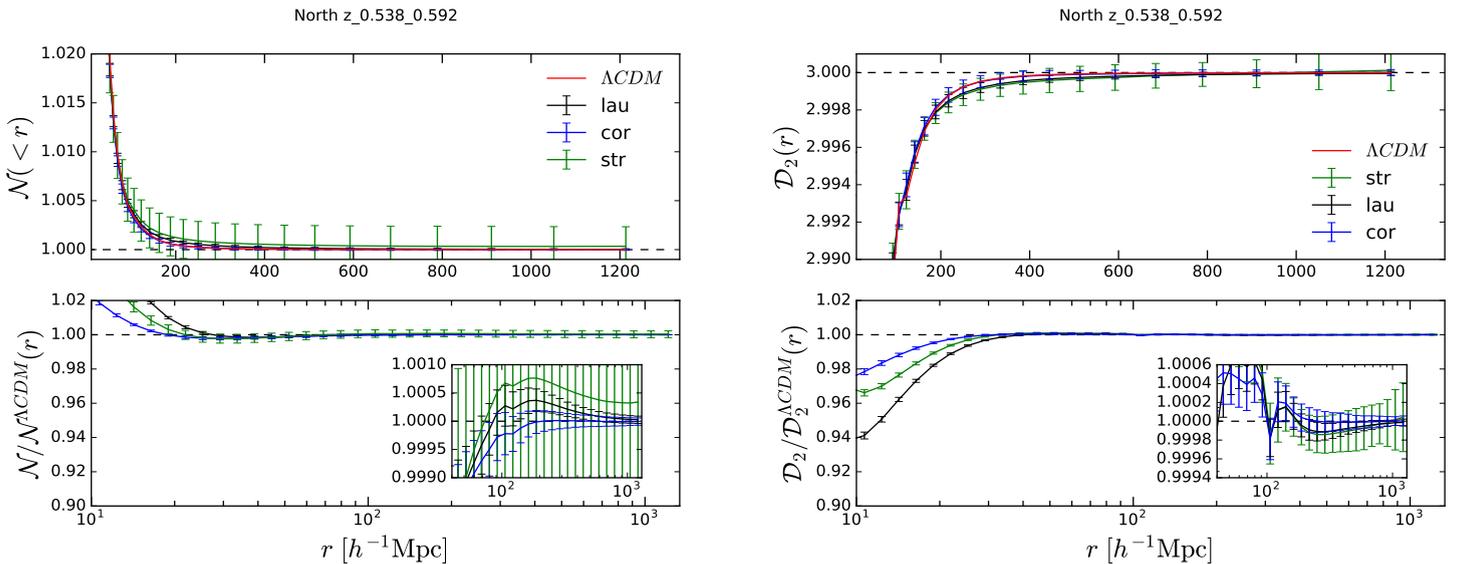


FIGURE 4.23: Top: the scaled counts-in-spheres, $\mathcal{N}(< r)$, (left) and the fractal correlation dimension, $\mathcal{D}_2(r)$, (right) for matter distribution, with str(green), lau (black) and cor (blue) estimators, compared to Λ CDM model (red). Bottom: the ratio to Λ CDM model for both estimator.

Furthermore, we compare the resulting 1σ dispersion as a function of scales, r , for the different estimators in Fig. 4.24. We can see that for the observable \mathcal{N} the str-estimator reaches 1 order of magnitude larger dispersion than the cor-estimator at scales larger than $100h^{-1}\text{Mpc}$. But the lau-estimator is about 5% more precise than the cor-estimator at scales larger than $30h^{-1}\text{Mpc}$. For the observable \mathcal{D}_2 , we detect for str-estimator a 1σ dispersion up to 7 times larger than the cor-estimator. While, the ratio between lau and cor estimators is a monotonically decreasing function at scales $[20 - 1300]h^{-1}\text{Mpc}$ reaching up to 40%. This renders the lau-estimator more precise than the cor-estimator. However, the scale of interest of measuring the homogeneity scales are about $60h^{-1}\text{Mpc}$ where the difference in precision is only 5%.

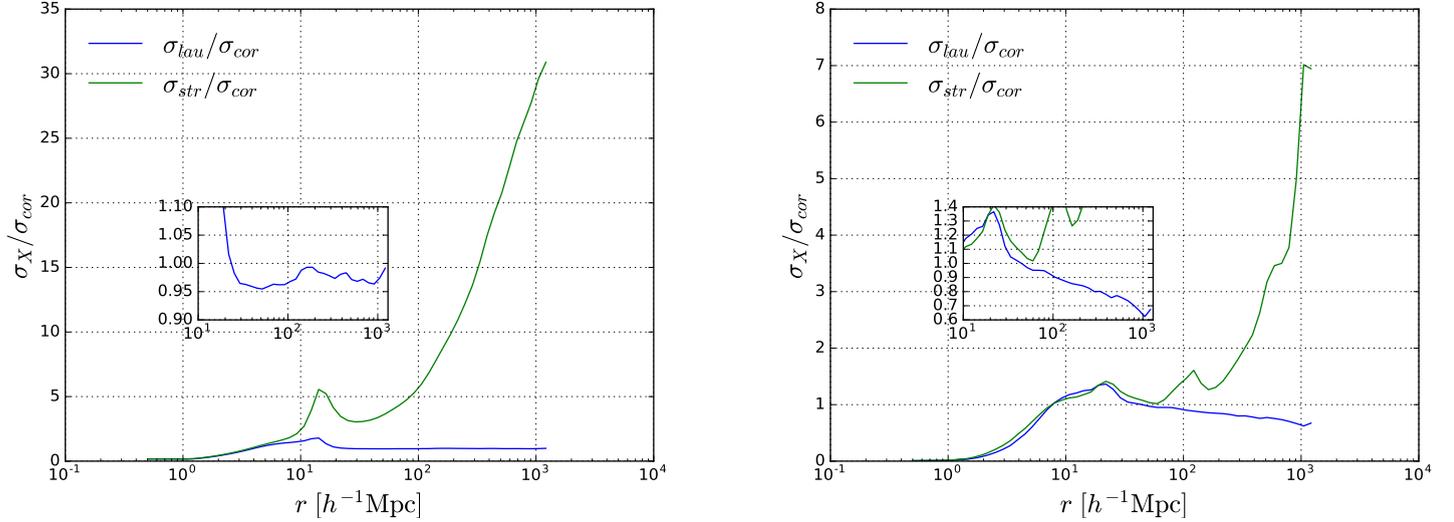


FIGURE 4.24: Ratio of the 1σ of the observables \mathcal{N} (left) and \mathcal{D}_2 (right) for estimators *lau* (blue) and *str* (green) against *cor* as defined in table 4.2 as a function of scales, r . We zoom in at scales $r \in [10 - 1300]h^{-1}\text{Mpc}$.

For these reasons, we decided to use *cor*-estimators since it is more accurate but only slightly less precise than the *lau*-estimators for both observables \mathcal{N} and \mathcal{D}_2 at the scale of interests.

4.3.10 RSD- \mathcal{R}_H correlations

As discussed in chapter 4.2.3, to perform the measurement of the homogeneity scale \mathcal{R}_H on the matter distribution we perform before a measurement on the bias, b and peculiar velocities, σ_p , to account for Redshift Space Distortions. Therefore, we are interested in the correlation, r , between the homogeneity scale and the bias and peculiar velocities. To calculate the correlation coefficient we use Eq. 4.26, the 1000 QPM mock catalogs for the different redshift slices described in table 3.1 on the North or South galactic caps. We measure the bias and peculiar velocities for each individual QPM mock catalog according to chapter 4.2.2. The measurement of the homogeneity scale for the galaxy or matter distribution is in accordance with chapter 4.2.1 or Fig. 4.2.3, respectively.

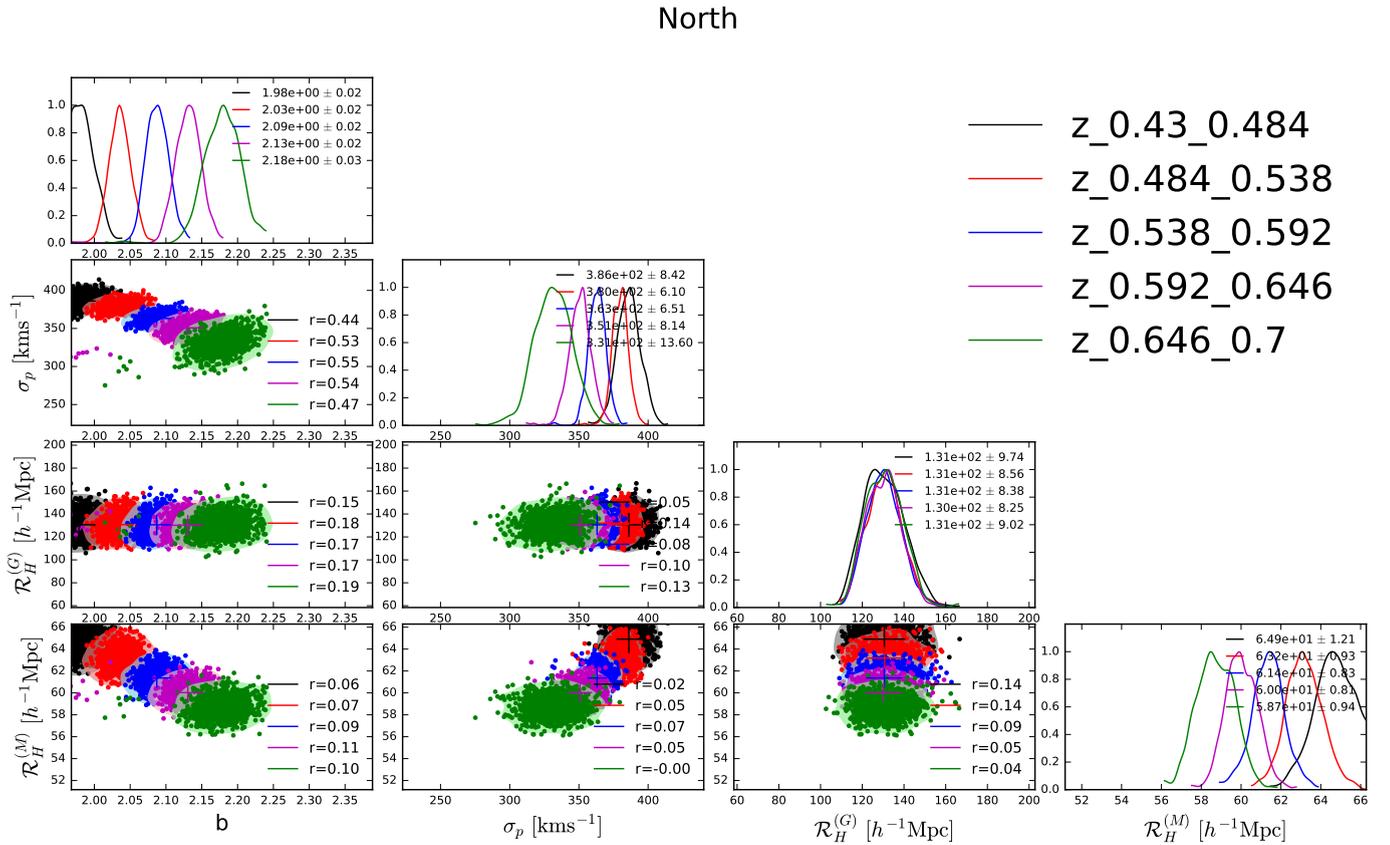


FIGURE 4.25: Homogeneity scale for the galaxy distribution measurement, $\mathcal{R}_H^{(G)}$ or matter distribution, $\mathcal{R}_H^{(M)}$, versus bias, b , or peculiar velocities, σ_p , measurements for 1000 QPM mock catalogs on the North Galactic Cap for the different redshift bins color-coded. The contours show the 1σ and 2σ confidence regions. The normalized histograms with the mean and 1σ values are displayed for the two measurements. The correlation coefficient, r , is also displayed.

The measurement between the homogeneity scale on the galaxy distribution, $\mathcal{R}_H^{(G)}$, or the matter distribution $\mathcal{R}_H^{(M)}$, bias, b , and peculiar velocities, σ_p is illustrated in Fig. 4.25 for the different redshift which are color-coded for the North Galactic Cap. Furthermore, the correlation coefficient, r , is computed for each redshift slice. On the diagonal panels are the normalised histograms of the bias, peculiar velocities homogeneity scale for the galaxy and matter distributions, from top to bottom respectively. For each redshift bin the the mean and 1σ values of the measurements are also displayed. We can observe that the $\mathcal{R}_H^{(G)}$ does not vary with the bias or the redshift. However, the bias increases with redshift with a precise measurement of about $\sim 1\%$. We can see that the $\mathcal{R}_H^{(G)}$ does not vary with the σ_p as well. On

the other hand, the peculiar velocities are slightly decreasing with redshift. Similar results we obtain for the homogeneity scale on the South Galactic Cap as we illustrate in Fig. 4.26.

South

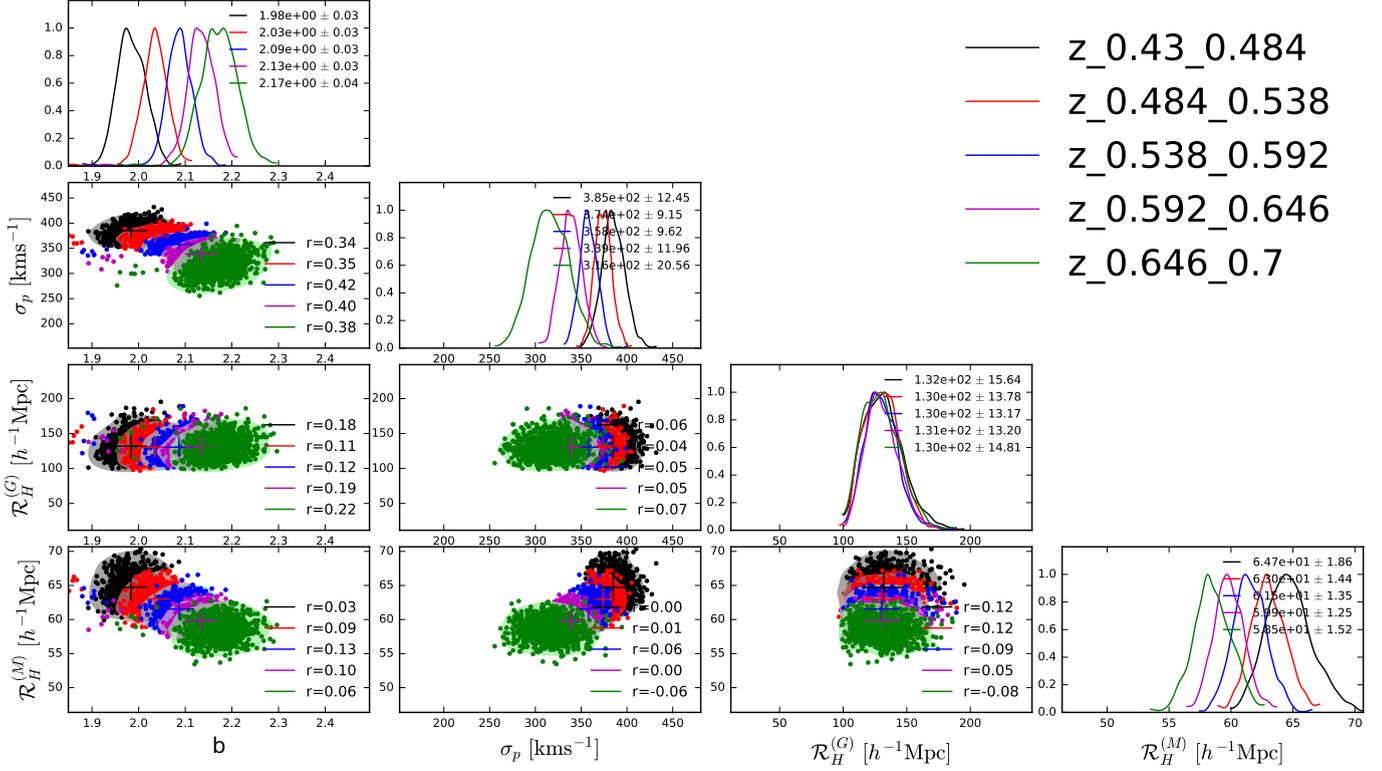


FIGURE 4.26: Same as Fig. 4.25 but for the South Galactic Cap.

Finally, in left panel of Fig. 4.27, we present the correlation coefficient, r , as a function of redshift, z , for the different combination between the measurements on the homogeneity scale and bias for North or South galactic cap and galaxy and matter distribution. We observe that the correlations are always positive, reaching up to 2.5% between the homogeneity on the galaxy distribution and the different variables on the South galactic cap. Additionally, in the right panel of Fig. 4.27, we present the correlation coefficient between the homogeneity scale and the peculiar velocities measurement for the same combinations. We observe less correlation between the different measurement reaching up to 1.1% rendering the correlation negligible. We conclude that the correlation between \mathcal{R}_H and b or σ_p is small and we do not account for them into our analysis.

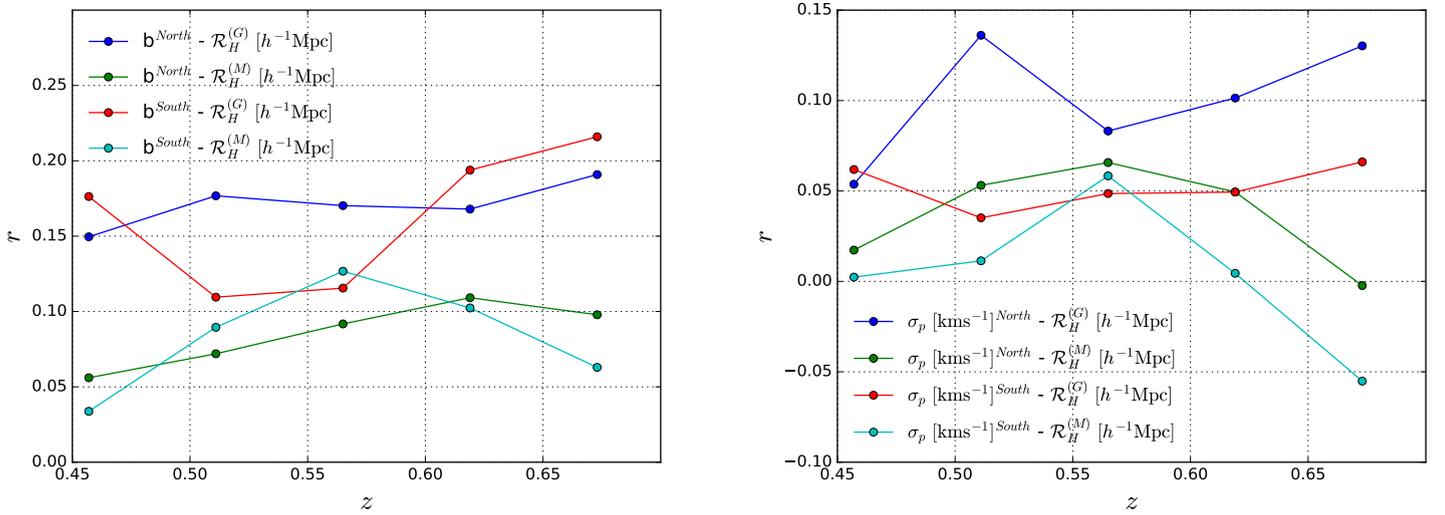


FIGURE 4.27: Correlation coefficient, r , as a function of redshift, z , between bias, b (left), or peculiar velocities, σ_p (right), and the homogeneity scale, \mathcal{R}_H for the North or South galactic caps for the matter (M) or galaxy (G) distribution, as measured for 1000 QPM mock catalogs.

4.3.11 Weight Scheme Systematic tests

We perform the measurement of the bias, peculiar velocities, homogeneity scale of the galaxy and matter distribution for different weighing schemes as described in chapter 3.4.2. Each galaxy has a total weight that is a combination of different systematic effect on the observed local galaxy density. First we perform the measurements by switching off each individual weight and keeping the rest of the weights on. Then we switch off all the weights at once. The results of the bias and peculiar velocities are shown Fig. 4.28, on the top left and top right panel respectively. In colors we represent the measurement of the homogeneity scale for the different weights that we switch off. We label with "No" when we switch off all the weights. With "NoStar, NoSee, NoFKP" when we switch of the weights w_{star} , w_{see} , w_{fkp} , respectively. Notice that we switch off both the redshift failure weight w_{zf} and the weight of closed pairs, w_{cp} , namely "NoZCP". Finally, with "All" we represent the null measurement, i.e. the basis measurement of our analysis. On the bottom panels we show the relative measurement in respect of the null measurement. We observe that the measurement of bias for switching of the w_{star} , w_{see} , w_{fkp} is not affected for all redshift bins. For removing the weight w_{nz} and w_{cp} or all, we see that we have a systematic error less 2% only for the intermediate redshift bins $0.484 \leq z \leq 0.538$ and $0.538 \leq z \leq 0.592$. However, this systematic is within the statistical errors. For the case of the peculiar velocities we observe a systematic error of less than 3%, for the same removal of weights w_{nz} and w_{cp} , within the statistical errors. Therefore the measurement of bias and peculiar velocities are driven by statistics.

We present the results for the systematic error of the homogeneity scale measurement for the galaxy and matter distribution, in left and right panel of Fig. 4.29, respectively. We observe no significant systematic error for the redshift bins larger than $z \geq 0.5$ for the homogeneity scale of the galaxy distribution, i.e. the systematic errors is less than 2%, consistent within the 5% precision of the statistical measurement. Furthermore, we observe a 5% systematic error when we turn off the weights w_{fkp} or all the weights but within the statistical errors of 10%. The homogeneity scale of the matter distribution follows the same picture. The maximum systematic error is only at 1% significance when we turn off all the combination of the weights for the redshift slice $0.484 \leq z \leq 0.538$. However this is within the 1σ statistical error of 2% as it is true for all other redshift slices.

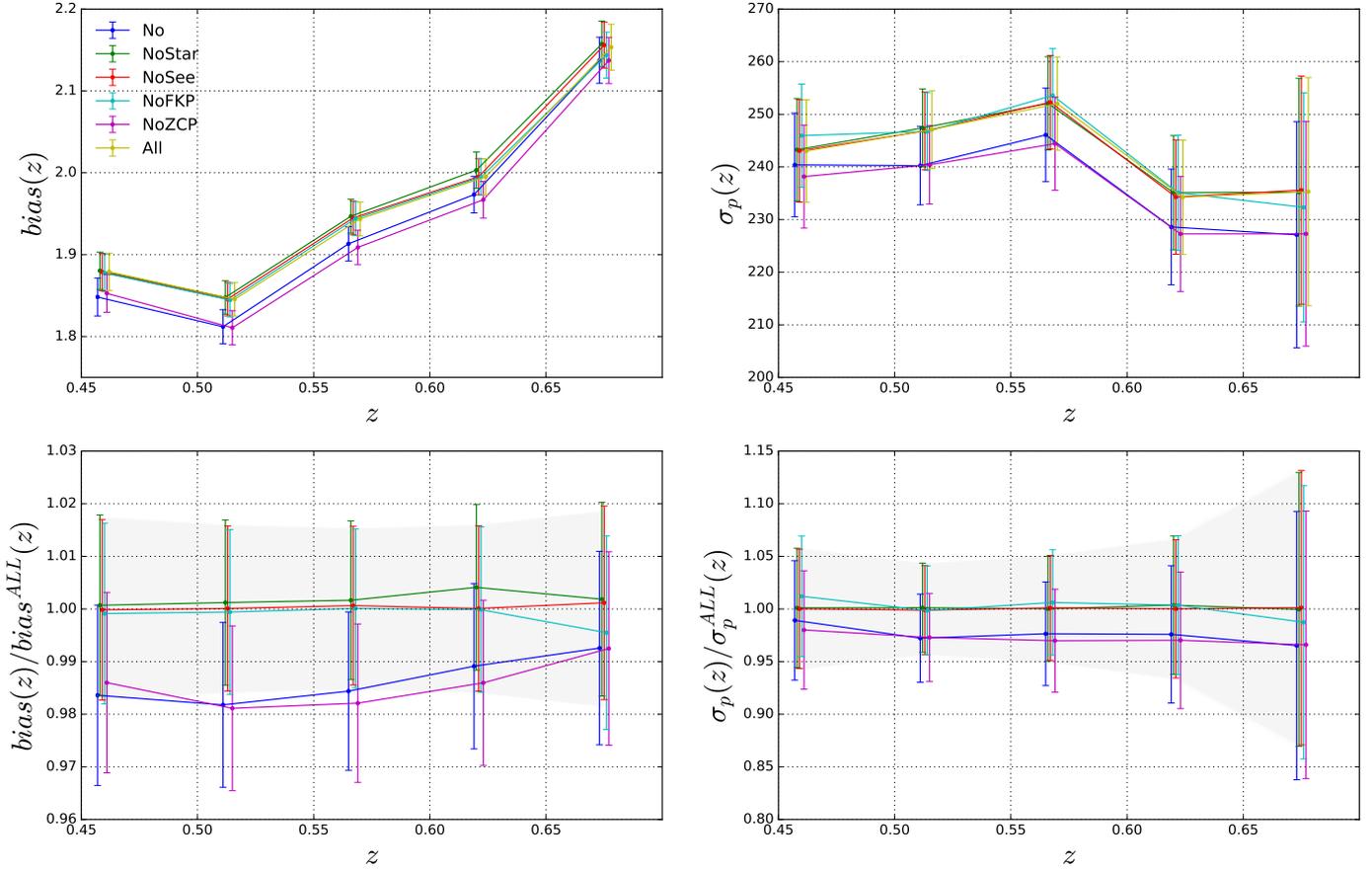


FIGURE 4.28: Systematic measurement of the bias, b and peculiar velocities, σ_p as a function of redshift bins, z for the North galactic cap. Different colors correspond to different weighting scheme. Top figures show the measurement and bottom figures show the percentage ratio from the measurement applying all the weights. Shaded regions shows the 1σ error of the measurement applying all the weights.

Similar results we obtain for the South galactic cap for the bias and peculiar velocities displayed in Fig. 4.30. We see that we observe no significant systematic for the homogeneity scale measurement on the south galactic cap represented in Fig. 4.31.

4.4 Bias Gain

As we show in appendix A.8, assuming a scale independent bias, the precision on the measurement of the correlation function on the matter distribution is not affected by the bias:

$$\frac{\sigma_{\xi^G}}{\xi^G}(r) \simeq \frac{\sigma_{\xi^M}}{\xi^M}(r) \quad (4.35)$$

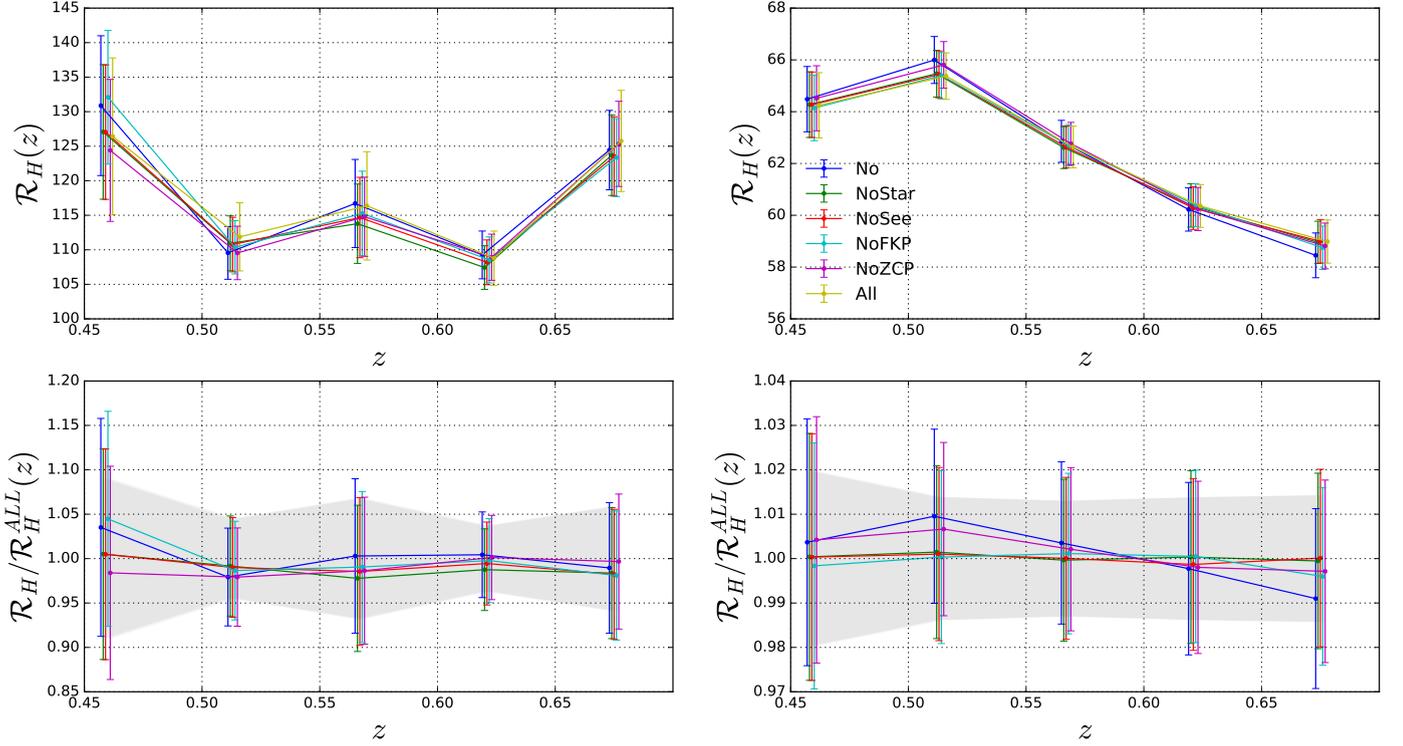


FIGURE 4.29: Systematic measurement for the homogeneity scale, \mathcal{R}_H , of the galaxy (left) and matter (right) distributions as a function of redshift, z , for the North galactic cap. Different colors correspond to different combination of the weighting scheme. Top figures show the measurement and bottom figures show the percentage ratio from the measurement applying all the weights. Shaded regions shows the 1σ error of the measurement applying all the weights.

While we show that for the scaled counts-in-spheres at large scale limit, i.e. $\mathcal{N}^M(< r) - 1 \simeq 0$, we have that:

$$\frac{\sigma_{\mathcal{N}^M}(r)}{\mathcal{N}^M(r)} \simeq \frac{\sigma_{\mathcal{N}^G}(r)}{\mathcal{N}^G(r) + (b^2 - 1)} \quad (4.36)$$

Notice that there is a gain for a biased tracer, while for an unbiased tracer there is no gain ($b^2 = 1$). Similar we have for the Fractal Dimension Precision at the large scale limit, i.e. $\mathcal{D}_2^G(r) - 3 \simeq 0$, that :

$$\frac{\sigma_{\mathcal{D}_2^M}(r)}{\mathcal{D}_2^M(r)} \simeq \frac{\sigma_{\mathcal{D}_2^G}(r)}{\mathcal{D}_2^G(r) + 3(b^2 - 1)} \quad (4.37)$$

In order to verify this calculation, we perform a test by using the 1000 QPM mock catalogs. We measure the fractal correlation dimension on the galaxy distribution and the matter distribution with the reconstruction method described in chapter 4.2.3. Then, we use the measurement of the galaxy distribution but we correct with the bias using Eq. 4.37. We perform the same measurement for normalized-counts-in-spheres for the galaxy and matter distribution as well as a measurement for the galaxy distribution but correcting for bias using equation Eq. 4.36.

The results from the fractal correlation dimension are shown in the top panel of Fig. 4.32. We see that the precision for the fractal correlation dimension of the matter distribution with the reconstructed method, \mathcal{D}_2^{nat} , shown with red line matches the one with the fractal correlation dimension of the galaxy distribution corrected with bias, i.e. using Eq. 4.37, which is shown with green line. Similar results

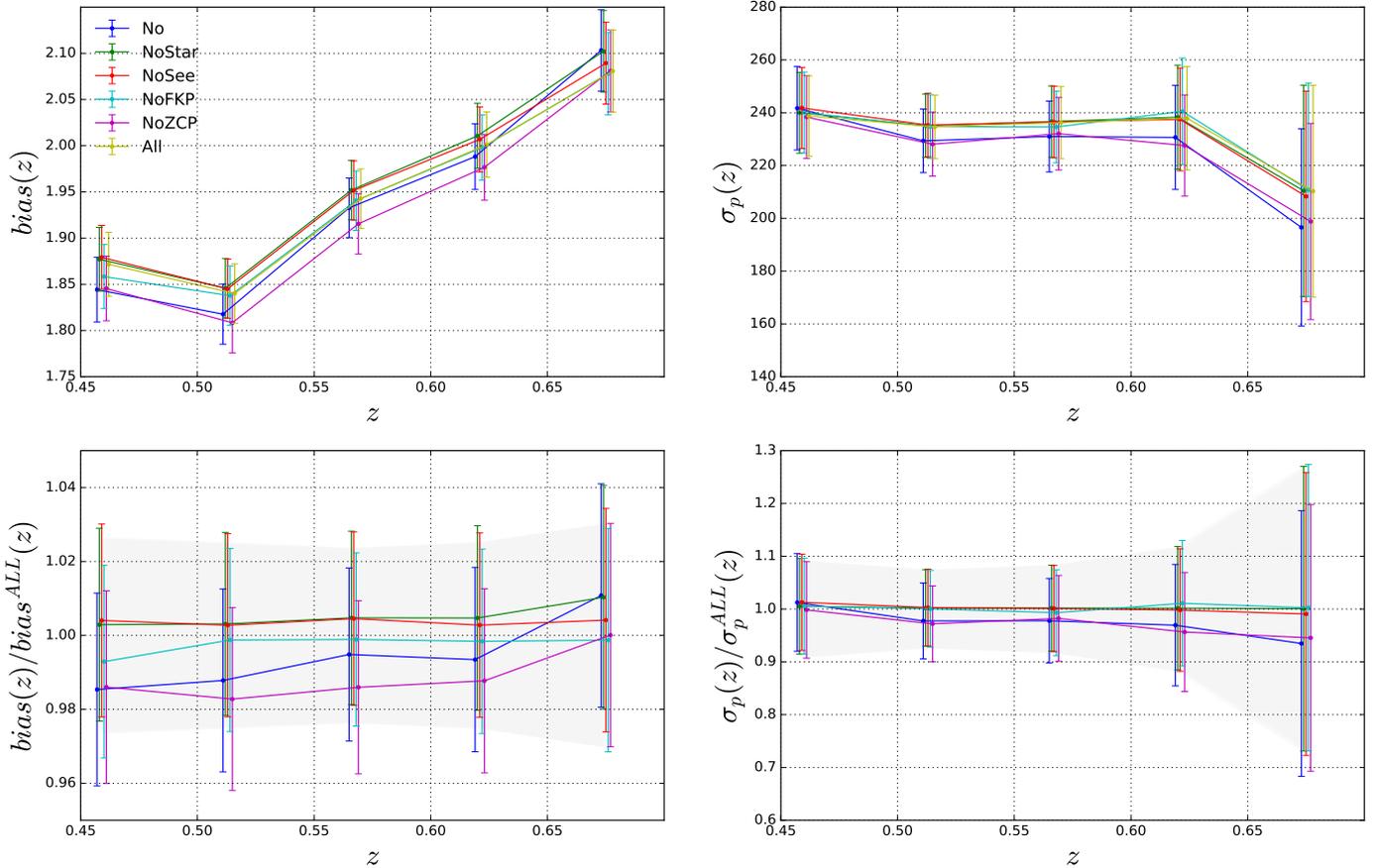


FIGURE 4.30: Same figure as Fig. 4.28 for the south galactic cap

we obtain for the scaled-counts-in-spheres, \mathcal{N} , which are shown in the bottom panel of Fig. 4.32. We see that the scaled counts-in-spheres Eq. 4.36 (bottom). The Eq. 4.37 and Eq. 4.36 tell us that the precision of measuring the observable for the matter distribution from a biased tracer gains in precision from the individual bias measurement.

4.5 Homogeneity scale at 0.1%

The choice of a 1% threshold to define the homogeneity scales is arbitrary. Therefore, we can define them for instance at 0.1% as:

$$\mathcal{D}_2(R_H^{\mathcal{D}_2=2.997}) = 2.997 \quad \text{or} \quad \mathcal{N}(R_H^{\mathcal{N}=1.001}) = 1.001, \quad (4.38)$$

namely per-Mille definition. Figure 4.33 shows that the measured homogeneity scale for matter distribution is compatible with Λ CDM, with $\chi_{red}^2 = 5.82/6$ in the NGC and $\chi_{red}^2 = 7.98/6$ in the SGC.

We observe that the definition of the homogeneity scale with a more strict threshold, 0.1%, result in a Homogeneity scale about three times higher than the one with 1%, i.e. $\mathcal{R}_H^{0.1\%,NGC} = 157.2 \pm 4.5 h^{-1} \text{Mpc}$

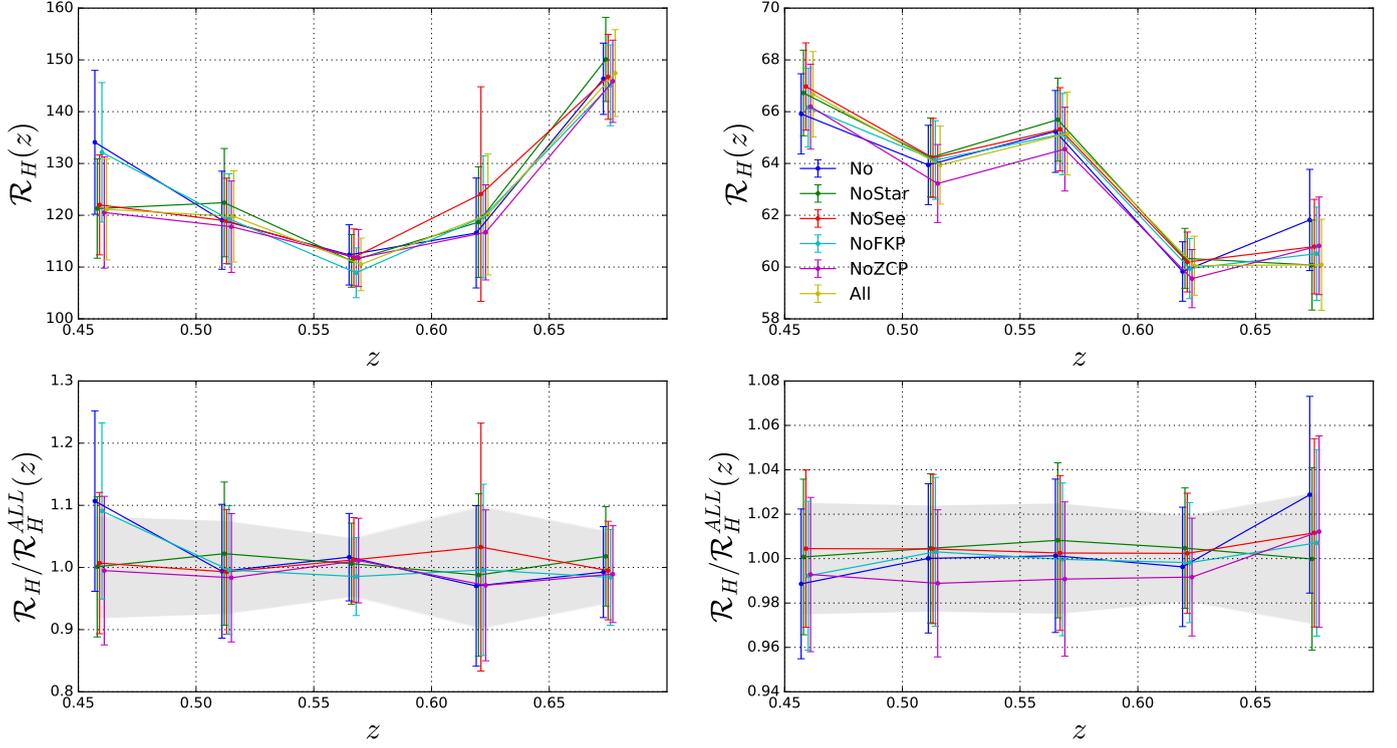


FIGURE 4.31: Same figure as Fig. 4.29 for the south galactic cap

and $\mathcal{R}_H^{0.1\%,SGC} = 174.9 \pm 4.7 h^{-1} \text{Mpc}$. This per-Mille definition has less precision than the one we use for our basis analysis. This is true since the former definition probes larger scales where the volume density drops at 10% as shown in Fig. 4.9, i.e. probing less statistics. Finally, using the definition of Eq. 4.38, it is possible to compare the results of both of the analyses that I was involved in: Galaxy analysis and QSO analysis (to a smaller extent).

4.6 Alternative Definition of \mathcal{R}_H results

In this section we explore the error definition of the homogeneity scale. In particular, we use the minimum scale that corresponds to the fractal dimension value plus the one sigma error that is less or equal to the nominal value for homogeneity, 3 as we explain in chapter 4.2.3. This is formulated as:

$$\mathcal{R}_H^E := \min \{ r \in \mathbb{R} : (\mathcal{D}_2 + \sigma_{\mathcal{D}_2})(r) > 3 \} . \quad (4.39)$$

In Fig. 4.34, the weighted mean between the North galactic cap and South galactic cap for fractal correlation dimension, $\mathcal{D}_2(r)$, at different scales, r , and redshift bins is displayed. We determine the homogeneity scale as defined in Eq. 4.39 for the matter distribution (left) and the galaxy distribution (right). Different colours correspond to the \mathcal{R}_H^E for the different redshift bins. The rest of the fractal correlation dimension is color with grey color for all redshift bins. The error homogeneity scale is reached at different scales for different redshift bins.

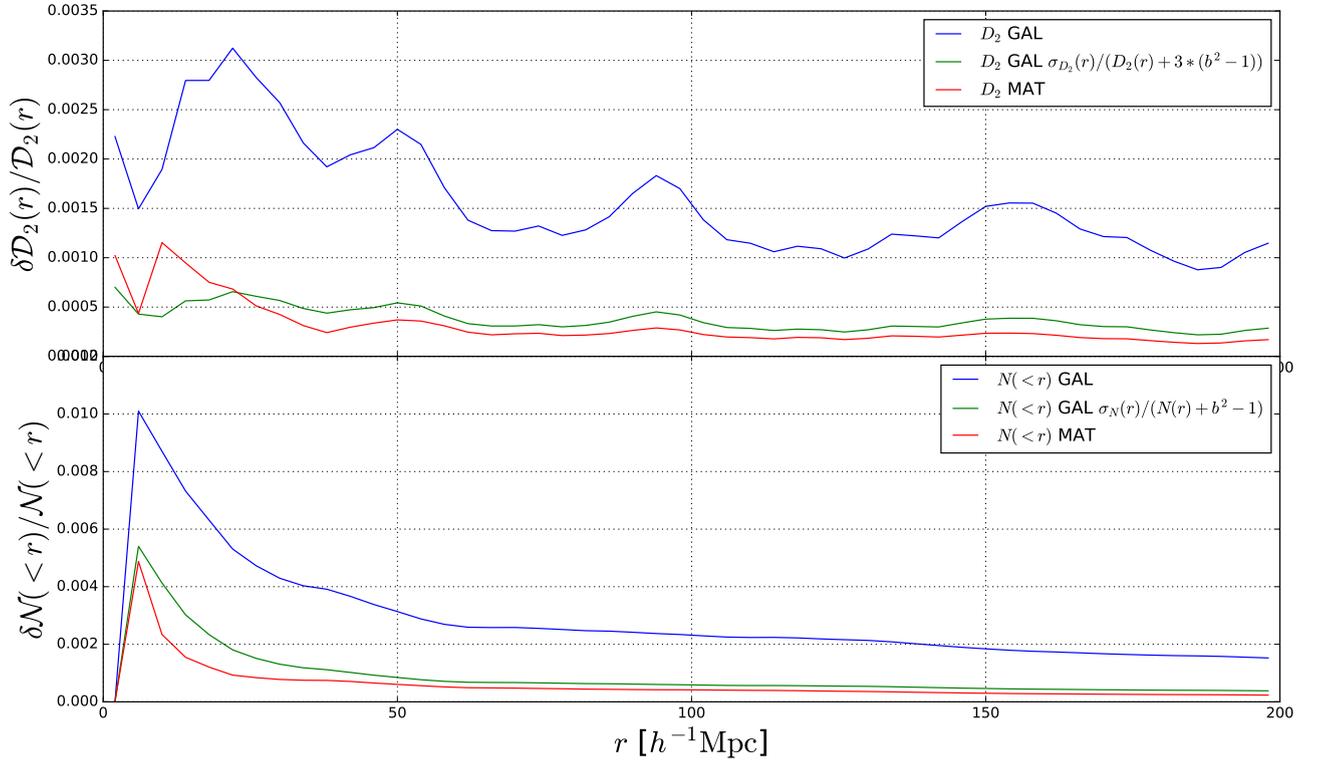


FIGURE 4.32: *Top*: D_2 -precision *Bottom*: N -precision for galaxy (blue) , matter (red) and the prediction-relation (green)

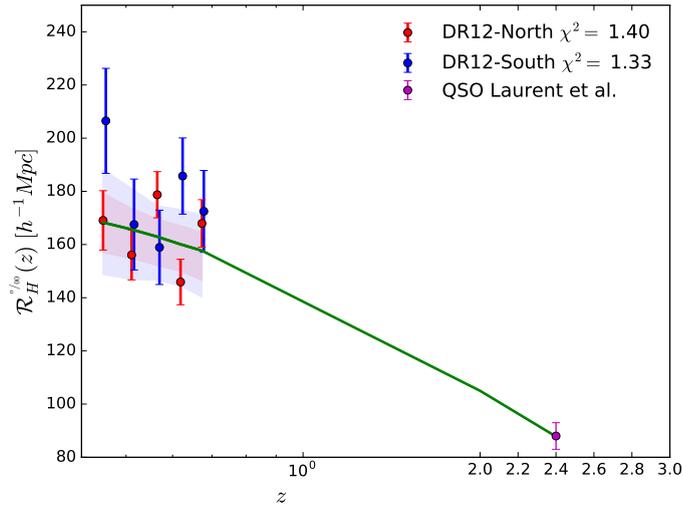


FIGURE 4.33: The homogeneity scale at 0.1% level, $R_H^{D_2=2.997}(z)$, measured in the NGC (red) and in the SGC (blue) as a function of redshift. The purple point is the result obtained with quasars in the NGC, in the redshift range $2.2 \leq z \leq 2.8$ by Laurent et al. [166]. The green line is the Λ CDM model prediction. The shaded areas indicate the 1σ range for the 1000 QPM mock catalogs.

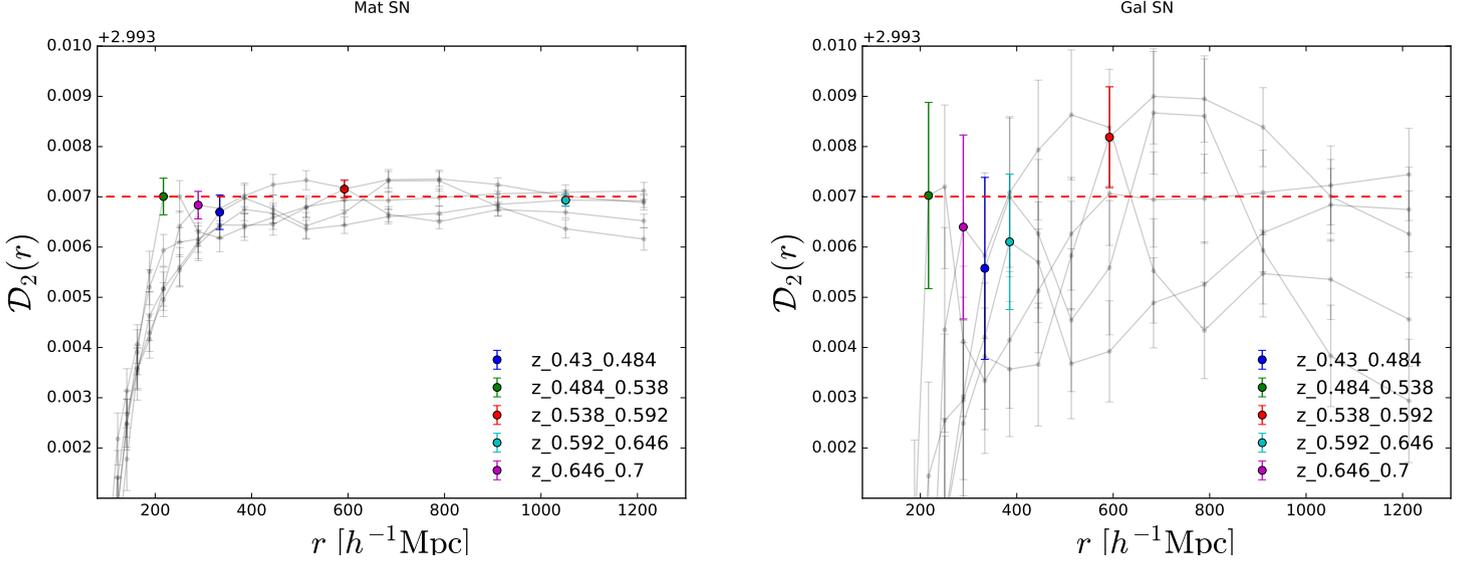


FIGURE 4.34: \mathcal{R}_H^V -estimator defined by equation 4.39 for 5 z -bins. *Left:* Galaxy *Right:* Matter.

In addition, we compute this estimator (\mathcal{R}_H^E) for the 1000 QPM mock catalogs. The results are summarised in Fig. 4.35 for the North galactic cap (left) or South galactic cap(right) for the matter or galaxy distribution. In Fig. 4.35, we can observe that the mean and 1σ dispersion of the error homogeneity scale for 1000 QPM mock catalogs is $\mathcal{R}_{H,mock}^{E,NGC} \simeq 360 \pm 120 h^{-1} \text{Mpc}$ for the north galactic cap and $\mathcal{R}_{H,mock}^{E,SGC} \simeq 300 \pm 120 h^{-1} \text{Mpc}$ for the south galactic cap. Therefore, the test with the mock catalogs suggest a signal to noise ratio for all redshift bins about $S/N \simeq 3$. Furthermore, we observe that $\mathcal{R}_{H,mock}^{E,NGC} > \mathcal{R}_{H,mock}^{E,SGC}$. This is consistent with the fact that the volume of the survey is larger on the North Galactic Cap than the South Galactic Cap, $V_{surv}^{NGC} < V_{surv}^{SGC}$. This is true, since we shown in chapter 5.5 that the sensitivity of the fractal correlation dimension and therefore the error of the Error homogeneity scale is proportional to the inverse of the volume of the survey, $\sigma_{\mathcal{D}_2} \propto \sigma_\xi \propto \sigma_P \propto \frac{1}{\sqrt{V}}$. Then, since $\sigma_{\mathcal{D}_2}^{SGC} > \sigma_{\mathcal{D}_2}^{NGC}$, then $(\mathcal{D}_2 + \sigma_{\mathcal{D}_2})^{SGC}$ approaches $\mathcal{D}_2 = 3$ in scales less than $(\mathcal{D}_2 + \sigma_{\mathcal{D}_2})^{NGC}$.

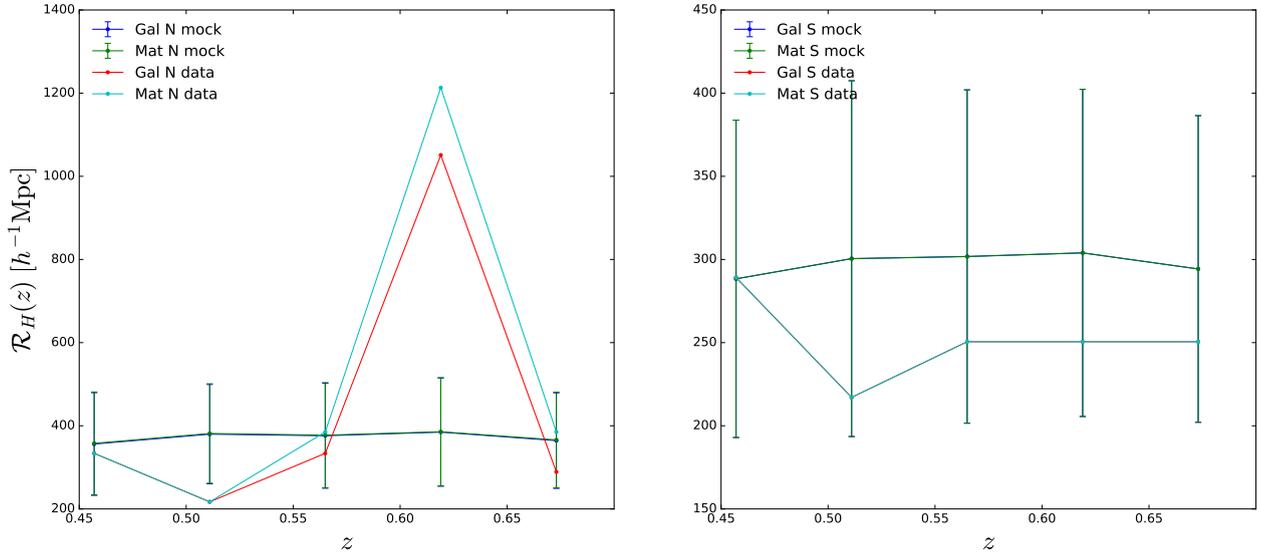
Furthermore, we observe that the data are within 1σ region of the mock catalogs, for almost every redshift bin. However, at $z \sim 0.62$ we have an outlier for both North and South galactic cap, $\mathcal{R}_{H,data}^{E,XGC} \simeq 1100 h^{-1} \text{Mpc}$, for which we have not found an explanation other than a statistical fluctuation.

From Fig. 4.35, we infer that both galaxy (Gal) and matter (Mat) distribution result in the same mean and variance of Error Homogeneity scale. Thus, we marginalise galaxy and matter information. Then we perform a Kolmogorov-Smirnov test[81] (KS-test) to see how likely is that the data distribution is drawn from the mock distribution. The result is summarised in table 4.8, where D^{KS} is the KS-statistic and D_{crit}^{KS} is the critical value for $\alpha = 0.001$ confidence level (C.L.).

C.L. = $\alpha = 0.001$	NGC	SGC
D^{KS}	0.196	0.334
D_{crit}^{KS}	0.617	0.617
$p - value$	0.791	0.172

TABLE 4.8: KS-test for North and South Galactic Cap. [See text for details]

The KS-statistic does not pass its critical value. Hence, the KS-test for this estimator of the homogeneity scale, \mathcal{R}_H^E , suggest that data and mocks are drawn from the same distribution at 21% and 83% C.L. for

FIGURE 4.35: Mean and dispersion of \mathcal{R}_H^E -estimator for NGC and SGC.

North and South galactic caps respectively. Finally, we conclude from this test that this estimator for the homogeneity scale depends on the the survey characteristics and it has a large dispersion, rendering this estimator not useful for precision measurements in cosmology.

4.7 3D Homogeneity or 2D Spatial Isotropy?

In large scale structure studies, we observe the redshift position of galaxies, z . Thus, the observed density of galaxies can be expressed by:

$$\rho(z, \theta, \phi) = \frac{d^3 N(z, \theta, \phi)}{dz d\Omega} \quad (4.40)$$

which is only a function of redshift and angles. Then, we can, easily, relate the observed density with the reconstructed density via any jacobian:

$$\frac{d^3 N(x, y, z)}{dx dy dz} = \frac{d^3 N(z, \theta, \phi)}{dz d\Omega} J(z, \theta, \phi). \quad (4.41)$$

However, the $J(z, \theta, \phi)$ depends on the relation between distances and redshift. This relation is model dependent. This makes the reconstruction of the density field a quantity that relies on the choice of the fiducial cosmology. One could imagine differently the possibility of having a different choice of a model. Different kind of models are described by a different metrics. We present the current possible models and their corresponding metric in table 4.9.

Universe	Metric Name
Homogeneous & Isotropic	FLRW
Non Homogeneous but Isotropic	LTB[182–184]
Fractal	RS[112]

TABLE 4.9: Different Universe and their metrics

In standard cosmography, we measure distances with FLRW metrics. This means that our Jacobian is defined via:

$$J_{FLRW}(z; p_{fid}) = \frac{H(z; p_{fid})}{cd_A^2(z; p_{fid})} \quad (4.42)$$

which is only a function of z for any possible fiducial cosmology $p_{fid} = (h, \vec{\Omega})$, where $\vec{\Omega}$ are the vector containing the energie densities of the different species of the universe described in chapter 1.2.2.

As we will saw in this thesis, the analysis of CMASS galaxy sample [144] and Quasar sample [166] shows that we observed a homogeneous 3D spatial distribution with a spectroscopic survey on scales larger than $r > 300h^{-1}\text{Mpc}$. This statement is quantified by the following relation:

$$\rho(x, y, z) \equiv \frac{d^3N(x, y, z)}{dxdydz} = \text{constant} . \quad (4.43)$$

Since we constrained by the choice of an FLRW cosmography, we are lead to the conclusion that

$$\frac{d^3N(z, \theta, \phi)}{dzd\Omega} \propto J_{FLRW}^{-1}(z; p_{fid}) . \quad (4.44)$$

Therefore, the observed density is not a function of angles. This is the definition of spatial isotropy. Thus, we conclude that a 3D spatial homogeneity analysis in the FLRW context is only an 2D spatial isotropy analysis. Thus, the CMASS galaxy sample and the QSO sample are spatially isotropically distributed under the choice of a Jacobian which is only a function of z for any fiducial parametrization of an FLRW metric system. However, we expect minor changes from different parametrisation.

On the other hand, one can observe that if we assume a different kind of model, we assume different jacobian $J(z, \theta, \phi)$. Then, we may end up in a different conclusion.

4.8 Conclusion and Discussion

We have used the data release 12 of BOSS CMASS galaxy sample to study the transition to cosmic homogeneity over a volume of $5.1 h^{-3} \text{Gpc}^3$. We do not consider the correlation function, $\xi(r)$, to study homogeneity because its definition requires an average density, which is only defined for a homogeneous sample. We rather use the counts-in-spheres, $N(< r)$, i.e. the average number of objects around a given object, and its logarithmic derivative, the fractal correlation dimension, $D_2(r)$, to test for a possible fractal universe. For a homogeneous sample, $N(< r) \propto r^3$ and $D_2 = 3$. We define a characteristic homogeneity scale, \mathcal{R}_H , as the value for which D_2 reaches the homogeneous value within 1%, i.e. $D_2(\mathcal{R}_H) = 2.97$.

For the galaxy distribution, we get $3 - \langle D_2 \rangle = (0.6 \pm 1.3) \times 10^{-3}$ at 1σ over the range $300 h^{-1} \leq r \leq 1300 h^{-1} \text{Mpc}$, which is consistent with homogeneity assumption of the universe and a transition to homogeneity at a characteristic scale, $\mathcal{R}_H^{(G)} = 114.2 \pm 5.8 h^{-1} \text{Mpc}$. However, in our analysis we use a random catalog to take into account the geometry and the completeness of the survey. The redshift distribution of this catalog is taken from the data. Therefore, this study is insensitive to a possible isotropic variation of the density with redshift, $\rho = \rho(z)$. This means that we can only check for spatial isotropy, $\rho(z, \theta_1) = \rho(z, \theta_2)$. Note that the same is true for all galaxy redshift surveys [185]. Contrary, this spatial isotropy can be obtained without using any fiducial model as discussed by Laurent et al. [166]. So by assuming the Copernican principle our data imply homogeneity of the galaxy sample without any prior ΛCDM assumption.

If we use a similar to Yadav et al. [158] estimator for the transition to homogeneity, we find agreement between the mocks and data. This estimator provides a qualitative estimate of the homogeneity scale in the range of $150 \leq r \leq 400 h^{-1}\text{Mpc}$. This further confirms the transition to homogeneity in both matter and galaxy distributions, despite the lower precision of this estimator, i.e. 30% compared to our estimator which is near %-level.

Furthermore, we can perform the measurement in an alternative way to check the ΛCDM model. In particular, we fit the CMASS galaxy two point correlation function in the range $1 < r < 40 h^{-1}\text{Mpc}$ to obtain the galaxy bias relative to the ΛCDM prediction for the matter correlation function as long as the peculiar velocities to correct for *Redshift Space Distortions*. We correct our measurement of $N(< r)$ for this bias and the peculiar velocities, in order to get the result for the matter distribution, that we finally compare to the ΛCDM prediction.

For the distribution of matter, we achieved a measurement of $3 - \langle \mathcal{D}_2(r > 300h^{-1}\text{Mpc}) \rangle_z = (1.7 \pm 1.0) \times 10^{-4}$ at 1σ . We obtained a characteristic scale $\mathcal{R}_H^{(M)} = 61.9 \pm 0.8 h^{-1}\text{Mpc}$ at an average $z = 0.538 - 0.592$ for the transition to homogeneity. This measurement of \mathcal{R}_H is more precise than the previous measurements of Scrimgeour et al. [93] Collaboration by a factor 5. While Sarkar et al. [178], only give a qualitative measurement of \mathcal{R}_H .

We also investigate the redshift evolution of our observables. We find that the homogeneity scale for the matter distribution, $\mathcal{R}_H^{(M)}$, is decreasing with time as expected if clustering is increasing with time. We find accordance with ΛCDM model with a reduced $\chi^2 = 0.89(0.61)$ for the North(South) Galactic Cap for the 6 redshift bins, i.e. accounting also the measurement of Laurent et al. [166] at redshifts $2.2 \leq z \leq 2.8$. However, the characteristic scale of the galaxy distribution, $\mathcal{R}_H^{(G)}$, remains almost constant with time or redshift, z . The homogeneity scale is an intrinsic property of the galaxy sample. Therefore, this result is consistent with the fact that the CMASS sample has properties that remain constant with time.

Moreover, we performed several test to ensure that the measurement is unbiased against several choices on the performance of the measurement. Firstly, we checked for a potential systematic against alternative modelling of *Redshift Space Distortions*. Furthermore, we ensure that measuring the fractal correlation dimension on several fractal distributions with the survey mask does not bias the intrinsic properties of these distributions, i.e. the fractal dimension of the distribution. Moreover, we perform the basis measurement of the homogeneity scale with different combination of the weighting scheme, observing no significant systematic errors. Furthermore, we check for alternative estimators of the homogeneity scale and we found that the optimum one (most accurate and the 2nd least precise) is the one we developed in this thesis. In a further study we found out that the bias and peculiar velocities, used to measure the homogeneity on the matter distribution are slightly correlated with the characteristic scale, i.e. $r \leq 20\%$ for all possible combination between of the quantities $(\mathcal{R}_H^{(G)}, \mathcal{R}_H^{(M)}, b, \sigma_p)$ at all redshift bins, rendering their correlation negligible. Additionally, we find that using a different definition for the homogeneity scale, i.e. the minimum scale at which the fractal correlation dimension becomes consistent with the nominal value within 0.1%, $\mathcal{D}_2(\mathcal{R}_H^{\mathcal{D}_2=2.997}) = 2.997$, namely per-Mille definition, gives scales which are more than twice the scales obtained with the adopted definition of our basis analysis, $\mathcal{D}_2(\mathcal{R}_H^{\mathcal{D}_2=2.97}) = 2.97$.

Furthermore, we find that the precision of the observables of the fractal correlation dimension and the scaled-counts-in-spheres for the matter distribution gains about 1 order of magnitude in respect of the same observable for the galaxy distribution, due to the bias correction, which is not the case for the 2-point-correlation function that it is being used in regular large scale structure analysis. This renders

our observables, counts-in-spheres and fractal correlation dimension a potential precision observables of cosmological inferences for large scale structure cosmography.

We also conclude that this measurement of the homogeneity scale is only a validation of the Cosmological Principle. This is true since we cannot infer distances from observational data, i.e. redshift, without the assuming in prior a Homogeneous universe with an FLRW metric. Therefore future projects should relax the homogeneity assumption and use the same observable and estimators for a model independent measurement of the characteristic scale of the transition to cosmic homogeneity for a confirmation test.

Finally, we are unable to perform measurements outside the observable universe. Therefore, we rely on the assumption that the observable universe is a fair sample, i.e. the properties of the observable universe are the same as the unobserved one. However, since the measurements done by Hogg et al. [83] in 2005 have a homogeneous universe at redshift, $z \simeq 0.3$. While in this study we have shown that the galaxies even beyond the observable universe, of 2005, i.e. the observable universe 2017 at redshift $z \simeq 0.56$, the universe has also a homogeneous behaviour. Therefore by induction, we conclude that the unobserved universe will look uniform as well, with a model dependent measurements of course.

This study was published in an article in JCAP that is attached in this manuscript.

Exploring cosmic homogeneity with the BOSS DR12 galaxy sample

Pierros Ntelis,^{a,1} Jean-Christophe Hamilton,^a Jean-Marc Le Goff,^b Etienne Burtin,^b Pierre Laurent,^b James Rich,^b Nicolas Guillermo Busca,^a Jeremy Tinker,^m Eric Aubourg,^a Hélion du Mas des Bourboux,^b Julian Bautista,^f Nathalie Palanque Delabrouille,^b Timothée Delubac,^g Sarah Eftekharzadeh,^d David W. Hogg,^c Adam Myers,^d Mariana Vargas-Magaña,^j Isabelle Pâris,^e Partick Petitjean,^l Graziano Rossi,^h Donald P. Schneider,^k Rita Tojeiroⁱ and Christophe Yèche^b

^aAPC, Université Paris Diderot-Paris 7, CNRS/IN2P3, CEA, Observatoire de Paris, 10, rue A. Domon & L. Duquet, Paris, France

^bIRFU, CEA, Université Paris-Saclay, F-91191, Gif-sur-Yvette, France

^cCenter for Cosmology and Particle Physics, New York University, 4 Washington Place, Meyer Hall of Physics, New York, NY 10003, USA

^dDepartment of Physics and Astronomy, University of Wyoming, Laramie, WY 82071, USA

^eAix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, 13388, Marseille, France

^fDepartment of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA.

^gLaboratoire d'astrophysique, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland

^hDepartment of Astronomy and Space Science, Sejong University, Seoul, 143-747, Korea

ⁱSchool of Physics and Astronomy, University of St Andrews, St Andrews, KY16 9SS, UK

^jInstituto de Física, Universidad Nacional Autónoma de México, Apdo. Postal 20-364, México

^kDepartment of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802

^lInstitut d'Astrophysique de Paris, CNRS-UPMC, UMR7095, 98bis bd Arago, Paris, 75014 France

^mDepartment of Physics and Center for Cosmology and Particle Physics, New York University

Abstract. In this study, we probe the transition to cosmic homogeneity in the Large Scale Structure (LSS) of the Universe using the CMASS galaxy sample of BOSS spectroscopic survey which covers the largest effective volume to date, $3 h^{-3} \text{ Gpc}^3$ at $0.43 \leq z \leq 0.7$. We study the scaled counts-in-spheres, $\mathcal{N}(< r)$, and the fractal correlation dimension, $\mathcal{D}_2(r)$, to assess the homogeneity scale of the universe using a *Landy & Szalay* inspired estimator.

Defining the scale of transition to homogeneity as the scale at which $\mathcal{D}_2(r)$ reaches 3 within 1%, i.e. $\mathcal{D}_2(r) > 2.97$ for $r > \mathcal{R}_H$, we find $\mathcal{R}_H = (63.3 \pm 0.7) h^{-1} \text{ Mpc}$, in agreement at the percentage level with the predictions of the ΛCDM model $\mathcal{R}_H = 62.0 h^{-1} \text{ Mpc}$. Thanks to the large cosmic depth of the survey, we investigate the redshift evolution of the transition to homogeneity scale and find agreement with the ΛCDM prediction. Finally, we find that \mathcal{D}_2 is compatible with 3 at scales larger than $300 h^{-1} \text{ Mpc}$ in all redshift bins.

These results consolidate the Cosmological Principle and represent a precise consistency test of the ΛCDM model.

DRAFT

Contents

1	Introduction	2
2	Dataset	4
2.1	The BOSS survey	4
2.2	Data sample	4
2.3	Mock-Catalogues	5
3	Methodology	6
3.1	Observables	6
3.2	Estimators	6
3.3	Homogeneity scale definition	7
4	Theoretical Model	8
4.1	Prediction for $\xi(r)$, $\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$	8
4.2	Correction for bias and RSD	9
5	Results	10
5.1	Analysis range	10
5.2	Covariance matrices	10
5.3	Homogeneity scale for the galaxy distribution	11
5.4	Determining the bias and the velocity dispersion	12
5.5	Homogeneity scale for the matter distribution	13
5.6	Consistency with Λ CDM	14
5.7	Constraints on fractal correlation dimension	14
6	Analysis Robustness	17
6.1	Bias in the fractal correlation dimension	17
6.2	Sensitivity to RSD model	18
7	Conclusions	19
A	From $\xi(r)$ to $\mathcal{N}(< r)$	21
B	Choice of estimator for $\mathcal{N}(< r)$	21
C	Tuning errors in RSD Analysis	21
D	Test of spline fit on QPM mock catalogues	22
E	Homogeneity scale at 0.1%	22

1 Introduction

Most models of modern Cosmology are based on solutions of General Relativity for an isotropic and homogeneous universe. The standard model, known as Λ CDM, is mainly composed of a Cold Dark Matter (CDM) and Λ corresponds to a cosmological constant. This model shows excellent agreement with current data, be it from Type Ia supernovae [1, 2], temperature and polarisation anisotropies in the Cosmic Microwave Background [3] or Large Scale Structure [4–6]. The two main assumptions of this model are the validity of General Relativity and the *Cosmological Principle* [7] that states that the Universe is isotropic and homogeneous on large enough scales, or equivalently that the Universe statistical properties are both rotationally and translationally invariant on large scales.

Isotropy is well tested through several probes at various cosmic epochs: at $z \approx 1100$, Cosmic Microwave Background temperature anisotropies, corresponding to density fluctuations in the young Universe, have been shown to be of order 10^{-5} [8]. In the more recent Universe, the distribution of sources in X-ray [9] and radio [10] surveys strongly supports isotropy. Large spectroscopic galaxy surveys, such as the baryon oscillation spectroscopic survey (BOSS) of the third Sloan digital sky survey (SDSS-III), show no evidence for anisotropies in the projected galaxy distribution in volumes of a few Gpc^3 [11].

We should stress, however, that these observations test isotropy at a given redshift or after projection over a given range of redshifts. This "projected" isotropy is a weaker assumption than "spatial" isotropy, which is an isotropy at each redshift¹. Combining spatial isotropy and the Copernican principle, which states that our position in the Universe is not privileged, implies that the Universe is homogeneous [7, 12–14]. However, as shown by Durrer et al. [15], this implication is not true if we only have projected isotropy. So CMB isotropy, for instance, cannot be combined with the Copernican principle to prove homogeneity.

It is therefore important to test homogeneity. Large three dimensional spectroscopic surveys offer an excellent occasion to strengthen this aspect of the cosmological model with accurate observations. Most of the studies conducted so far found a transition to homogeneity in the galaxy distributions at scale between 70 and 150 h^{-1} Mpc [16–27] therefore strengthening the cosmological principle. Some studies did not find a transition to homogeneity [28–34]. However none of those studies reached scale larger than 200 h^{-1} Mpc, where homogeneity becomes clear.

However, these 3D surveys investigate the statistical properties on the observer past light-cone and not inside of it, so they do not observe galaxies at the same epoch. An attempt to overcome this limitation was to use star formation history in order to study the homogeneity inside the past light-cone [35], but this is model dependent. Another possibility [36] is to use a combination of secondary CMB probes, including integrated Sachs-Wolfe, kinetic Sunyaev-Zel'dovich and Rees-Sciama effects.

In this article, we use the Data Release 12 of the SDSS-III/BOSS spectroscopic galaxy sample to search for a transition to homogeneity [37]. The BOSS CMASS galaxy catalogue covers a volume of 5.1 h^{-3} Gpc^3 in the redshift range $0.43 < z < 0.7$ and contains nearly one million galaxies in 8500 square degrees. It represents the largest effective volume sampled as of today² and therefore offers a unique opportunity to observe the transition to homogeneity.

¹More precisely, spatial isotropy is the assumption that $\rho(r, \theta_1) = \rho(r, \theta_2)$ for any given (r, θ_1, θ_2) , while projected isotropy is the assumption that $\rho(\theta_1) = \rho(\theta_2)$ for any (θ_1, θ_2) , where $\rho(\theta) = \int \rho(r, \theta)W(r)dr$ and $W(r)$ is the window function.

²The effective volume of a survey is given by $V_{\text{eff}} = V \frac{nP}{1+nP}$, which describes the statistical power of the

A similar study of homogeneity has been performed with the DR12 quasar sample of BOSS [38]. We mostly follow the method introduced in [26] to measure the "fractal correlation dimension" of the distribution, $\mathcal{D}_2(r)$, as an indicator of the transition to homogeneity [37].

Due to the possible evolution of the tracer with redshift, a redshift survey is blind to strictly radial variation of the density [39]. So strictly speaking we can only test homogeneity up to a radial variation, that is we can test spatial isotropy (and then combining spatial isotropy with the Copernican principle we can prove homogeneity). On the other hand, as discussed by Laurent et al. [38], we can demonstrate spatial isotropy independently of a fiducial cosmology ³.

Throughout this study, unless stated otherwise, we use a fiducial flat Λ CDM cosmological model with the following parameters:

$$\mathbf{P}_{\text{cosmo}} = (\omega_{\text{cdm}}, \omega_b, h, n_s, \ln [10^{10} A_s]) = (0.1198, 0.02225, 0.6727, 0.9645, 3.094) , \quad (1.1)$$

where $\omega_b = \Omega_b h^2$ and $\omega_{\text{cdm}} = \Omega_{\text{cdm}} h^2$ are the reduced fractional density of baryons and cold dark matter, respectively; $h = H_0 / [100 \text{ km s}^{-1} \text{ Mpc}^{-1}]$, with H_0 the Hubble constant; and finally n_s and A_s are, respectively, the spectral index and the amplitude of the primordial scalar power spectrum. The numerical values in Eq. 1.1 are from Planck 2015 $TT, TE, EE + lowP$ analysis [40].

The paper is organized as follows. Section 2 describes the galaxy data sample used in this analysis and section 3 the methodology to quantify cosmic homogeneity. In section 4 we introduce the model that is adjusted to the observed data, while in section 5 we present our results. In section 6 we test the robustness of our method, and we conclude in section 7.

sample. The CMASS sample has $V_{\text{eff}}^{\text{CMASS}} \simeq 2.9h^{-3} \text{Gpc}^3$, while the QSO sample has $V_{\text{eff}}^{\text{QSO}} \simeq 0.21h^{-3} \text{Gpc}^3$.

³ Indeed let's assume a homogeneous 3D spectroscopic survey with a constant density $\rho(r, \theta, \phi) = \text{Cst}$. This density is the product of the Jacobian $J(z) = H(z)/cD^2(z)$ with the observed density of sources $dN(\theta, \phi, z)/d\Omega dz$ which is a function of angles and redshift. As a result the observed density is necessarily a function of redshift only following: $dN(\theta, \phi, z)/d\Omega dz \propto J^{-1}(z)$, which is the definition of spatial isotropy.

2 Dataset

2.1 The BOSS survey

BOSS[41] is dedicated to studying the 3D distribution of $\sim 1.4 \times 10^6$ galaxies and $\sim 10^5$ quasars and their Lyman- α forests within an effective area of $\sim 10,000 \text{ deg}^2$ [42, 43]. Aluminum plates are set on the focal plane of the 2.5m telescope[44] and drilled with 1000 holes corresponding to targets to be observed. An optical fiber is fixed on each hole, allowing 1000 spectra to be measured per exposure [45] with a typical redshift uncertainty of a few tens of $\text{km} \cdot \text{s}^{-1}$ [46]. Several color cuts are used to select massive galaxy in the redshift range $0.43 < z < 0.7$ in the imaging data from SDSS-I-II and SDSS-III/BOSS in (u, g, r, i, z) bands [47]. These cuts are designed to result in a stellar mass limit that is constant with redshift, according to a passively evolving model [48]. The large masses ensure a strong bias with respect to the underlying dark matter field, providing a high signal-to-noise ratio on the two-point correlation function. The resulting sample of galaxies is called CMASS and consists in $\sim 10^6$ galaxies with accurate spectroscopic redshifts. The CMASS sample has been used to measure the Baryonic Acoustic Oscillation feature (BAO) with unprecedented accuracy, putting strong constraints on cosmological parameters [49]. We use the same galaxy sample to perform a measurement of the transition to homogeneity benefiting from the signal-to-noise enhancement due to the large bias of this galaxy sample.

We weight galaxies with [50]:

$$w_{gal} = (w_{cp} + w_{noz} - 1) \times w_{star} \times w_{see} \times w_{FKP} , \quad (2.1)$$

Here, the close-pair weight, w_{cp} , accounts for the fact that, due to fiber coating, one cannot assign optical fibers on the same plate to two targets that are closer than $62''$. The w_{noz} weight accounts for targets for which the pipeline failed to measure the redshift. The w_{star} and w_{see} weights correct for the dependance of the observed galaxy number density with the stellar density and with seeing, respectively. Finally, we use the FKP weight, w_{FKP} , [51] in order to reduce the variance of the two-point correlation function estimator.

z	NGC	SGC
0.430-0.484	101,383	40,170
0.484-0.538	174,468	63,518
0.538-0.592	151,084	56,805
0.592-0.646	97,155	37,179
0.646-0.700	47,289	17,899
0.430-0.700	571,379	215,571

Table 1. DR12 data sample in 5 redshift intervals in north (NGC) and south (SGC) galactic caps.

2.2 Data sample

We divide our data sample into 5 redshift intervals, as defined in table 1, to study the evolution of the clustering of the CMASS galaxy sample. The angular and redshift distributions of the sample are shown in figure 1. We use a flat Λ CDM model with parameters defined by equation 1.1 to convert the redshift measurements to comoving distances:

$$d_{\text{comov}}(z) = \frac{c}{H_0} \int_0^z \frac{dz'}{E(z')} , \quad (2.2)$$

where $E(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}$. A final sample with a total effective volume of $3.8 h^{-3} \text{Gpc}^3$ [50] is obtained, which is significantly larger than the effective volume used in previous studies, such as $0.6 h^{-3} \text{Gpc}^3$ for WiggleZ [26], $1 h^{-3} \text{Gpc}^3$ for a DR7 LRG galaxy sample study [27] or $0.2 h^{-3} \text{Gpc}^3$ for the SDSS II LRG analysis [16].

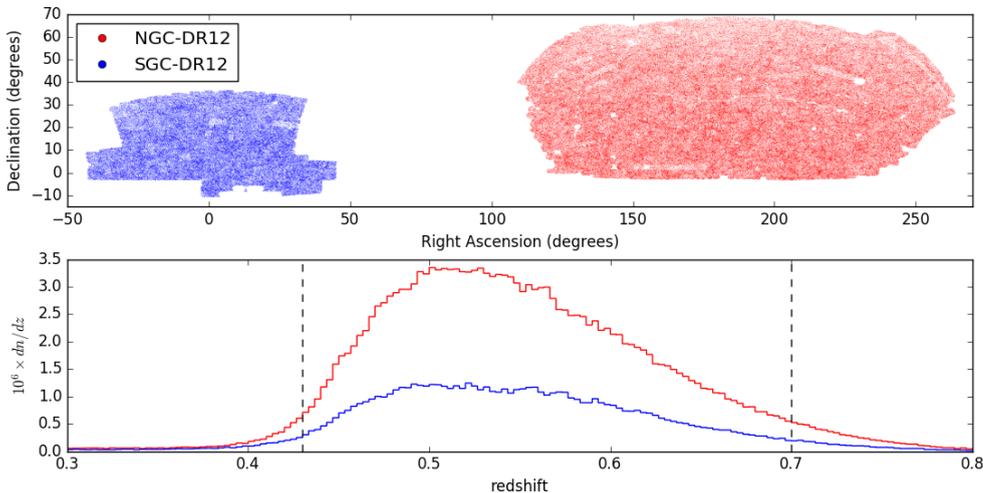


Figure 1. Right ascension and declination profiles (top) and redshift distribution (bottom) of the DR12 sample for the North (NGC) and South (SGC) Galactic Caps. The vertical dashed lines define the redshift interval used in our analysis.

2.3 Mock-Catalogues

Mock catalogues are an important tool for determining uncertainties in galaxy surveys and tuning data analysis pipelines. In our analysis we use 1000 mock catalogues constructed with the quick-particle-mesh (QPM) algorithm [52] for BOSS. These are expected to be more realistic than mocks based on second-order Lagrangian perturbation theory, although not as much as the highly time-consuming N-body simulations. These mock catalogues use a different flat Λ CDM cosmology than the one we use in our analysis, but this difference is accounted for in our analysis:

$$\mathbf{p}_{\text{qpm}} = (\Omega_{\text{cdm}}, \Omega_b, h, n_s, \sigma_8) = (0.274, 0.046, 0.7, 0.95, 0.8) \quad (2.3)$$

where σ_8 is the variance of the matter power spectrum computed in spheres of $8 h^{-1} \text{Mpc}$ radius.

3 Methodology

In this section we present the method we have used to study the transition to homogeneity in the distribution of the BOSS CMASS galaxy catalogue. We follow Scrimgeour et al. [26] in assessing the homogeneity of the catalogue through the fractal dimension, $D_2(r)$, but use significantly different methods.

3.1 Observables

The ‘‘counts-in-spheres’’, $N(< r)$, is defined as the average number of objects in a sphere of radius r , it should obviously scale as r^3 for an homogeneous distribution. This scaling is the basis for the definition of the fractal dimension:

$$D_2(r) \equiv \frac{d \ln N(< r)}{d \ln r}. \quad (3.1)$$

For a homogeneous distribution one gets $D_2(r) = 3$, while for a fractal distribution $D_2 < 3$ is a measure of the fractal dimension. These simple expressions are no longer valid in the case of a survey that has a peculiar geometry and a non-uniform completeness. We correct for these geometrical effects by using catalogues of random points that uniformly fill the survey volume. We use random samples that are five times larger than the data sample in order to ensure that statistical fluctuations due to the random points are significantly smaller than those due to the data. We therefore define a ‘‘scaled counts-in-spheres’’

$$\mathcal{N}(< r) = \frac{N_{gal}(< r)}{N_{rand}(< r)} \quad (3.2)$$

which is the ratio between the counts-in-spheres of the galaxy distribution, $N_{gal}(< r)$, and the counts-in-spheres of the random point distribution inside the same survey geometry, $N_{rand}(< r)$. This quantity is now expected to be independent of r (ratio of two quantities $\propto r^3$) for a homogeneous distribution and therefore the fractal dimension is redefined as:

$$\mathcal{D}_2(r) \equiv \frac{d \ln \mathcal{N}(< r)}{d \ln r} + 3, \quad (3.3)$$

in order to be equal to 3 for a homogeneous distribution.

$\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$ are the two observables we will consider in the analysis, and we show in the next section that there are different ways to obtain them from the data.

3.2 Estimators

We define the following quantities from our data and random catalogues:

- $dd(r) = \frac{DD(r)}{n_g(n_g - 1)/2}$: the normalized number of galaxy pairs distant by r ,
- $rr(r) = \frac{RR(r)}{n_r(n_r - 1)/2}$: the normalized number of random-point pairs distant by r ,
- $dr(r) = \frac{DR(r)}{n_g n_r}$: the normalized number of galaxy random-point pairs distant by r ,

where n_g and n_r are the total number of galaxies and random points, respectively. The definition of $\mathcal{N}(< r)$ in Eq. 3.2 recalls the Peebles-Hauser estimator, $\widehat{\xi}(r) = dd(r)/rr(r) - 1$, for the two-point correlation function [53]. This estimator is known to be less efficient than the more sophisticated Landy-Szalay estimator [54].

$$\widehat{\xi}_{ls}(r) = \frac{dd(r) - 2dr(r) + rr(r)}{rr(r)}, \quad (3.4)$$

which has minimal variance on scales where $\xi(r) \ll 1$ [54].

Laurent et al. [38] defines a Landy-Szalay inspired estimator for the counts-in-spheres:

$$\widehat{\mathcal{N}}_{lau}(< r) = 1 + \frac{\int_0^r (dd(s) - 2dr(s) + rr(s)) ds}{\int_0^r rr(s) ds}. \quad (3.5)$$

Alternatively, we can directly compute the counts-in-spheres from the Landy-Szalay estimator of the two-point correlation function itself, as explained in appendix A:

$$\widehat{\mathcal{N}}_{cor}(< r) = 1 + \frac{3}{r^3} \int_0^r \widehat{\xi}_{ls}(s) s^2 ds. \quad (3.6)$$

These two estimators are expected to be closer to optimal than the estimator of equation 3.2. They result in two different estimators for the fractal dimension that we name accordingly $\widehat{\mathcal{D}}_{2,lau}(r)$ and $\widehat{\mathcal{D}}_{2,cor}(r)$. In appendix B, we show that the *cor*-estimators, $\widehat{\mathcal{N}}_{cor}(< r)$ and $\widehat{\mathcal{D}}_{2,cor}(r)$, are less biased than the *lau*-estimators. In our final analysis, unless stated otherwise, we use the *cor*-estimators and we drop the *lau* indice and the hat for sake of simplicity.

Since we are interested in matter distribution, we must use the galaxy bias, b , to convert the counts-in-spheres measured for galaxies to the one we would get if we were measuring the whole matter (mostly dark matter). We also have to account for redshift space distortions and we explain in details how we do that in section 5.4.

3.3 Homogeneity scale definition

Following Scrimgeour et al. [26], we define an homogeneity scale R_H for both \mathcal{N} and \mathcal{D}_2 observables, as the scale for which the observable approaches its homogeneous value within 1%. We then have two a-priori different transition-to-homogeneity scales defined by:

$$\mathcal{D}_2(R_H^{\mathcal{D}_2=2.97}) = 2.97 \quad \text{and} \quad \mathcal{N}(R_H^{\mathcal{N}=1.01}) = 1.01. \quad (3.7)$$

These are, of course, arbitrary definitions for the homogeneity scales but they do not depend on the survey and can be used to test cosmological models and compare different survey measurements as long as the same definitions are used in all cases [26].

Other authors have proposed different ways of defining the homogeneity scale. In particular Yadav et al. [18] proposed to define it as the scale at which \mathcal{D}_2 cannot be distinguished from 3 within the survey errors. This is obviously another proper definition because it is independent of the bias and similar conclusions can be drawn as we demonstrate at the end of section 5.7. Furthermore, using this estimator on the mock catalogues (See section 5.7), we find larger variations on the estimated homogeneity scale in respect with the one obtained by Eq. 3.7. For these reasons, we have made the choice in this study to use the arbitrary but universal definition in Eq. 3.7, although we show that beyond $300 h^{-1}\text{Mpc}$, we do find consistency with $\mathcal{D}_2 = 3$ within our measurement errors.

4 Theoretical Model

4.1 Prediction for $\xi(r)$, $\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$

We use the CLASS software [55] to compute the Λ CDM theoretical prediction for the two-points correlation function and the observables we consider in the analysis. CLASS computes the theoretical matter power spectrum, $P_{\delta\delta}^{(r)}(k)$, where the (r) exponent indicates that it is calculated in real space, by opposition to redshift space. In order to get a prediction for the redshift-space galaxy power-spectrum, $P_{gg}^{(s)}(k)$ where the (s) exponent indicates redshift space, we model redshift-space-distortion (RSD):

- On large scales, galaxies are falling into large gravitational potentials, which tends to sharpen their distribution along the line-of-sight in redshift space. This is known as the Kaiser effect [56], which results in:

$$P_{gg}^{(s)}(k, \mu; b) = b^2(1 + \beta\mu^2)^2 P_{\delta\delta}^{(r)}(k), \quad (4.1)$$

where b is the linear (scale independent) bias between galaxy and matter distributions, $\beta = f/b$ where f is the linear growth rate, which can be approximated by $f \approx \Omega_m^{0.55}(z)$ [57] and $\mu = \cos\theta$, with θ the angle relative to the line-of-sight.

- On small scales, galaxies have a velocity dispersion whose projection on the line-of-sight in redshift space gives rise to the "finger-of-God" effect (FoG). These distortions can be modeled with a simple Gaussian orientation-dependent and scale-dependent damping model, which takes into account the pairwise peculiar-velocity dispersion of the galaxies σ_p [58]:

$$\ln D(k, \mu; \sigma_p) = -\frac{1}{2} \left(\frac{k\sigma_p\mu}{H_0} \right)^2. \quad (4.2)$$

This damping factor represents well actual data down to scales where $k\sigma_p \sim H_0$ [58], which corresponds to $r \sim 15h^{-1}\text{Mpc}$.

Finally, accounting for both effects and integrating over all orientations relative to the line-of-sight, the redshift-space galaxy power spectrum reads:

$$P_{gg}^{(s)}(k; b, \sigma_p) = b^2 \int_0^1 (1 + \beta\mu^2)^2 D(k, \mu; \sigma_p) d\mu P_{\delta\delta}^{(r)}(k), \quad (4.3)$$

which can be analytically integrated over μ , leading to:

$$P_{gg}^{(s)}(k; b, \sigma_p) = P_{\delta\delta}^{(r)}(k) \times \frac{b^2 H_0}{2k^5 \sigma_p^5} \times \left[-2 \exp\left(-\frac{k^2 \sigma_p^2}{2H_0^2}\right) k\beta\sigma_p H_0 (k^2(2 + \beta)\sigma_p^2 + 3\beta H_0^2) + \sqrt{2\pi} \text{Erf}\left(\frac{k\sigma_p}{\sqrt{2}H_0}\right) (k^4 \sigma_p^4 + 2k^2 \beta \sigma_p^2 H_0^2 + 3\beta^2 H_0^4) \right] \quad (4.4)$$

Applying a Fast Fourier Transform to equation (4.4) results in the two-point correlation function with two parameters b and σ_p :

$$\xi^{(s)}(r; b, \sigma_p) = FFT \left[P_{gg}^{(s)}(k; b, \sigma_p) \right]. \quad (4.5)$$

Then we use equation (3.6) and (3.3) to compute the Λ CDM prediction for $\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$.

4.2 Correction for bias and RSD

The estimators presented in section 3.2 measure the clustering properties of the galaxy distribution, we need to convert them into estimators that describe the clustering properties of the total matter distribution. We fit $\xi(r)$ in the range $1 h^{-1} < r < 40 h^{-1}$ Mpc with our model of equation 4.5 to obtain b and σ_p . If we take into account only the bias and the Kaiser effect, the angle-averaged correlation function is just multiplied by a constant factor, a squared effective bias. Since this factor is independent of r , the same factor applies to $\mathcal{N}(< r) - 1$, which is an integral over $\xi(r')$ (see equation 3.6). Taking into account also finger-of-God effect, the multiplicative factor is no longer independent of r . To transform the estimated $\mathcal{N}(< r) - 1$ for galaxies into an estimation for matter, we multiply it by the ratio of our model for $\mathcal{N}(< r) - 1$ for matter (i.e. $b = 1$ and $\sigma_p = 0$) to our model for the best fit value of b and σ_p , as

$$\widehat{\mathcal{N}}(< r) = \frac{\mathcal{N}^{\text{model}}(< r; b = 1, \sigma_p = 0) - 1}{\mathcal{N}^{\text{model}}(< r; b, \sigma_p) - 1} \times [\widehat{\mathcal{N}}_{gal}(< r) - 1] + 1. \quad (4.6)$$

Then the fractal correlation dimension $\widehat{\mathcal{D}}_2(r)$ is obtained from $\widehat{\mathcal{N}}(< r)$ using equation 3.3.

Section 6.2 shows that taking into account the finger-of-God effect only contributes in a small change in the measurement of the homogeneity scale by typically 1%. So the error due to the imperfection in the modelling of this effect is negligible.

DRAFT

5 Results

In this section, we present the results of our analysis. We determine the range in comoving distance for which we can measure $\mathcal{N}(< r)$ with our sample. We quantify the uncertainties in our measurements using mock catalogues. We measure the homogeneity scale for the galaxy distribution, we fit the redshift-space-distortion parameters, measure the homogeneity scale for matter distribution. Finally we make comparisons with the Λ CDM model and estimate the average of \mathcal{D}_2 on large scales.

5.1 Analysis range

Figure 2 shows the number density of random pairs, dN/dr , divided by r^2 and normalized such that at small scales it is unity. This scaled number density is constant at small scales and then decreases with r due to the finite size of the survey. We analyse data up to a maximum $r = 1300 h^{-1}$ Mpc, where the scaled density goes down to 1%.

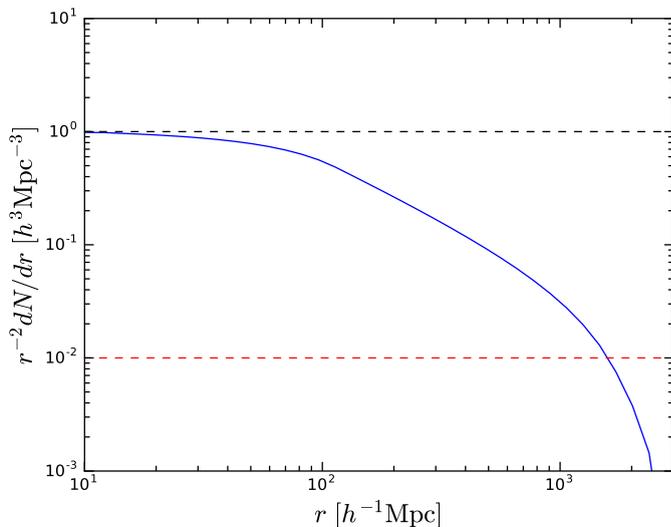


Figure 2. Scaled number density of random pairs, $r^{-2}dN/dr$, versus the comoving radius of spheres, r , normalized to be unity at small r .

5.2 Covariance matrices

We use a set of 1000 QPM mock catalogues to estimate the covariance matrices. We compute our observables for each of the catalogues and derive the bin-to-bin covariance matrices for each of the relevant observables, $\xi(r)$, $\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$.

Fig. 3 displays the resulting correlation matrices⁴ in the redshift bin $0.538 \leq z \leq 0.592$. The correlations matrices are similar in the other redshift bins. $\mathcal{N}(r)$ is more correlated on large scales than $\xi(r)$ because it is an integral over $\xi(r)$. On the other hand $\mathcal{D}_2(r)$, which is a derivative of $\mathcal{N}(r)$, is not very correlated. Therefore, when studying the redshift evolution in section 5.5, we consider only the homogeneity scale obtained with $\mathcal{D}_2(r)$, and not with $\mathcal{N}(r)$.

⁴We show the correlation matrices, $\rho_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}$, which have less dynamics than the covariance matrices.

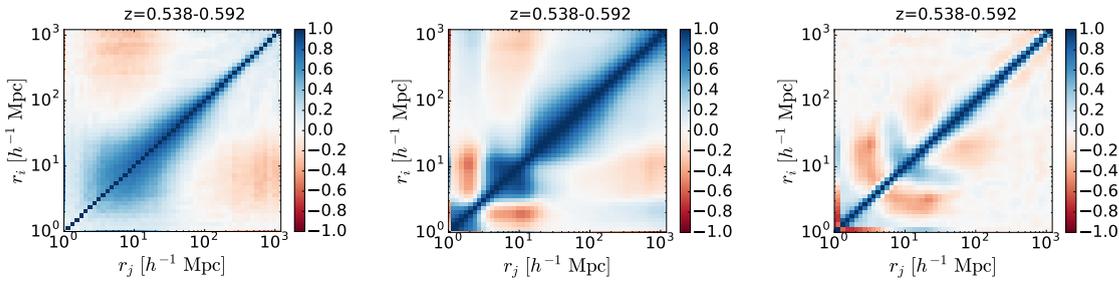


Figure 3. Correlation matrices for $\xi(r)$, $\mathcal{N}(<r)$ and $\mathcal{D}_2(r)$ for the total matter distribution in the redshift range $0.538 \leq z \leq 0.592$.

5.3 Homogeneity scale for the galaxy distribution

We first compute $\mathcal{N}^{gal}(<r)$ and $\mathcal{D}_2^{gal}(r)$ for the CMASS galaxy distribution using equation 3.6 and 3.3, without correcting for bias and redshift space distortions. This provides a measurement that does not rely on Λ CDM model to determine the bias. Results are shown in figure 4 for the $0.430 \leq z \leq 0.484$ redshift interval in the NGC. As described in section 5.2, the error are obtained from the 1000 QPM mock catalogues.

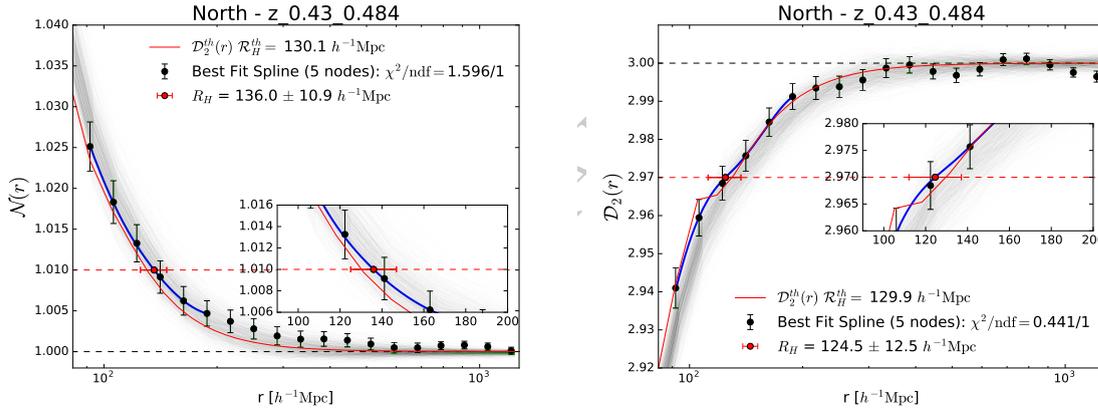


Figure 4. Scaled counts-in-spheres, $\mathcal{N}(<r)$, (left) and fractal correlation dimension, $\mathcal{D}_2(r)$, (right). The black points with error bars are the result for the galaxy distribution in the NGC in the $0.430 \leq z \leq 0.484$ redshift bin. The blue lines are the best spline fit of the data. The green lines are the results for the 1000 QPM mock catalogues. The red continuous line is the Λ CDM prediction for $b = 1.95$ and $\sigma_p = 252$ km/s.

The horizontal black-dashed lines in figure 4 indicate the value of the corresponding observable for an homogeneous distribution. The two observables reach homogeneity on large scales. The intersection of the data with the red-dashed lines (at $\mathcal{N} = 1.01$ and $\mathcal{D}_2 = 2.97$) defines the homogeneity scale. In order to determine it, we perform a spline fit over 6 data point around the intersection. The number of nodes of the spline fit is set to 5 in order to get an average χ^2 equal to the number of degrees of freedom for the 1000 QPM mocks. The error on R_H was obtained from the errors in the spline factors by error propagation. The results, presented in Table 3 for the case of \mathcal{D}_2 , are consistent with our Λ CDM model predictions. We finally stress the fact that the behaviour of the data is consistent with that of the 1000 QPM mock catalogues.

5.4 Determining the bias and the velocity dispersion

As illustrated in figure 5, we fit our model (Eq. 4.5) to the measured correlation function for the CMASS galaxy sample, in order to determine the galaxy bias and velocity dispersion. As explained in Appendix C, we account for theoretical uncertainties in the RSD modelling by boosting the covariance matrix on the relevant small scales. We thus achieve a good $\chi^2/\text{n.d.f.}$ in the range $[1, 40] h^{-1} \text{Mpc}$. Appendix C shows that this change in the covariance matrix at small scales has a negligible effect on the measurement of the homogeneity scale.

Results are given in table 2. The mean precision is 2.6% (NGC) and 4.1% (SGC) for the bias and 12% (NGC) and 23% (SGC) for the velocity dispersion. The bias is in agreement with values obtained by several authors [58–63] and the velocity dispersion is consistent with the typical CMASS-galaxy-sample velocity-dispersion, $\sigma_p \approx 240 \pm 50 \text{ km/s}$ [64].

z	NGC			SGC		
	b	σ_p [km/s]	χ_{red}^2	b	σ_p [km/s]	χ_{red}^2
0.430-0.484	1.879 ± 0.023	243.0 ± 9.7	0.44	1.872 ± 0.035	238.7 ± 15.2	0.85
0.484-0.538	1.846 ± 0.021	247.1 ± 7.4	0.67	1.840 ± 0.032	234.6 ± 12.1	0.66
0.538-0.592	1.944 ± 0.021	252.0 ± 8.9	0.91	1.943 ± 0.032	236.2 ± 13.7	0.59
0.592-0.646	1.995 ± 0.022	234.2 ± 10.9	1.47	2.001 ± 0.035	237.9 ± 19.6	0.79
0.646-0.700	2.153 ± 0.028	235.3 ± 21.7	0.79	2.081 ± 0.045	210.3 ± 40.1	0.50

Table 2. Fitted values of bias, b , and velocity dispersion, σ_p , in the different redshift bins, together with the corresponding reduced χ^2 for 24 degrees of freedom.

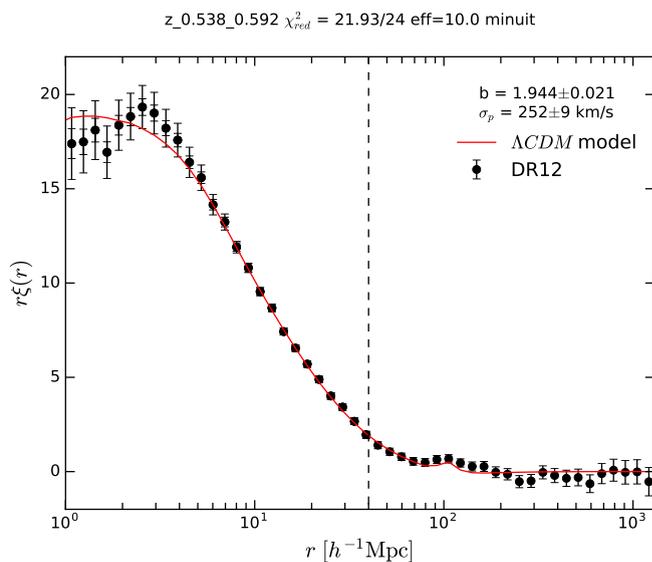


Figure 5. The correlation function of the CMASS galaxies in the $[0.538 - 0.592]$ redshift range. Data points have both their original error bars and the error bars enlarged to take into account the uncertainty of the RSD model on small scales (see section 5.4 and appendix C). The red line is the result of the fit performed over the range $[1, 40] h^{-1} \text{Mpc}$ (up to the vertical black-dashed line).

5.5 Homogeneity scale for the matter distribution

We use equation (4.6) with parameter b and σ_p obtained in section 5.4 to transform $\mathcal{N}^{gal}(< r)$ into $\mathcal{N}(< r)$ for matter. We then use Eq 3.3 to get $\mathcal{D}_2(r)$ for matter distribution. Results are shown in figure 6 for the redshift interval $0.538 \leq z \leq 0.592$. The two observables indicate homogeneity on large scales in this redshift interval, and in the four other intervals as well. As in section 5.3 we fit the data points to determine the homogeneity scales. We stress that the fit range, $40 < r < 100 h^{-1}$ Mpc, does not overlap with the fit range for determining the bias, $1 < r < 40 h^{-1}$ Mpc.

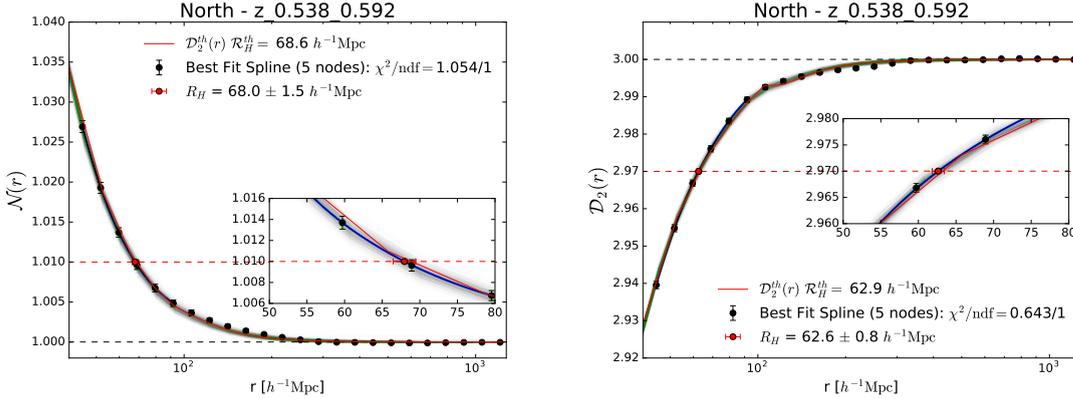


Figure 6. Same as figure 4 for matter distribution in the redshift interval $0.538 \leq z \leq 0.592$.

The results, presented in Table 3 for the case of \mathcal{D}_2 , are consistent with our Λ CDM model predictions. In the redshift interval $0.538 \leq z < 0.592$ we get a precision of 1.6%, which is a factor 5 better than Scrimgeour et al. [26] in spite of their wider redshift range, $0.5 < z < 0.7$. The more recent analysis by Sarkar et al. [27] does not give R_H for matter distribution.

Figure 7 shows that the measured homogeneity scale is compatible with Λ CDM, with reduced χ^2 smaller than unity both in the NGC and the SGC. The scale is increasing with time, from $z = 0.7$ to $z = 0.43$, as expected when clustering is increasing with time. The reduced χ^2 remains smaller than unity when adding the result from Laurent et al. [38], $\mathcal{R}_H(z = 2.4) = 26.2 \pm 0.9 h^{-1}$ Mpc.

z	$\mathcal{R}_H[h^{-1}\text{Mpc}]$ Matter		$\mathcal{R}_H[h^{-1}\text{Mpc}]$ Galaxy	
	NGC	SGC	NGC	SGC
0.430 – 0.484	64.2 ± 1.3	66.7 ± 1.6	124.5 ± 12.5	121.1 ± 9.8
0.484 – 0.538	65.4 ± 0.9	63.9 ± 1.5	111.9 ± 4.9	119.8 ± 8.8
0.538 – 0.592	62.6 ± 0.8	65.2 ± 1.6	116.4 ± 7.8	110.5 ± 5.1
0.592 – 0.646	60.4 ± 0.8	60.1 ± 1.1	108.8 ± 3.9	120.1 ± 11.7
0.646 – 0.700	59.0 ± 0.8	60.1 ± 1.8	125.8 ± 7.3	147.4 ± 8.4

Table 3. Homogeneity scale, $R_H^{D_2=2.97}(z)$, for the galaxy and matter distributions in the north and south galactic caps.

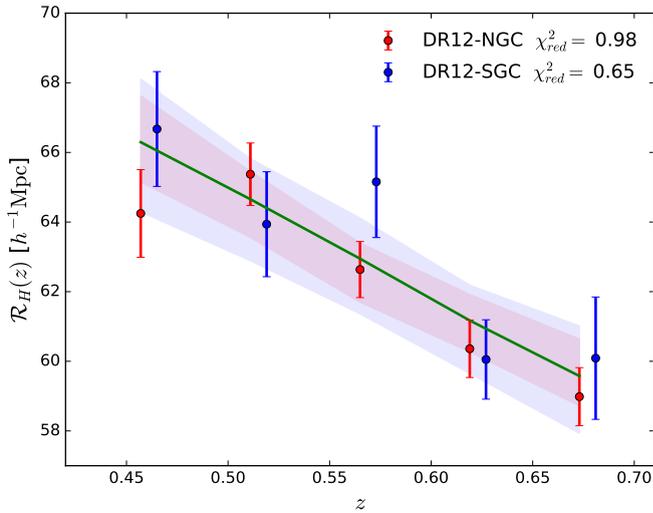


Figure 7. The homogeneity scale $R_H^{\mathcal{D}_2=2.97}(z)$ measured in the NGC (red) and in the SGC (blue) as a function of redshift. The green line is the Λ CDM model prediction. The shaded areas indicate the 1σ range for the 1000 QPM mock catalogues.

5.6 Consistency with Λ CDM

In order to characterize more precisely the agreement with the Λ CDM model, we fit $\mathcal{D}_2(r)$ from the data and the 1000 mock catalogues in the range $40 h^{-1} < r < 1300 h^{-1}$ Mpc with

$$\mathcal{D}_2^{\Lambda\text{CDM}} = \mathcal{D}_2(ar; \mathbf{p}_{\text{cosmo}}). \quad (5.1)$$

We fix the cosmological parameters $\mathbf{p}_{\text{cosmo}}$ and leave only free the a .

Figure 8 shows the results of the fits in different redshift bins for both NGC and SGC. The values of a , all consistent with 1, demonstrate a good agreement with Λ CDM model. Furthermore, the bottom plot of the figure shows that, for the data, the χ^2 are consistent with the number of degrees of freedom (23), as illustrated by the grey shaded area that indicates the expected 1σ extension, $\chi^2 = 23 \pm \sqrt{2} \times 23$. The mock catalogues, however, are a bit at variance with this grey area at large z .

5.7 Constraints on fractal correlation dimension

z	Matter		Galaxy	
	NGC ($\times 10^{-4}$)	SGC ($\times 10^{-4}$)	NGC ($\times 10^{-3}$)	SGC ($\times 10^{-3}$)
0.430 – 0.484	2.8 ± 3.3	2.5 ± 8.1	1.3 ± 1.5	1.1 ± 3.7
0.484 – 0.538	2.0 ± 3.1	1.9 ± 5.4	0.9 ± 1.4	0.8 ± 2.4
0.538 – 0.592	3.6 ± 1.7	2.4 ± 9.9	0.2 ± 0.9	1.2 ± 4.9
0.592 – 0.646	4.1 ± 2.0	0.9 ± 4.5	2.1 ± 1.1	0.5 ± 2.3
0.646 – 0.700	1.7 ± 3.8	-0.2 ± 2.8	1.0 ± 2.3	0.1 ± 1.6

Table 4. Fractal correlation dimension, $3 - \langle \mathcal{D}_2 \rangle_r(z)$, averaged over $300 h^{-1} < r < 1300 h^{-1}$ Mpc, with 1σ errors, in the NGC and SGC.

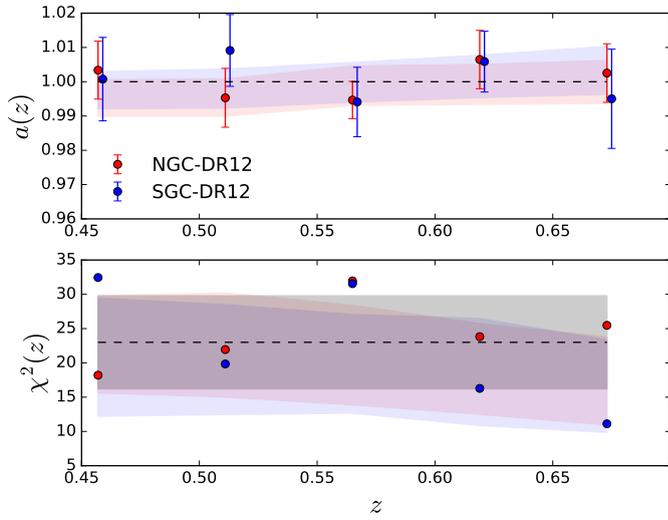


Figure 8. Top: value of the parameter α resulting from the fit of the NGC (red) and SGC (blue) $\mathcal{D}_2(r)$ data with the model of equation 5.1, in different redshift bins. The shaded areas correspond to the 1σ region for the 1000 QPM mock catalogues. Bottom: The corresponding χ^2 .

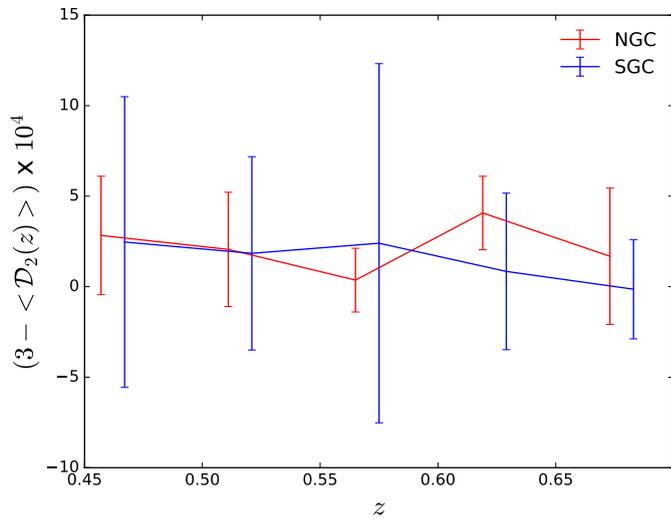


Figure 9. Fractal correlation dimension for the matter distribution, $3 - \langle \mathcal{D}_2 \rangle$, averaged over $300 h^{-1} < r < 1300 h^{-1}$ Mpc, in the different redshift bins.

In order to assess the level of homogeneity of the CMASS DR12 galaxy sample we compute the average of $\mathcal{D}_2(r)$ over the range $300 h^{-1} < r < 1300 h^{-1}$ Mpc, accounting for the covariance matrix. The results are presented in table 4 for the different redshift bins in the NGC and SGC. All results are compatible and we average them to get $3 - \langle \mathcal{D}_2 \rangle_{r,z} = (0.9 \pm 1.2) \times 10^{-3}$ (1σ).

Strictly speaking, it is not possible to transform the results for the galaxy distribution into results for matter distribution without using Λ CDM prediction to compute the galaxy

bias. However, Laurent et al. [38] show that, with a set of reasonable assumptions independent of Λ CDM, it is possible to obtain a lower limit for the tracer bias. In our case this results in $b > \sqrt{1.6}$ and $3 - \langle \mathcal{D}_2 \rangle < 7.12 \pm 9.48 \times 10^{-4}$ for matter distribution.

Alternatively we can assume Λ CDM, use it to get the galaxy bias and compute $\mathcal{D}_2(r)$. We average $\mathcal{D}_2(r)$ over $300 h^{-1} < r < 1300 h^{-1}$ Mpc and give the results in table 4 and in figure 9. Since the galaxy bias is significantly larger than one, the constraints are tighter for matter distribution than for galaxy distribution. Averaging over redshift and caps we get $3 - \langle \mathcal{D}_2 \rangle_{r,z} = (1.7 \pm 1.0) \times 10^{-4}$ (1σ). This is a strong consistency check of Λ CDM.

Using a similar estimator to Yadav et al. [18] on 1000 QPM mock catalogues, we find a minimum $160 h^{-1}$ Mpc, maximum $1250 h^{-1}$ Mpc, a spread of $120 h^{-1}$ Mpc and an average of $320 h^{-1}$ Mpc corresponding to a precision of about 30% which is less precise than the results, we present in this study. A KS-test [65] for this estimator suggest that data and mocks are drawn from the same distribution at 21% and 83% C.L. for NGC and SGC respectively.

DRAFT

6 Analysis Robustness

6.1 Bias in the fractal correlation dimension

We use homogeneous random catalogues to take into account the inhomogeneity of the survey, so we may wonder if this could bias the resulting $\mathcal{N}(< r)$ towards homogeneity. To search for such a possible bias, we generate 500 fractal realizations with a given value of the fractal correlation dimension, we pass them through our pipeline analysis and study the resulting \mathcal{D}_2 , as done by Laurent et al. [38].

Following Castagnoli and Provenzale [66], we create a cubic box of $L \approx 4 h^{-1}$ Gpc side, containing the whole survey at $z = 0.538 - 0.592$. We divide this box in $M = n^3$ sub-boxes of size L/n , where $n = 2$ and we give to each sub-box a survival probability p . We then repeat the procedure for each surviving sub-box. An infinite number of iterations would give a fractal distribution with:

$$D_2 = \frac{\log(pM)}{\log(n)} \quad (6.1)$$

We perform 9 iterations. After the last iteration, we populate each sub-box with random points that follow a Poisson law of mean $\lambda < 1$. Then we convert the cartesian coordinates to z, RA, DEC with the same FLRW metric as used in our analysis. Finally, we apply cuts to simulate the selection function of our galaxy survey. The value of λ is chosen so that after cuts the number of objects in the fractal distribution is approximately the same as in our survey.

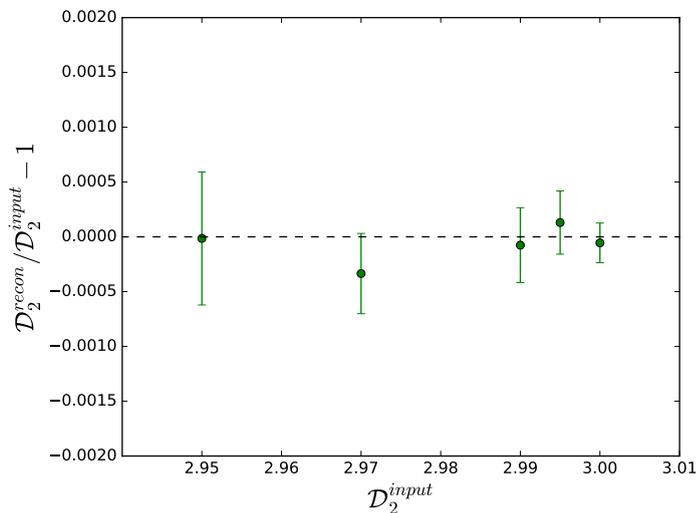


Figure 10. Reconstructed fractal correlation dimension, \mathcal{D}_2 , in the redshift interval $z = 0.538-0.592$, averaged over $r \in [15, 1300] h^{-1}$ Mpc.

At the last iteration the size of the sub-box is $15 h^{-1}$ Mpc. So we average the reconstructed \mathcal{D}_2 over $r \in [15, 1300] h^{-1}$ Mpc. Figure 10 presents this average for different input \mathcal{D}_2 . Fitting these points with a straight line, $y = \alpha x + \beta$, results in $\alpha = 1.00 \pm 0.03$ and $\beta = (-0.4 \pm 7.9) \times 10^{-2}$ with $\chi_{\text{red}}^2 = 0.9/3$. So, in contrast with Scrimgeour et al. [26], we do not observe any bias in the reconstructed fractal correlation dimension.

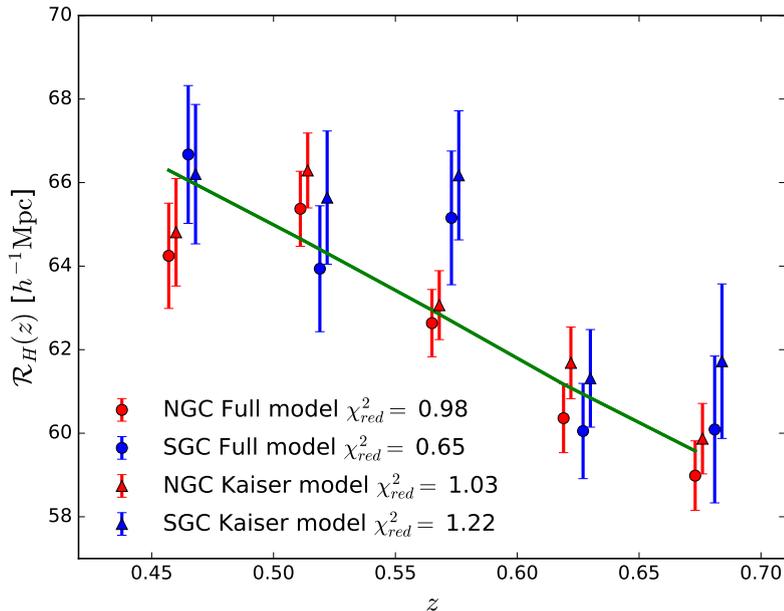


Figure 11. Measured homogeneity scale in different redshift bins, in the NGC and the SGC, with Kaiser (triangle) and full RSD model (circles).

6.2 Sensitivity to RSD model

We compare the homogeneity scale obtained with two different modelling of redshift space distortions. Namely, the model with a scale-independent bias and the Kaiser effect, see equation (4.1) :

$$\xi(r; b) = FFT [P_{gg}(k, \mu; b)] , \quad (6.2)$$

and our full model, which includes in addition a modeling of the finger-of-God effect (Eq. 4.3):

$$\xi(r; b, \sigma_p) = FFT [P_{gg}(k, \mu; b, \sigma_p)] \quad (6.3)$$

Figure 11 presents the homogeneity scale for these two models in the five redshift bins for the north and south galactic caps. We note that using the full RSD model lowers the χ^2 , confirming that this is a better model than the purely linear Kaiser model. The full model tends to lower R_H , by up to slightly more than 1σ , or about 1%. This means that we should not use the purely linear Kaiser model. The remaining error due to the imperfection of our full RSD model is most probably a fraction of the difference between Kaiser and full models, and therefore negligible with respect to our statistical error.

7 Conclusions

We use the data release 12 of BOSS CMASS galaxy sample to study the transition to cosmic homogeneity over a volume of $5.1 h^{-3} \text{ Gpc}^3$. We do not consider the correlation function, $\xi(r)$, to study homogeneity because its definition requires an average density, which is only defined for a homogeneous sample. We rather use the counts-in-spheres, $N(< r)$, i.e. the average number of objects around a given object, and its logarithmic derivative, the fractal correlation dimension, $D_2(r)$. For a homogeneous sample, $N(< r) \propto r^3$ and $D_2 = 3$. We define a characteristic homogeneity scale, \mathcal{R}_H , as the value for which D_2 reaches the homogeneous value within 1%, i.e. $D_2(\mathcal{R}_H) = 2.97$.

For the distribution of galaxies, we get $3 - \langle D_2 \rangle = (0.6 \pm 1.3) \times 10^{-3}$ at 1σ over the range $300 h^{-1} \leq r \leq 1300 h^{-1} \text{ Mpc}$, consistent with homogeneity and a transition to homogeneity at $\mathcal{R}_H = 114.2 \pm 5.8 h^{-1} \text{ Mpc}$. However, our analysis makes use of a random catalog to take into account the geometry and the completeness of the survey. The redshift distribution of this catalogue is taken from the data. We are therefore insensitive to a possible isotropic variation of the density with redshift, $\rho = \rho(z)$. In other words, we can only check for spatial isotropy, $\rho(z, \theta_1) = \rho(z, \theta_2)$. We stress that the same is true for all galaxy redshift surveys [39]. On the other hand, this spatial isotropy can be obtained without using any fiducial model as discussed by Laurent et al. [38]. So if we assume the Copernican principle our data imply homogeneity of the galaxy sample without any Λ CDM assumption.

Using an estimator similar to that of Yadav et al. [18], we find agreement between the mocks and data, further confirming the transition to homogeneity in the matter and galaxy distributions, despite the lower precision of this estimator (30% compared to our estimator which is near %-level). This estimator provides a qualitative estimate of the homogeneity scale in the range of $150 h^{-1} \leq r \leq 400 h^{-1} \text{ Mpc}$.

Alternatively, we can make a cross check of Λ CDM model. Thus, we fit the CMASS galaxy correlation function in the range $1 h^{-1} < r < 40 h^{-1} \text{ Mpc}$ to obtain the galaxy bias relative to the Λ CDM prediction for the matter correlation function. We correct our measurement of $N(< r)$ for this bias in order to get the result for the matter distribution, that we finally compare to the Λ CDM prediction.

For the matter distribution, we get $3 - \langle \mathcal{D}_2(r > 300 h^{-1} \text{ Mpc}) \rangle_z = (1.7 \pm 1.0) \times 10^{-4}$ at 1σ and a transition to homogeneity at a characteristic scale $\mathcal{R}_H = 61.9 \pm 0.8 h^{-1} \text{ Mpc}$ at an average $z = 0.538 - 0.592$. This measurement of \mathcal{R}_H is more precise than previous measurement by Scrimgeour et al. [26] by a factor 5, while Sarkar et al. [27], only give a qualitative measurement of \mathcal{R}_H . We also investigate the redshift evolution of our observables. We find that the homogeneity scale is decreasing with time as expected if clustering is increasing with time. We find accordance with Λ CDM model with a reduced $\chi^2 = 0.89(0.61)$ for the North(South) Galactic Cap for the 6 redshift bins.

Acknowledgements

We would like to thank Jean-Philippe Uzan, Cyril Pitrou and Ruth Durrer for useful suggestions and discussions. PN would like to thank James Rich, although a coauthor, for fruitful discussions on the understanding of this work. PN would like also to thank Mikhail Stolpovskiy and Ranajoy Banerji for useful suggestion on the computational side of this project.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University,

Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

MV is partially supported by Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPITT) No IA102516, Proyecto Conacyt Fronteras No 281 and Proyecto DGTIC SC16-1-S-120.

GR acknowledges support from the National Research Foundation of Korea (NRF) through NRF-SGER 2014055950 funded by the Korean Ministry of Education, Science and Technology (MoEST), and from the faculty research fund of Sejong University in 2016

This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

DRAFT

A From $\xi(r)$ to $\mathcal{N}(< r)$

The probability of finding a galaxy within a volume dV around another galaxy depends on the two-point correlation function $\xi(\vec{r})$ [67]:

$$dP = \bar{\rho} [1 + \xi(\vec{r})] dV . \quad (\text{A.1})$$

The counts-in-spheres of the distribution of galaxies is then related to the correlation function:

$$N(< r) = \int dP = \bar{\rho} \int [1 + \xi(\vec{r}')] dV . \quad (\text{A.2})$$

Assuming $\xi(\vec{r}) = \xi(r)$ we get:

$$N(< r) = 4\pi\bar{\rho} \int_0^r [1 + \xi(r')] r'^2 dr' . \quad (\text{A.3})$$

For the random homogeneous distribution, $N_R(< r) = \bar{\rho} \frac{4\pi}{3} r^3$, so

$$\mathcal{N}(< r) = \frac{N(< r)}{N_R(< r)} = \frac{3}{r^3} \int_0^r [1 + \xi(r')] r'^2 dr' = 1 + \frac{3}{r^3} \int_0^r \xi(r') r'^2 dr' . \quad (\text{A.4})$$

B Choice of estimator for $\mathcal{N}(< r)$

In section 3.2, we consider two estimators for $\mathcal{N}(< r)$, defined by equations (3.5) and (3.6). We compute $\mathcal{N}(< r)$ and the resulting $\mathcal{D}_2(r)$ with the two estimators for the 1000 QPM mock catalogues. Figure 12 compares the mean of the 1000 mocks to the Λ CDM model. The result with the cor estimator are much closer to the Λ CDM model.

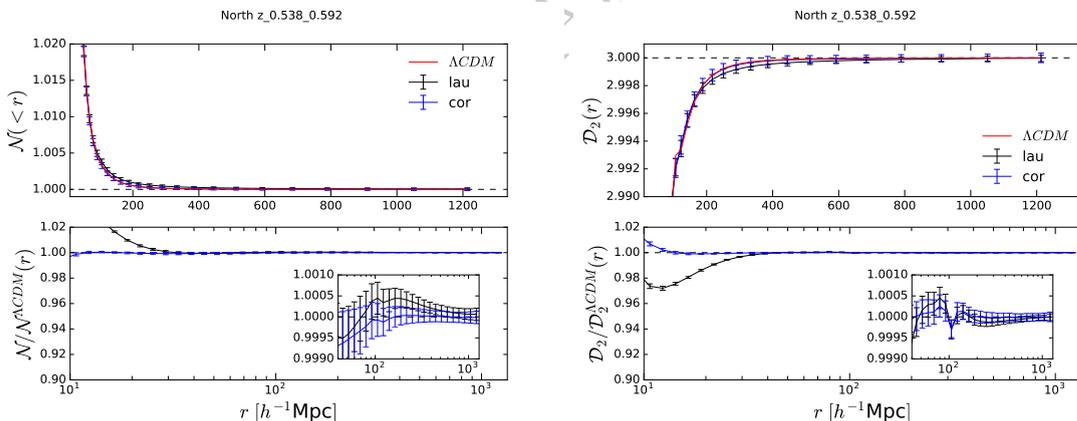


Figure 12. Top: the scaled counts-in-spheres, $\mathcal{N}(< r)$, (left) and the fractal correlation dimension, $\mathcal{D}_2(r)$, (right) for matter distribution, with lau (black) and cor (blue) estimators, compared to Λ CDM model (red). Bottom: the ratio to Λ CDM model for both estimator.

C Tuning errors in RSD Analysis

The theoretical model for redshift space distortions (Eq. 4.3) is not perfectly accurate at the smallest scales due to the real nonlinear behaviour of gravity at these scales. In order to ensure satisfying $\chi^2/\text{n.d.f.}$ for the RSD fitting, we boost the error on $\xi(r)$ at the relevant scales in an empirical way as

$$C'_{ij} = C_{ij} (1 + \delta_{ij} \Delta_i \Delta_j e^2) . \quad (\text{C.1})$$

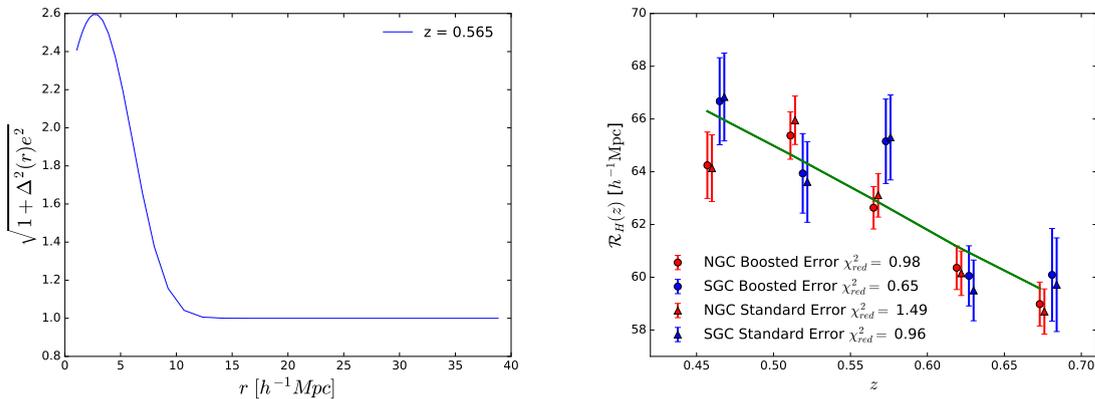


Figure 13. Left: Boosting factor on the error in the correlation function. Right: $R_H^{D_2=2.997}(z)$ with (circle) and without (triangle) boosting the errors, for NGC (red) and SGC (blue).

Here δ_{ij} is the usual Kronecker symbol; e is a parameter that measures the amount of error boosting we apply; Δ_i is the theoretical inaccuracy, estimated as the relative difference between our model and the average of the 1000 QPM mock catalogues. Figure 13 (left) presents the resulting boosting factor on the error in the correlation function. It appears to be significant only on scales smaller than $10 h^{-1}$ Mpc. Fig. 13 (right) shows that the reconstructed homogeneity scales measured with \mathcal{D}_2 is not significantly modified by the error boosting.

D Test of spline fit on QPM mock catalogues

We perform a spline fit of \mathcal{D}_2 for the 1000 QPM mock catalogues in order to obtain the homogeneity scale at 1% (Eq. 3.7). The fit is performed in the range $r \in [40, 100] h^{-1}$ Mpc with 6 data points and 1 degree of freedom. The distribution of the χ^2 of the mock should therefore follow a χ^2 -law for 1 degree of freedom. In table (5) we show the mean and the error on the mean of the distribution of the corresponding χ^2 for the 1000 QPM mock catalogues. The test is successful in both NGC and SGC.

z	χ_{NGC}^2	χ_{SGC}^2
0.430-0.484	1.00 ± 0.05	0.99 ± 0.04
0.484-0.538	0.99 ± 0.05	1.00 ± 0.04
0.538-0.592	1.02 ± 0.05	1.00 ± 0.04
0.592-0.646	0.99 ± 0.04	1.00 ± 0.05
0.646-0.700	1.02 ± 0.05	1.00 ± 0.04

Table 5. Mean and error over the 1000 QPM mock catalogues for the χ^2 of the spline fit with 1 degree of freedom, in the NGC and SGC for the five redshift bins.

E Homogeneity scale at 0.1%

The choice of a 1% threshold to define the homogeneity scales is arbitrary. We can define them for instance at 0.1% as:

$$\mathcal{D}_2(R_H^{D_2=2.997}) = 2.997 \quad \text{or} \quad \mathcal{N}(R_H^{\mathcal{N}=1.001}) = 1.001 \quad (\text{E.1})$$

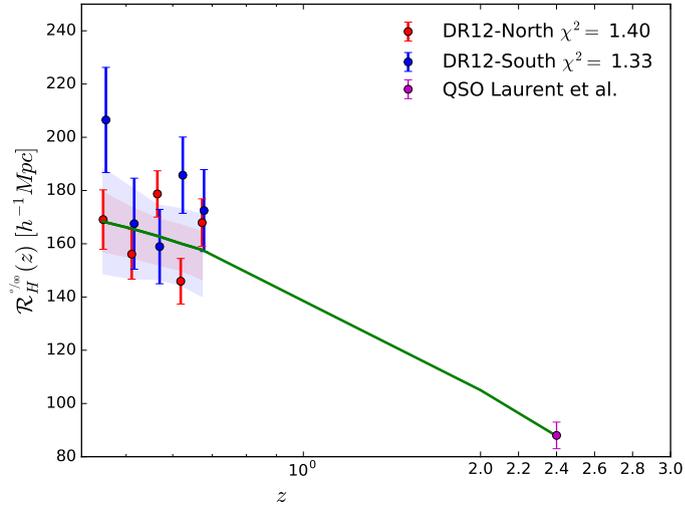


Figure 14. The homogeneity scale at 0.1% level, $R_H^{\mathcal{D}_2=2.997}(z)$, measured in the NGC (red) and in the SGC (blue) as a function of redshift. The purple point is the result obtained with quasars in the NGC, in the redshift range $2.2 \leq z \leq 2.8$ by Laurent et al. [38]. The green line is the Λ CDM model prediction. The shaded areas indicate the 1σ range for the 1000 QPM mock catalogues.

Figure 14 shows that the measured homogeneity scale for matter distribution is compatible with Λ CDM, with $\chi_{red}^2 = 5.82/6$ in the NGC and $\chi_{red}^2 = 7.98/6$ in the SGC.

References

- [1] Perlmutter, S., G. Aldering, G. Goldhaber, et al. Measurements of ω and λ from 42 high-redshift supernovae. *The Astrophysical Journal*, 517:565, 1999.
- [2] Riess, A. G., A. V. Filippenko, P. Challis, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116:1009, 1998.
- [3] Adam, R., P. Ade, N. Aghanim, et al. Planck 2015 results. i. overview of products and scientific results. 2015. [arXiv:1502.01582](https://arxiv.org/abs/1502.01582).
- [4] Percival, W. et al. Parameter constraints for flat cosmologies from cmb and 2dfgrs power spectra, 2002. *Mon. Not. R. Astron. Soc.*, 337:1068. [arXiv:astro-ph/0206256](https://arxiv.org/abs/astro-ph/0206256).
- [5] Parkinson, D., S. Riemer-Sørensen, C. Blake, et al. The wigglez dark energy survey: final data release and cosmological results. *Physical Review D*, 86:103518, 2012. [arXiv:1210.2130](https://arxiv.org/abs/1210.2130).
- [6] Heymans, C., E. Grocutt, A. Heavens, et al. Cfhtlens tomographic weak lensing cosmological parameter constraints: Mitigating the impact of intrinsic galaxy alignments. *Monthly Notices of the Royal Astronomical Society*, 432:2433–2453, 2013. [arXiv:1303.1808](https://arxiv.org/abs/1303.1808).
- [7] Ellis, G. F. Cosmology and verifiability. *Quarterly Journal of the Royal Astronomical Society*, 16:245–264, 1975.
- [8] Fixsen, D., E. Cheng, J. Gales, et al. The Cosmic Microwave Background Spectrum from the Full COBE FIRAS Data Set. *The Astrophysical Journal*, 473:576, 1996. [astro-ph/9605054](https://arxiv.org/abs/astro-ph/9605054).
- [9] Schärf, C., M. Treyer, K. Jahoda, et al. The 2-10 keV xrb dipole and its cosmological implications. 1999. [arXiv:astro-ph/9908187](https://arxiv.org/abs/astro-ph/9908187).
- [10] Blake, C. and J. Wall. Detection of the velocity dipole in the radio galaxies of the nrao vla sky survey. 2002. [astro-ph/0203385](https://arxiv.org/abs/astro-ph/0203385).
- [11] Hamilton, J.-C. What have we learned from observational cosmology? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 46:70–85, 2014.
- [12] Maartens, R. Is the universe homogeneous? *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 369:5115–5137, 2011. [arXiv:1104.1300](https://arxiv.org/abs/1104.1300).
- [13] Clarkson, C. Establishing homogeneity of the universe in the shadow of dark energy. *Comptes Rendus Physique*, 13:682–718, 2012.
- [14] Straumann, N. Minimal assumptions leading to a robertson-walker model of the universe. *Helvetica Physica Acta*, 47:379–383, 1974.
- [15] Durrer, R., J.-P. Eckmann, F. S. Labini, et al. Angular projections of fractal sets. *EPL (Europhysics Letters)*, 40:491, 1997.
- [16] Hogg, D. W., D. J. Eisenstein, M. R. Blanton, et al. Cosmic homogeneity demonstrated with luminous red galaxies. *The Astrophysical Journal*, 624:54, 2005. [arXiv:astro-ph/0411197v1](https://arxiv.org/abs/astro-ph/0411197v1).
- [17] Yadav, J., S. Bharadwaj, B. Pandey, et al. Testing homogeneity on large scales in the sloan digital sky survey data release one. *Monthly Notices of the Royal Astronomical Society*, 364: 601–606, 2005. [arXiv:astro-ph/0504315](https://arxiv.org/abs/astro-ph/0504315).

- [18] Yadav, J. K., J. Bagla, and N. Khandai. Fractal dimension as a measure of the scale of homogeneity. *Monthly Notices of the Royal Astronomical Society*, 405:2009–2015, 2010. [arXiv:1001.0617](#).
- [19] Sarkar, P., J. Yadav, B. Pandey, et al. The scale of homogeneity of the galaxy distribution in sdss dr6. *Monthly Notices of the Royal Astronomical Society: Letters*, 399:L128–L131, 2009. [arXiv:0906.3431](#).
- [20] Martinez, V. J. and P. Coles. Correlations and scaling in the qdot redshift survey. *The Astrophysical Journal*, 437:550–555, 1994.
- [21] Guzzo, L. Is the universe homogeneous?(on large scales). *New Astronomy*, 2:517–532, 1997. [arXiv:astro-ph/9711206](#).
- [22] Martínez, V. J., M.-J. Pons-Borderia, R. A. Moyeed, et al. Searching for the scale of homogeneity. *Monthly Notices of the Royal Astronomical Society*, 298:1212–1222, 1998. [arXiv:astro-ph/9804073](#).
- [23] Scaramella, R., L. Guzzo, G. Zamorani, et al. The eso slice project [esp] galaxy redshift survey: V. evidence for a $d=3$ sample dimensionality. 1998. [arXiv:astro-ph/9803022](#).
- [24] Amendola, L. and E. Palladino. The scale of homogeneity in the las campanas redshift survey. *The Astrophysical Journal Letters*, 514:L1, 1999. [arXiv:astro-ph/9901420](#).
- [25] Pan, J. and P. Coles. Large-scale cosmic homogeneity from a multifractal analysis of the pscz catalogue. *Monthly Notices of the Royal Astronomical Society*, 318:L51–L54, 2000. [arXiv:astro-ph/0008240v1](#).
- [26] Scrimgeour, M. I., T. Davis, C. Blake, et al. The wigglez dark energy survey: the transition to large-scale cosmic homogeneity. *Monthly Notices of the Royal Astronomical Society*, 425:116–134, 2012. [arXiv:1205.6812](#).
- [27] Sarkar, P., S. Majumdar, B. Pandey, et al. The many scales to cosmic homogeneity: Use of multiple tracers from the sdss. 2016. [arXiv:1611.07915](#).
- [28] Appleby, S. and A. Shafieloo. Testing isotropy in the local universe. *Journal of Cosmology and Astroparticle Physics*, 2014:070, 2014. [arXiv:1405.4595](#).
- [29] Labini, F. S., M. Montuori, and L. Pietronero. Comment on the paper by l. guzzo" is the universe homogeneous?". 1998. [arXiv:astro-ph/9801151](#).
- [30] Coleman, P. H. and L. Pietronero. The fractal structure of the universe. *Physics Reports*, 213:311–389, 1992.
- [31] Labini, F. S., M. Montuori, and L. Pietronero. Scale-invariance of galaxy clustering. *Physics Reports*, 293:61–226, 1998. [arXiv:astro-ph/9711073](#).
- [32] Joyce, M., M. Montuori, and F. S. Labini. Fractal correlations in the cfa2-south redshift survey. *The Astrophysical Journal Letters*, 514:L5, 1999. [arXiv:astro-ph/9901290](#).
- [33] Labini, F. S. Breaking the self-averaging properties of spatial galaxy fluctuations in the sloan digital sky survey–data release six. *Astronomy & Astrophysics*, 508:17–43, 2009. [arXiv:astro-ph/9901290](#).
- [34] Labini, F. S. Very large-scale correlations in the galaxy distribution. *EPL (Europhysics Letters)*, 96:59001, 2011. [arXiv:1110.4041](#).
- [35] Hoyle, B., R. Tojeiro, R. Jimenez, et al. Testing homogeneity with galaxy star formation histories. *The Astrophysical Journal Letters*, 762:L9, 2012.

- [36] Zibin, J., A. Moss, et al. Nowhere to hide: closing in on cosmological homogeneity. 2014. [arXiv:1409.3831](#).
- [37] Wu, K. K., O. Lahav, and M. J. Rees. The large-scale smoothness of the universe. *Nature*, 397: 225–230, 1999.
- [38] Laurent, P., J.-M. Le Goff, E. Burtin, et al. A $14 h^{-3} \text{ Gpc}^3$ study of cosmic homogeneity using BOSS DR12 quasar sample. 2016. [arXiv:1602.09010](#).
- [39] Mustapha, N., C. Hellaby, and G. Ellis. Large-scale inhomogeneity versus source evolution: can we distinguish them observationally? *Monthly Notices of the Royal Astronomical Society*, 292: 817–830, 1997.
- [40] Planck Collaboration and Ade, P. A. R. and Aghanim, N. and Arnaud, M. and Ashdown, M. and Aumont, J. and Baccigalupi, C. and Banday, A. J. and Barreiro, R. B. and Bartlett, J. G. and et al., t. . P. 2016. [arXiv:1502.01589](#).
- [41] Dawson, K. S., D. J. Schlegel, C. P. Ahn, et al. The baryon oscillation spectroscopic survey of sdss-iii. *The Astronomical Journal*, 145:10, 2012.
- [42] Alam, S., F. D. Albareti, C. A. Prieto, et al. The eleventh and twelfth data releases of the sloan digital sky survey: final data from sdss-iii. *The Astrophysical Journal Supplement Series*, 219: 12, 2015. [arXiv:1501.00963v3](#).
- [43] Eisenstein, D. J., D. H. Weinberg, E. Agol, et al. Sdss-iii: Massive spectroscopic surveys of the distant universe, the milky way, and extra-solar planetary systems. *The Astronomical Journal*, 142:72, 2011.
- [44] Gunn, J. E., W. A. Siegmund, E. J. Mannery, et al. The 2.5 m telescope of the sloan digital sky survey. *The Astronomical Journal*, 131:2332, 2006.
- [45] Smee, S. A., J. E. Gunn, A. Uomoto, et al. The multi-object, fiber-fed spectrographs for the sloan digital sky survey and the baryon oscillation spectroscopic survey. *The Astronomical Journal*, 146:32, 2013. [arXiv:1208.2233](#).
- [46] Bolton, A. S., D. J. Schlegel, É. Aubourg, et al. Spectral classification and redshift measurement for the sdss-iii baryon oscillation spectroscopic survey. *The Astronomical Journal*, 144:144, 2012. [arXiv:1207.7326](#).
- [47] Fukugita, M., T. Ichikawa, J. Gunn, et al. The sloan digital sky survey photometric system. *The Astronomical Journal*, 111:1748, 1996.
- [48] Maraston, C., G. Strömbäck, D. Thomas, et al. Modelling the colour evolution of luminous red galaxies—improvements with empirical stellar spectra. *Monthly Notices of the Royal Astronomical Society: Letters*, 394:L107–L111, 2009.
- [49] Aubourg, É., S. Bailey, F. Beutler, et al. Cosmological implications of baryon acoustic oscillation measurements. *Physical Review D*, 92:123516, 2015.
- [50] Reid, B., S. Ho, N. Padmanabhan, et al. Sdss-iii baryon oscillation spectroscopic survey data release 12: galaxy target selection and large-scale structure catalogues. *Monthly Notices of the Royal Astronomical Society*, 455:1553–1573, 2016.
- [51] Feldman, H. A., N. Kaiser, and J. A. Peacock. Power spectrum analysis of three-dimensional redshift surveys. *arXiv preprint astro-ph/9304022*, 1993.
- [52] White, M., J. L. Tinker, and C. K. McBride. Mock galaxy catalogues using the quick particle mesh method. *Monthly Notices of the Royal Astronomical Society*, 437:2594–2606, 2014.

- [53] Peebles, P. J. E. and M. G. Hauser. Statistical analysis of catalogs of extragalactic objects. iii. the shane-wirtanen and zwicky catalogs. *The Astrophysical Journal Supplement Series*, 28:19, 1974.
- [54] Landy, S. D. and A. S. Szalay. Bias and variance of angular correlation functions. *The Astrophysical Journal*, 412:64–71, 1993.
- [55] Blas, D., J. Lesgourgues, and T. Tram. The cosmic linear anisotropy solving system (class). part ii: approximation schemes. *Journal of Cosmology and Astroparticle Physics*, 2011:034, 2011.
- [56] Kaiser, N. Clustering in real space and in redshift space. *Monthly Notices of the Royal Astronomical Society*, 227:1–21, 1987.
- [57] Linder, E. V. and R. N. Cahn. Parameterized beyond-einstein growth. *Astroparticle Physics*, 28: 481–488, 2007. [arXiv:astro-ph/0701317v2](#).
- [58] Ballinger, W., J. Peacock, and A. Heavens. Measuring the cosmological constant with redshift surveys. *arXiv preprint astro-ph/9605017*, 1996. [arXiv:astro-ph/9605017v1](#).
- [59] White, M., M. Blanton, A. Bolton, et al. The clustering of massive galaxies at $z \approx 0.5$ from the first semester of boss data. *The Astrophysical Journal*, 728:126, 2011. [arXiv:1010.4915v2](#).
- [60] Nuza, S. E., A. G. Sánchez, F. Prada, et al. The clustering of galaxies at $z \approx 0.5$ in the sdss-iii data release 9 boss-cmass sample: a test for the Λ cdm cosmology. *Monthly Notices of the Royal Astronomical Society*, 432:743–760, 2013. [arXiv:1202.6057](#).
- [61] Ho, S., A. Cuesta, H.-J. Seo, et al. Clustering of sloan digital sky survey iii photometric luminous galaxies: the measurement, systematics, and cosmological implications. *The Astrophysical Journal*, 761:14, 2012.
- [62] Comparat, J., E. Jullo, J.-P. Kneib, et al. Stochastic bias of colour-selected bao tracers by joint clustering–weak lensing analysis. *Monthly Notices of the Royal Astronomical Society*, 433: 1146–1160, 2013. [arXiv:1302.5655v2](#).
- [63] Gil-Marín, H., J. Noreña, L. Verde, et al. The power spectrum and bispectrum of sdss dr11 boss galaxies–i. bias and gravity. *Monthly Notices of the Royal Astronomical Society*, 451:5058–5099, 2015.
- [64] Thomas, D., O. Steele, C. Maraston, et al. Stellar velocity dispersions and emission line properties of sdss-iii/boss galaxies. *Monthly Notices of the Royal Astronomical Society*, page stt261, 2013. [arXiv:1207.6115v2](#).
- [65] Kolmogorov–Smirnov test. Encyclopedia of Mathematics.
- [66] Castagnoli, C. and A. Provenzale. From small-scale fractality to large-scale homogeneity—a family of cascading models for the distribution of galaxies. *Astronomy and Astrophysics*, 246:634–643, 1991.
- [67] Peebles, P. The large-scale structure of the universe (research supported by the national science foundation. princeton, nj, princeton university press, 435 p., 1980). In *ADS*, volume 25, page 82, 1980.

Chapter 5

Cosmic Homogeneity as a Standard Ruler

In this chapter, we discuss how to use the characteristic scale of transition to homogeneity as a standard ruler to measure the properties of the Universe, i.e. perform cosmography with it. Furthermore, we estimate forecasts of this ruler for future survey instruments. This work was done in collaboration with Nicolas Busca, J.M. Le Goff.

5.1 Does the \mathcal{R}_H bring any new extra cosmological information?

To answer this question, we first need to know if the measurement of the homogeneity scale is independent of the BAO peak position, since both are obtained from the same data. To know this we need to perform the following simple test, i.e. describe the correlation of the measurement homogeneity scale with that of the BAO scale. In order to perform a blind analysis test, we measure the Homogeneity scale and the BAO peak position, each with two different methods. First, we describe the methods used to measure the two scales and then we present the result. If we are convinced that the homogeneity scale is independent of the measurement of the BAO scale then we proceed to characterise the cosmological information the homogeneity scale provides. Then we compare this information to the one from the BAO scale. We do that in two steps. Firstly, we provide a fisher matrix analysis on the two scales for a set of cosmological parameters. This will show us which parameters the homogeneity scale is sensitive to. We then deploy an MCMC parameter estimation algorithm to explore the cosmological parameter space that the homogeneity scale is sensitive to, and compare it to the BAO scale cosmological parameter sensitivity. Furthermore, in this chapter we provide an estimate of the sensitivity of the homogeneity scale from future surveys. In order to do so, we will simulate ground based and satellite photometric and spectroscopic surveys, that is, their density, volume and photometric redshift error.

5.1.1 Determination of the BAO peak Position

In order to determine the BAO peak position, we first use the Landy & Szalay estimator to measure the 2-point correlation function from the galaxy sample data, as discussed in section 4.1.4.1. Then in the range of $r = [40, 180]h^{-1}\text{Mpc}$, we fit the 2 point correlation function with two different methods. In both methods, in order to avoid the result from being driven by the full shape of the 2pt correlation function, which is subjected to systematics, we model the 2-point correlation function in such a way that the result is only driven by the location of the BAO peak position[186]. We then describe the application to both methods of a broadband model. This model can be described by the following formula:

$$bb(r; p_{bb}) = p_1 + \frac{p_2}{r} + \frac{p_3}{r^3} \quad (5.1)$$

where $p_{bb} = (p_1, p_2, p_3)$ are the broadband parameters.

For the first method, we model the measurement of the BAO peak position using the usual Gaussian model [187], described by:

$$\xi^{(1)}(r; \mathcal{R}_{BAO}, A, \sigma_{peak}, p_{bb}) = A \exp\left[-\frac{1}{2}\left(\frac{r - \mathcal{R}_{BAO}}{\sigma_{peak}}\right)^2\right] + bb(r; p_{bb}) \quad (5.2)$$

where \mathcal{R}_{BAO} is the BAO scale parameter, A and σ_{peak} are the amplitude and the smoothing scale of a Gaussian function, respectively.

In the second method, we use the usual correlation function described in section 4.2.2 with the fiducial cosmology as a template. We then model the correlation function as:

$$\xi^{(2)}(r; \alpha_{iso}, p_{bb}) = b^2 \xi(\alpha_{iso} * r; p_{cosmo}^{fid}) + bb(r; p_{bb}) \quad (5.3)$$

where we model the *isotropic dilatation* parameter as:

$$\alpha_{iso} = r_s / r_s^{fid} \quad (5.4)$$

In order to determine the BAO peak position, we measure the parameter $r_s = \mathcal{R}_{BAO}$ by marginalising on the rest of the parameters. For the first set of parameters we marginalise the nuisance parameters defined by:

$$p_n^{(1)} = (A, \sigma_{peak}, p_{bb}) \quad (5.5)$$

. For the second method we marginalise the nuisance parameters defined by:

$$p_n^{(2)} = (b, p_{bb}) . \quad (5.6)$$

Therefore, in both methods, we take into consideration uncertainties due to *redshift space distortions* introduced in our measurement.

5.1.2 Determination of the Homogeneity Scale

To determine the homogeneity scale we use two different fitting methods. In both methods we use the estimator that is described by Eq. 4.7. For the first method, we use the spline-interpolation method, i.e. we fit a spline function on the fractal correlation dimension observable \mathcal{D}_2 of the galaxy distribution at ranges $r = [90 - 200]h^{-1}\text{Mpc}$ as described in chapter 4.1.6. This gives us the homogeneity scale for the first method, namely $\mathcal{R}_H^{(1)}$. For the second method, we use the same algorithm but we substitute

the spline-interpolation with a polynomial interpolation at ranges $r = [10 - 1300]h^{-1}\text{Mpc}$. This gives us the homogeneity scale from the second method, namely $\mathcal{R}_H^{(2)}$. We use the *cor*-estimator for the Fractal Correlation dimension table 4.2 for both cases.

5.1.3 Blind Cross-Correlation of $\mathcal{R}_H - \mathcal{R}_{BAO}$ plane

Using 1000 QPM mock catalogues, we determine the $\mathcal{R}_{BAO}^{(1)}$ and $\mathcal{R}_{BAO}^{(2)}$ as determined by the method using the equation 5.2 and Eq. 5.3 respectively. In Fig. 5.1, we plot the Homogeneity scale versus the BAO peak position both determined via their first method as described in section 5.1.2. The contour shows that there is a 19% correlation between the measurement of the two scales at redshift slice $0.538 \leq z < 0.592$ on the North Galactic Cap. Notice that both measurements are done on the galaxy distribution rather than on the total matter distribution. This is in order not to measure the bias before the parameter estimation with all the cosmological parameters. We do this in order to simplify the measurement of the homogeneity scale by not requiring a bias measurement beforehand. This renders the measurement less model dependent and furthermore, we can investigate any possible correlation of the bias parameter with the cosmological parameters.

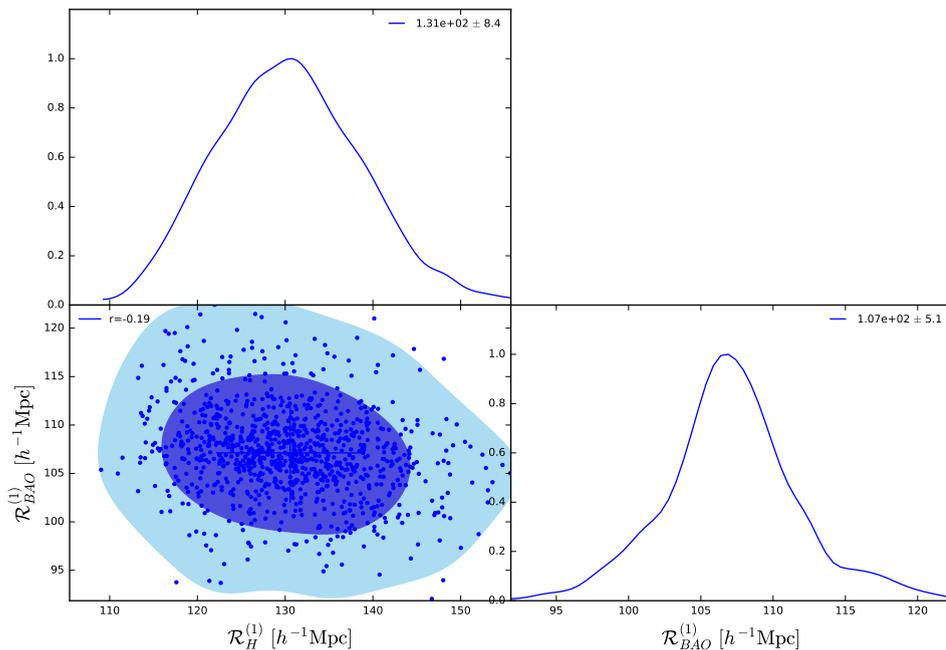


FIGURE 5.1: Contour plot with $1\sigma(2\sigma)$ in dark blue (light blue) for the Homogeneity scale versus the BAO scale for the $0.538 \leq z < 0.592$. See text for details.

Furthermore, as a blind test, we perform the measurement using the other two methods. We cross-correlate the results of all different combinations in order to calculate the correlation coefficient, r . In Fig. 5.2, we present the results. The top panel of Fig. 5.2 shows that the absolute value of the correlation coefficient for the different methods is less than 30 for all redshift slices. Moreover, in the bottom panel of Fig. 5.2, we show the non z -dependence of the measurement of the BAO peak position and the Homogeneity scale. This means that the correlation between the two scale measurements, the homogeneity scale measurement and the BAO measurement is negligible.

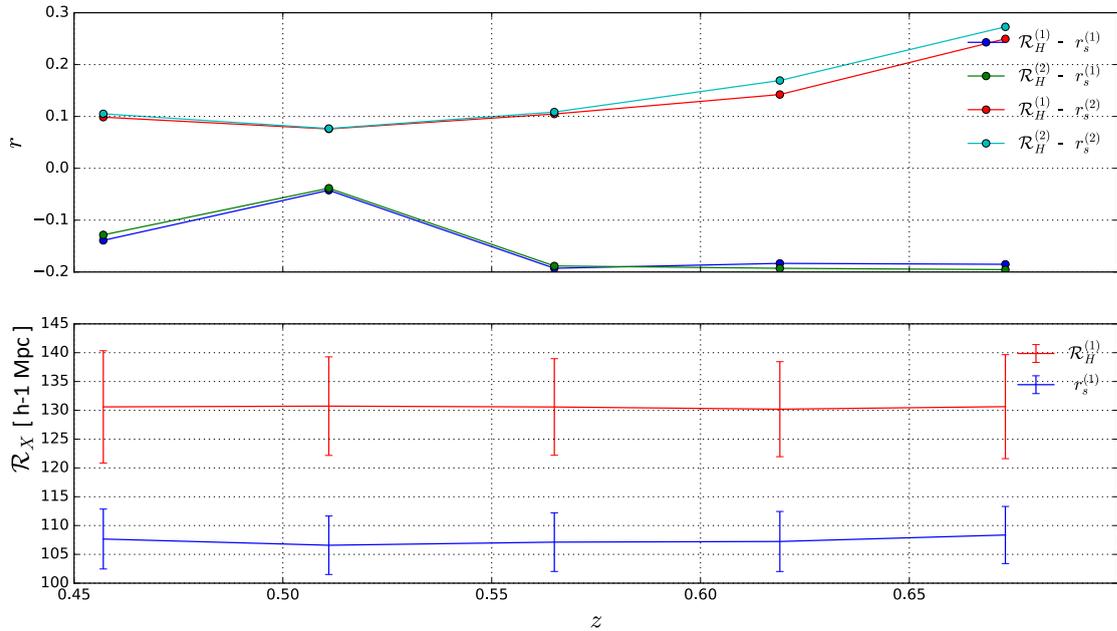


FIGURE 5.2: *Upper*: Correlation coefficient (r) as a function of z for methods measuring \mathcal{R}_H and r_s . *Bottom*: \mathcal{R}_H and r_s measurement as a function of z for 1000 QPM mock catalogues. See text for explanation.

Additionally, in table 5.1 we present the weighted average between the measurements for each scale using the two different method. The weighted average of BAO peak is consistent with measurements found in [188].

Galaxies	\mathcal{R}_H [h ⁻¹ Mpc]	r_s [h ⁻¹ Mpc]
<i>Method 1</i>	130.5 ± 3.9	107.4 ± 2.3
<i>Method 2</i>	130.9 ± 3.7	108.0 ± 3.6

TABLE 5.1: Weighted average over 5 redshift bins for the Homogeneity and BAO scale for the two different methods.

The aforementioned results show that the measurements are blind to the methods used to determine both estimators for both scales (\mathcal{R}_H and R_{BAO}). Furthermore, it is demonstrated that there is correlation less than 30% between the measurements of the two scales for the galaxy distribution for either of the redshift slices that we use in our analysis. Therefore, the homogeneity scale measurement, \mathcal{R}_H , brings information additional to that of BAO scale measurement, on the measured cosmological parameters. Since we are convinced that the homogeneity scale, as a ruler, gives extra independent information from previous rulers (such as the BAO scale), we can now quantify that amount of information.

5.2 Fisher Matrix \mathcal{R}_H vs r_s

The Fisher matrix formalism plays a fundamental role in predicting the errors obtained from a given experimental set up. Therefore it is used extensively on most experimental designs [189]. The fisher matrix is defined as

$$F_{ij} = \sum_b \frac{1}{\sigma_b^2(b)} \frac{\partial o(b)}{\partial p_i} \frac{\partial o(b)}{\partial p_j} \quad (5.7)$$

where $o(b)$ is the observable in the bin, b , of a measurement of an experimental setup, $\sigma_o(b)$ is the error of the observable and p_i are the parameters upon which the estimator depends. We use as a fiducial model the $w_0w_a\Lambda\text{CDM}$ -model, i.e. we assume for a time dependent equation of state¹. This means that the evolution of the Dark Energy fluid is a more complicated function than $\Omega_\Lambda \cdot (1+z)^{3(1+w)}$. The next order approximation is the redshift dependence of equation of state as $w(z) = w_0 + \frac{z}{1+z}w_a$, where w_0 is the equation of state parameter today and w_a is the difference between the equation of state parameter at high redshift, $z \gg 1$, and the one today. We are overall interested in the following parameters, namely:

$$\mathbf{p}_{w_0w_a\Lambda\text{CDM}} = \{h, n_s, \omega_b, \omega_{\text{cdm}}, n_s, \ln 10^{10} A_s, \Omega_\Lambda, w_0, w_a\} \quad (5.8)$$

to search for a possible impact using the homogeneity scale measurement as a probe. Note that we have assumed that $\sigma_o(b; \mathbf{p}) = \sigma_o(b)$, which means that the errors of the observable do not depend on the variation of the parameters. It has been shown by Fisher[190] that the inverse of the Fisher matrix is the covariance matrix with the minimum variance that one can obtain for this observable.²

From table 4.4 within the redshift bin $0.538 \leq z \leq 0.592$ on the North Galactic Cap, we can obtain an estimate of the precision on the homogeneity scale of the matter distribution, as $0.8/62.6 = 1.3\%$, or the galaxy distribution, as $7.8/116.4 = 6.7\%$. In table 5.1, we have shown that the precision of the Baryon Acoustic Oscillation scale for the galaxy distribution is $5.1/107 = 4.8\%$. We have shown in section A.8 that the precision of the BAO scale does not change for the matter distribution. Therefore, we adopt these precisions for each observable. We then apply Eq. 5.7 for the theoretical prediction of the homogeneity scale, $\mathcal{R}_H(z; \mathbf{p})$ and for the comoving size of the BAO,

$$\mathcal{R}_{\text{BAO}}(z; \mathbf{p}) = r_s(z; \mathbf{p})/a(z_*), \quad (5.9)$$

where $a(z_*)$ is the scale factor at the recombination epoch, z_* . We then use the property of the Fisher

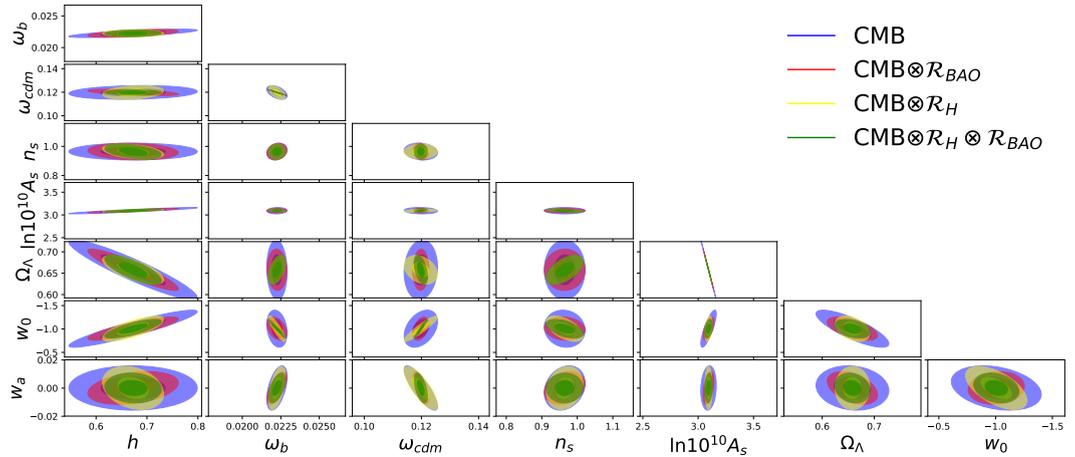


FIGURE 5.3: Contours of 68% (shaded) and 95% (light) C.L. of the CMB observable (blue), CMB combined with the BAO scale observable (red), CMB combined with the Homogeneity scale observable (blue) and CMB combined with the Homogeneity scale observable and the BAO scale observable (green) for the 7 cosmological parameters of $w_0w_a\Lambda\text{CDM}$ -model.

matrix to add Fisher matrices of different data sets as explained in appendix A.11. This property helps in two ways. First, we add the information for each observable, $\mathcal{R}_H(z; \mathbf{p})$ or $\mathcal{R}_{\text{BAO}}(z; \mathbf{p})$ for the different

¹Chevallier-Polarski-Linder parametrization[6]

² A simple derivation is given in appendix A.11.

redshift bins described in table 3.1, i.e.

$$F_{ij}^X = \sum_z F_{ij}^X(z)$$

where X is the $\mathcal{R}_H(\mathbf{p}_{w_0wa\Lambda CDM})$ or $\mathcal{R}_{BAO}(\mathbf{p}_{w_0wa\Lambda CDM})$ and (i, j) are indices of the matrix that run through our cosmological parameters $\mathbf{p}_{w_0wa\Lambda CDM}$. Then, using the information from the CMB measurements from Planck satellite on the estimation on the error of the parameters[6], we quantify the Fisher Matrix of CMB³. To quantify the fisher matrix of the CMB, we simply invert the covariance matrix of the parameters we are interested in. We then add the Fisher Matrix of each observable with the Fisher Matrix of the CMB.

The result for the matter homogeneity comoving scale and the BAO acoustic comoving scale, are shown in Fig. 5.3. We observe that the Homogeneity Scale breaks the degeneracies, in particular the degeneracies of the $\omega_{cdm}, \Omega_\Lambda, w_0, h$. We also observe that using both observables (the BAO and the homogeneity scale) we end up further constraining all cosmological parameters. Note that we have assumed the same precision for all redshift bins for simplicity. Notice also as well that the fisher matrix method assumes that the observable has a linear dependence on the parameters that we take into consideration. Therefore, it needs verification of parameter exploration from the data. Therefore, we are going to explore those parameters with the MCMC algorithm on the next chapter.

5.3 MCMC on \mathcal{R}_H

In order to explore the parameter space with the homogeneity scale, we use the following procedure. We measure the homogeneity scale for the galaxy distribution using a fiducial cosmology. We then measure the correlation function for the galaxy distribution as a function of scale at a range $r \in [10 - 40]h^{-1}$ Mpc. We do this step since it is necessary to obtain an estimate of the bias to use for the theoretical prediction of the homogeneity scale for the galaxy distribution, since it is biased from the homogeneity scale of the matter distribution. Therefore, we model our χ^2 with two parts. The first part takes the information of the theoretical prediction of the homogeneity scale with our measurement for the galaxy distribution. The second part estimates the bias information according to the galaxy sample using the 2pt correlation function. So, the first part is formulated as:

$$\chi_{\mathcal{R}_H}^2(z; b, p_T) = \left(\frac{\mathcal{R}_H^G(z; p_F) - \mathcal{R}_H^{G,th}(z; b, p_T) \times \alpha(z; p_F, p_T)}{\sigma_{\mathcal{R}_H^G}(z)} \right)^2 \quad (5.10)$$

where $\mathcal{R}_H^G(z; p_F)$ is the measurement of the homogeneity scale for the galaxy distribution at redshift z and p_F is the fiducial cosmology given by Eq. 3.4. The 1σ deviation of the homogeneity scale is given by $\sigma_{\mathcal{R}_H^G}(z)$ at each redshift slice. The theoretical prediction of the homogeneity scale is given by the term $\mathcal{R}_H^{G,th}(z; b, p_T)$ at redshift z where p_T are the parameters of the "true" cosmology that we are going to explore. Notice that we correct for the fiducial cosmology by the factor:

$$\alpha(z; p_F, p_T) = \frac{d_V(z; p_F)}{d_V(z; p_T)}$$

³We have used the Planck 2013 Data Release[191] since they have estimates of the likelihood of the cosmological parameters only for CMB measurements that we need to use independently. In 2015 Data Release they have only combinations of different datasets including large scale structures that we already use <http://pla.esac.esa.int>.

where $d_V(z; p_T)$ is the volume distance given by Eq. 4.22. We model the second part of the χ^2 via:

$$\chi_\xi^2(z; b, p_T) = \left[\xi^G(r_i; p_F) - \xi_{th}^G(\alpha^{-1} \cdot r_i; b, p_T) \right] C_{r_i r_j}^{-1} \left[\xi^G(r_i; p_F) - \xi_{th}^G(\alpha^{-1} \cdot r_i; b, p_T) \right]^T \quad (5.11)$$

where $\xi^G(r_i; p_F)$ is the measured correlation function at the fiducial cosmology, p_F at ranges $r \in [10 - 40]h^{-1}\text{Mpc}$ as a function of scales r_i , $\xi_{th}^G(\alpha^{-1} \cdot r_i; p_T)$ is the theoretical predictions at the cosmology we explore, i.e. the "true" set of cosmological parameters, p_T , and α^{-1} is the inverse of Eq. 5.3 that we use to correct for the fiducial cosmology. Moreover, $C_{r_i r_j}^{-1}$ is the covariance matrix of the correlation function at the scales we are considering. Note that we have used only the Kaiser model to constrain the bias parameter at small scales since it is the simplest one. Furthermore, the full RSD model has no significant deviation from the Kaiser model on the measurement of the homogeneity scale for the matter distribution (less than 1%) as we have shown in chapter 4.3.6, therefore it is justified to use only the Kaiser model here. We execute the MCMC algorithm[192] at each redshift bin by minimizing for p_T the following quantity:

$$\chi^2(z; b, p_T) = \chi_{\mathcal{R}_H}^2(z; b, p_T) + \chi_\xi^2(z; b, p_T)$$

We are focused in the simplest case exploring the 3 basic cosmological parameters, i.e. we consider the cold dark matter energy density ratio given by $\omega_{cdm} = \Omega_{cdm}h^2$ and the curvature ratio Ω_k of the universe and we fix the rest of the parameters to the fiducial values[6]. The parameters we are exploring are listed in table 5.2. Notice that we can infer the measurement the total matter of the universe, since $\Omega_m = (\omega_{cdm} + \omega_b)/h^2$ for the fixed values of h and ω_b . Furthermore, we can infer the Ω_Λ parameter from the exploration of the Ω_k parameter since $\Omega_\Lambda = 1 - \Omega_k - \Omega_m$. We further explore the bias parameter between luminous and dark matter, b . The last parameter we explore is the dimensionless hubble constant, h .

Parameter	Prior Mean Value	Prior Range	Prior Type
<i>bias</i>	1.94	-0.02 - 0.02	gaussian
ω_{cdm}	0.1198	-0.0015 - 0.0015	gaussian
Ω_k	0.0040	-0.015 - 0.015	gaussian
h	0.6727	-0.0066 - 0.0066	gaussian

TABLE 5.2: Prior measurements of the bias and the cosmological parameters space that we explore with the homogeneity scale of the galaxy distribution \mathcal{R}_H^G .

In Fig. 5.4, we present the result of the Eq. 5.3 with a corner-plot⁴ for the 4 aforementioned parameters, i.e. $b, h, \Omega_M, \Omega_\Lambda$. In blue is the measurement of cosmological parameters using the homogeneity scale without applying the priors nor the weights given by Eq. 3.6. In red we show the measurement with all the priors applied and we do not consider the weights as well. In green we do not use the priors but we apply all the weights on the measurement of the homogeneity scale. Finally in yellow, we use both all the priors and all the weights. These results show that without using the priors and applying or not the weights of our measurement we have no significant deviations from the estimated parameters, i.e. the deviation is less than 3% for all four parameters but within the 1σ deviation of the measurement. We further observe that there is an anti-correlation between the h parameter and the Ω_m parameter, $r \sim 0.9$. The correlation between the bias and the Ω_Λ reduces from 80% to 20% when we use the prior CMB measurement and the bias measurement rendering the correlation insignificant. Furthermore, we observe that when we use the prior information there is a slight correlation between the Ω_Λ and h as well as between Ω_Λ and Ω_m . The result using all the priors from the CMB measurement and the bias parameter shows that the deviation is less than 1% for the b and Ω_m parameters, while for the Ω_Λ is

⁴The corner plot is a matrix of several plots of a set of parameters where usually the 1σ (68%) and 2σ (95%) Confidence Level are displayed with contours. On the diagonal of this matrix the normalised histogram of each parameter is displayed.

2%. Therefore, the measurement of the homogeneity scale is a comparable way to measure cosmological parameters.

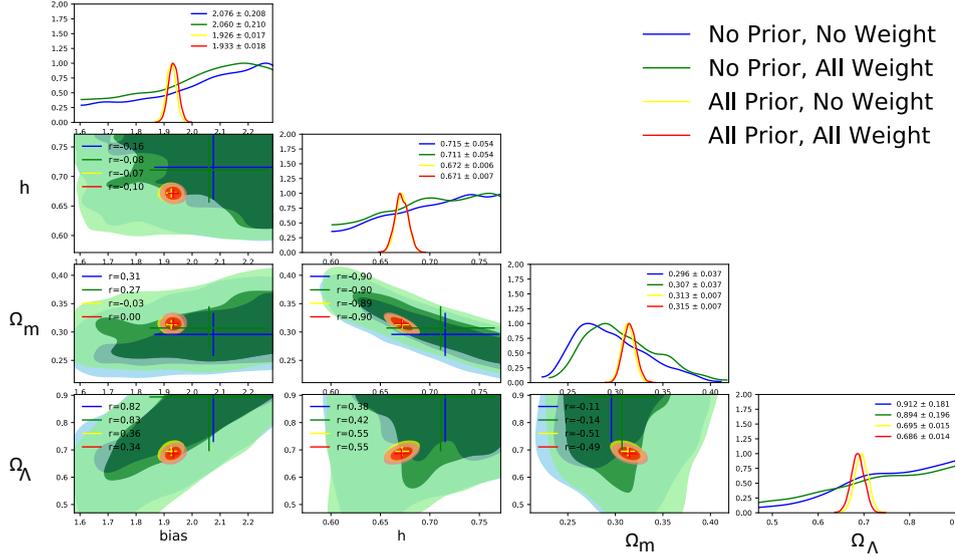


FIGURE 5.4: Contours of 68% (shaded) and 95% (light) C.L. of the homogeneity scale of the galaxy distribution for the 3 cosmological parameters of Λ CDM-model at the intermediate redshift slice of North Galactic Cap. Different colors represent different configurations of the measurement [See text for details.]

In Fig. 5.5 the estimation of 68% and 95% contours for the 3 cosmological parameters is displayed for the 5 redshift bins described by table 3.1 on the North Galactic Cap. The measurement was performed using all the weights and all the priors. We observe that using the priors we reach a 1% precision. The total matter density of the universe today remains consistent with the values obtained at all different epochs within 1σ confidence level. Its value for the middle redshift bin is $\Omega_m(0.538 \leq z \leq 0.592) = 0.315 \pm 0.007$ which is in accordance with the values obtained by Planck Satellite measurement, $\Omega_m^{Planck} = 0.316 \pm 0.009$. This measurement is also consistent with the values obtained by Laurent [193] that they use the two point correlation function method with a combination of CMASS, LOWZ and QSO samples of either BOSS or eBOSS survey.

Furthermore, we observe that the correlation between the individual biases and the dimensionless hubble constant and total matter density ratio are negligible, i.e. the correlation coefficient are $|r| \lesssim 0.16$ and $|r| \lesssim 0.07$ respectively. The correlation of bias and the Dark energy ratio density is insignificant as well, $r < 0.30$. Furthermore, we observe that the dimensionless hubble constant parameter is weakly correlated with Ω_Λ parameter, i.e. $r \sim 0.5$ for the individual redshift bins. We also see that the dimensionless hubble constant parameter is anti-correlated, or degenerate, with the Ω_Λ parameter with $r \sim -0.9$ for the individual redshift bins. Additionally, the total matter density is slightly anti-correlated with the dark energy density, $r \sim -0.50$ for the different redshift bins.

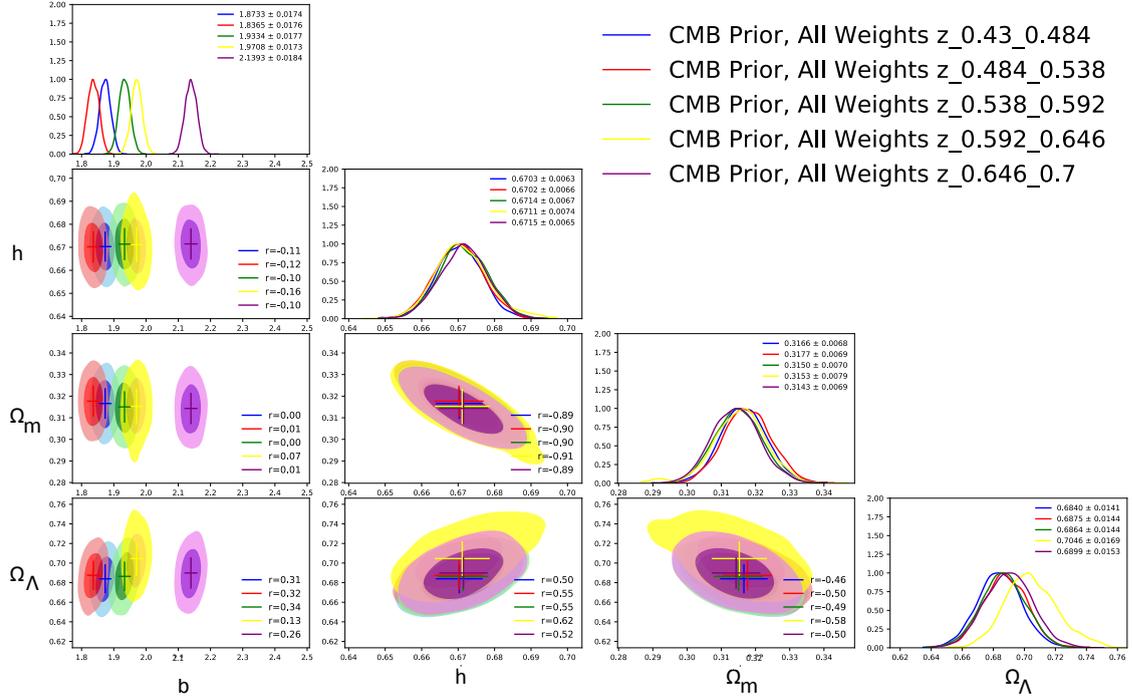


FIGURE 5.5: Contours of 68% (shaded) and 95% (light) C.L. of the homogeneity scale of the galaxy distribution for the 3 cosmological parameters of Λ CDM-model for the 5 redshift slices (color-coded) of the North Galactic Cap. [See text for details.]

Finally, we explore the parameter space of 3 cosmological parameters and the 5 bias parameters for the data for the 5 different redshift bins at once with and without using the priors. This is formally performed by minimizing the sum of the different $\chi^2(z)$ for the different parameters, i.e.:

$$\chi_{once}^2 = \sum_{z=0}^4 [\chi_{\xi}^2(z; b, p_T) + \chi_{\mathcal{R}_H}^2(z; b, p_T)] \quad (5.12)$$

Notice that we apply all the weights, as explained in chapter 3.4.2. Furthermore, we change our CMB measurement to the one of Planck 2013 results, which are more well known.

The result are shown in Fig. 5.6. In order to compare our measurement combined with the gaussian prior of CMB, namely CMB prior, we overplot the CMB prior in blue.

We see that the measurement of the CMB planck 13 of cosmological parameters and the bias measurement as prior knowledges combined with the Homogeneity scale the following. We obtain values on the biases consistent with the ones obtained in chapter 4.2.2 but more precise 19%. Furthermore, we observe a degeneracy between the dimensionless hubble constant with the total matter ratio density, $r \sim -0.95$. While the dimensionless hubble constant is less degenerate with the Dark Energy ratio density, $r \sim 0.79$. We also find no correlation between the 5 bias parameters and the 3 cosmological parameters as expected. We find a slight anticorrelation between the total matter ratio density and the Dark Energy ratio density of about 55%. The final result, using the CMB prior and the bias measurement prior, show that we obtain a value for the hubble constant, $H_0 = 69.4 \pm 0.18 \text{ km/s/Mpc}$ which is consistent but less precise than the one obtained $H_0^{Planck} = 67.7 \pm 0.1 \text{ km/s/Mpc}$. We find for the total matter ratio density to, $\Omega_m = 0.294 \pm 0.017$ which is less precise but comparable than the one reported by Planck $\Omega_m^{Planck} = 0.307 \pm 0.011$. For the Dark energy ratio density, we constrain $\Omega_\Lambda = 0.760 \pm 0.026$ which is in tension with the Planck value, $\Omega_\Lambda = 0.693 \pm 0.012$. Furthermore, we constrain using the

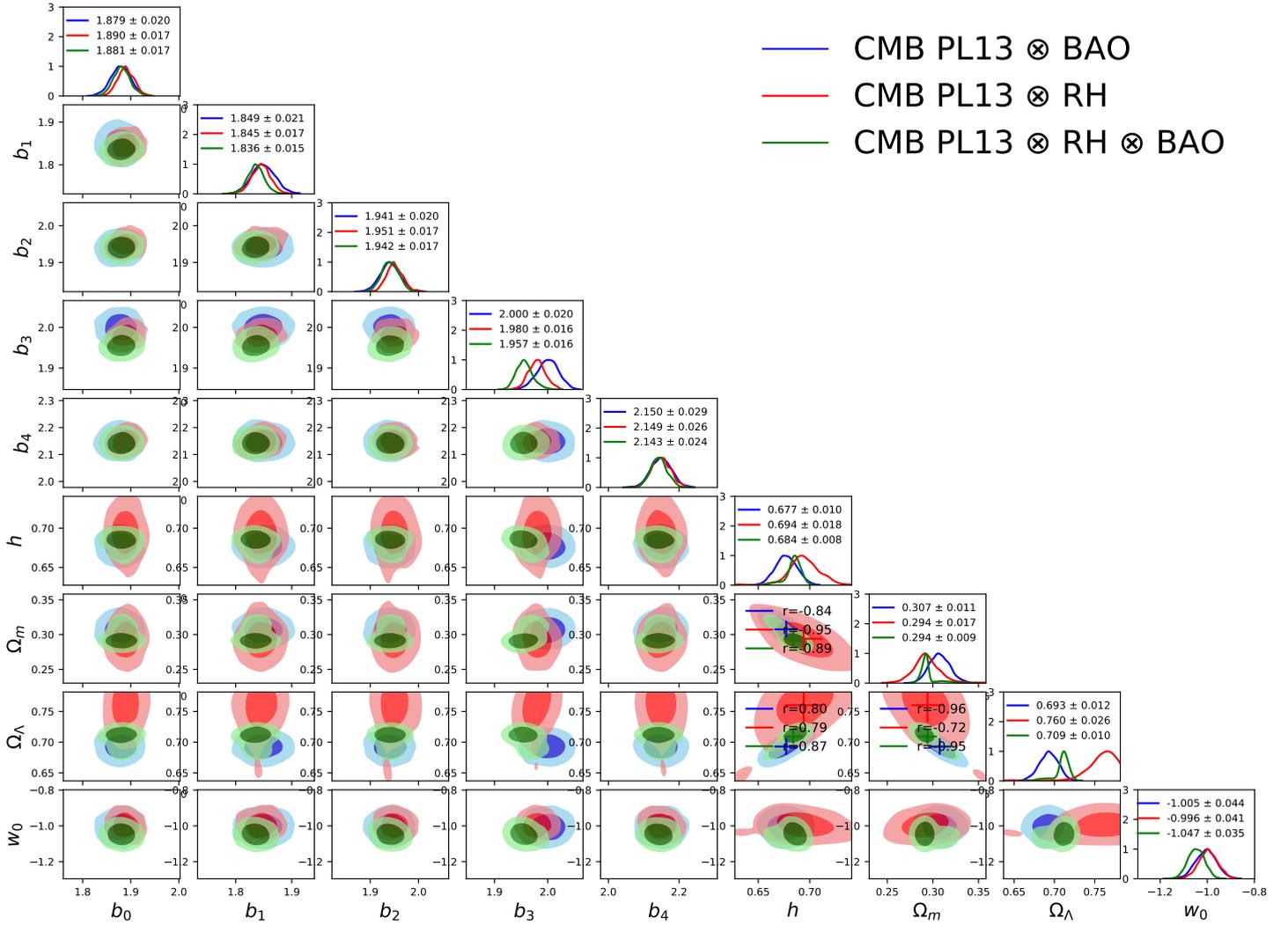


FIGURE 5.6: Contours of 68% (shaded) and 95% (light) C.L. of the homogeneity scale of the galaxy distribution for the 3 cosmological parameters of Λ CDM-model ($\Omega_m, \Omega_\Lambda, h$) and the 5 bias parameters of the 5 redshift slices of the North Galactic Cap. With blue is the CMB measurement of Planck 2013 combined with BAO, red CMB measurement combined with the \mathcal{R}_H and with green combine CMB, BAO and \mathcal{R}_H

homogeneity scale the equation of state parameter to a value, $w_0 = -0.996 \pm 0.041$ which is a better constrained obtained by the BAO and the CMB measurement, $w_0^{Planck-BAO} = -1.005 \pm 0.044$. Therefore, we find comparable constrains with the current constrains from Planck measurement [191] using the *Planck* + *WP* + *highL* + *BAO* data, last column of table 5. Finally, these results are consistent with our forecasts with the fisher matrix formalism explained in chapter 5.2, for example $\sigma_{\Omega_\Lambda}^{Fisher} / \sigma_{\Omega_\Lambda}^{MCMC} - 1 = 3\%$. Furthermore, the observed degeneracy between the Hubble constant and the Dark energy ratio density might brake by using the Alcock-Paczynski test[31] on the homogeneity scale. Finally, we observe that when we combine all three probes together, the BAO measurement the CMB

measurement and the homogeneity scale we end up constraining the cosmological parameters further. This renders the homogeneity scale a complementary cosmological probe.

5.4 Modelling the Future Surveys

In order to simulate a survey, we need first in our disposal the function that describes the number density of celestial objects such as galaxies or quasars as a function of redshift. However, we have two possibilities for that. The photometric surveys and the spectroscopic surveys. Firstly, we describe the method for a photometric survey. Mainly, we follow the method of Ma et al. [194].

Having only the photometric redshift of a galaxy, an observer can bin the galaxies with their photometric redshifts, z_{ph} , rather than the true redshift, z (i.e. the more precise redshift obtained from a spectroscopic survey). A given probability distribution, $p(z_{ph}|z)$ in z_{ph} at a given z , quantifies the overlap of the two distribution, i.e. photometric and true (spectroscopic). In general this distribution varies with z . Therefore, the true distribution of galaxies $n_i(z)$ that fall in the i -th photometric redshift bin with width $z_{ph}^{(i)} < z_{ph} < z_{ph}^{(i+1)}$ is given by:

$$n_i(z) = \int_{z_{ph}^{(i)}}^{z_{ph}^{(i+1)}} dz p(z|z_{ph}) . \quad (5.13)$$

While $n(z)$ is the overall galaxy redshift distribution parametrized by

$$n(z) \equiv dz \frac{d^3 N}{dz d\Omega} \propto \left(\frac{z}{z_0}\right)^a \times \exp\left[-\left(\frac{z}{z_0}\right)^b\right] \quad (5.14)$$

where the parameters (z_0, a, b) are the characteristic parameters for each different survey. The normalization is fixed by the total number of galaxies per steradian:

$$n^{(norm)} = \int_0^\infty dz n(z) \quad (5.15)$$

and analogously, we have that:

$$n_i^{(norm)} = \int_0^\infty dz n_i(z) \quad (5.16)$$

for the i th photometric redshift bin. In practice, the lower and upper bound (0 and infinity) turn to the minimum and maximum redshift probed by the survey, i.e. (z_{min}, z_{max}) . Notice that by construction we have $\sum_i n_i(z) = n(z)$. Therefore, regardless the complication of the photo- z probability distribution the redshift distribution of the photometric bins, i.e. the total distribution of galaxies, $n(z)$, remains unchanged. This way, we can distinguish the uncertainties due to the photometric redshifts of the individual survey galaxies characterised by $p(z_{ph}|z)$ from the uncertainties in the redshift distribution of the underlying total distribution of galaxies $n(z)$. This means that we are allowed to bin all of the galaxies together assuming that the underlying redshift distribution or selection function of the survey is considered to be known. Therefore, we consider, in practice, the known selection function to be described by a fair subsample of a spectroscopic galaxy survey. To simplify our modelling for the true photometric distribution of redshift, we consider it as a Gaussian function at each redshift, given by :

$$p(z_{ph}|z) = \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left[-0.5 \left(\frac{z - z_{ph} + z_{bias}}{\sigma_z}\right)^2\right] , \quad (5.17)$$

where the redshift bias, $z_{bias}(z)$, and standard deviation, $\sigma_z(z)$, can be any arbitrary function of redshift, z . This leads to the fact that the true distribution of galaxies, defined by Eq. 5.13, is written as:

$$n_i(z) = \frac{1}{2}n(z) [\text{erf}(x_{i+1}) - \text{erf}(x_i)] \quad (5.18)$$

where $x_i = (z_{ph}^i - z + z_{bias}(z)) / \sqrt{2}\sigma_z$ and $\text{erf}(x)$ is the error function. Notice that in the bibliography, there are two popular models for the photometric redshift distribution given by:

$$\text{Model I} : z_{bias}(z) = 0, \sigma_z = \sigma_F(1 + z) \quad (5.19)$$

$$\text{Model II} : z_{bias}(z) = 0, \sigma_z = 0.3\Theta(z - 1.0) + 0.2 \quad (5.20)$$

where $\Theta(x)$ is the step function and σ_F are the photometric redshift error at $z=0$. Ma et al. [194] have shown that the Model I is the simplest one. Therefore, we adopt this one to construct several photometric surveys by varying the parameters:

$$\mathbf{P}_{photo} = \{z_0, a, b, z_{min}, z_{max}, \sigma_F\} \quad (5.21)$$

as done by Schaan et al. [195]. However, the aforementioned models are the simplest one of the survey and do not take into account several imperfections such the ones described in chapter 3.4.2.

Now it is straightforward to model the density of a spectroscopic survey. One simply assumes that the survey observes the true underlying distribution of galaxies, and therefore assumes $p(z_{ph}|z) = 1$. Therefore the spectroscopic redshift density reduces to:

$$n_{spectro}(z) = n^{(norm)} \left(\frac{z}{z_0}\right)^a \times \exp\left[-\left(\frac{z}{z_0}\right)^b\right] \quad (5.22)$$

Therefore one varies the following parameters:

$$\mathbf{P}_{spectro} = \{z_0, a, b, z_{min}, z_{max}\} \quad (5.23)$$

for simulating the redshift number density function of a spectroscopic survey. This modelling of the cosmological surveys is the minimal one, since we have not consider experimental complications, i.e. we have consider the most simplistic selection function.

5.5 Modelling the sensitivity of \mathcal{R}_H

Here, we discuss the estimation of the sensitivity of the Homogeneity scale in spectroscopic and photometric galaxy surveys. In current galaxy surveys, the information of the large scale structures is obtained by analysing the correlation function that describes the 2-point statistics of galaxy and matter distributions of the universe and in particular, the BAO scale. Tegmark [196] provided a way to estimate the sensitivity obtained for a BAO measurement. In this study, we are going to use a similar procedure to obtain the sensitivity of the homogeneity scale. Firstly, we are going to discuss the sensitivity obtain from BAO studies and then we are going to demonstrate the sensitivity obtained from the homogeneity scale studies.

We make the usual assumption that the sensitivity of the BAO scale is proportional to the sensitivity of the measured Power Spectrum $P(k)$ of the large scale structures:

$$\frac{\delta R_{BAO}}{R_{BAO}} \propto \frac{\delta P}{P} \quad (5.24)$$

So we only need to model the precision of the power spectrum. We assume that the error of the Power Spectrum is the usual combination of the cosmic variance and the error coming from the shot-noise of the galaxy sample as shown by Blake et al. [197]:

$$\delta P = \frac{1}{\sqrt{m}} \left(P + \frac{1}{n} \right) \quad (5.25)$$

where the Power Spectrum is evaluated at the scales of first Baryon Acoustic Oscillation peak $P := P(k_*)$ and n is the density of the structured distribution. The quantity m is the total number of independent Fourier modes contributing to the measurement of the Power Spectrum. For the spectroscopic surveys it scales with the volume $m \propto V$. While for photometric surveys it scales with $m \propto V/\sigma_z$, where $\sigma_z = \sigma_F(1+z)$ is the photometric redshift error of the survey with optimistic and pessimistic scenarios given by the range $\sigma_F = (0.001, 0.01)$ for the current surveys [189]. Therefore we have:

$$\delta P \propto \frac{\sqrt{\sigma_z}}{\sqrt{V}} \left(P + \frac{1}{n_{tot}} \right) \quad (5.26)$$

where P is the normalized power spectrum and n is a normalized density.

Now, since the fractal correlation dimension is a function of the normalised count-in-spheres $\mathcal{D}_2 = \mathcal{D}_2(\mathcal{N})$ given by equation Eq. 4.3 and the normalised counts in spheres is a function of the power spectrum $\mathcal{N} = \mathcal{N}(P)$ since Eq. 4.18, we have that $\mathcal{D}_2 = \mathcal{D}_2(P)$. Thus, we can define a linear operator \hat{D}_2 acting on the power spectrum P :

$$\hat{D}_2[P] = 3 + \frac{d \ln}{d \ln r} \left(1 + \frac{1}{r^3} \int_0^r FT[P](s) s^2 ds \right) \quad (5.27)$$

Acting the operator on equation 5.26, we have that the sensitivity for each r is:

$$\delta \mathcal{D}_2(r) \propto \frac{\sqrt{\sigma_z}}{\sqrt{V}} \left[3 + \frac{d \ln}{d \ln r} \left(1 + \frac{1}{r^3} \int_0^r \int_{-\infty}^{+\infty} \left[P(k) + \frac{1}{n_{tot}} \right] e^{-iks} dk s^2 ds \right) \right] \quad (5.28)$$

Notice that the operator does not act on the ratio σ_z/V , since we treat it as a constant value. Then using the Dirac Delta property:

$$\int_{-\infty}^{+\infty} dk e^{-iks} = 2\pi \delta(s) \quad (5.29)$$

and the fact that

$$\int_{-\infty}^{+\infty} \delta(s) s^2 ds = 0 \quad (5.30)$$

with some basic algebra we see the the second term of Eq. 5.28, i.e. the one $\propto n_{tot}^{-1}$, vanishes, and therefore we end up to:

$$\delta \mathcal{D}_2(r) \propto \frac{\sqrt{\sigma_z}}{\sqrt{V}} \mathcal{D}_2(r) \quad (5.31)$$

which describes the error of Fractal correlation Dimension scaling with the Volume (V) and Redshift Error (σ_z) but not with density (n). Thus by also assuming that the fractional error on the Homogeneity scale is proportional to the fractional error of the fractal correlation dimension:

$$\frac{\delta \mathcal{R}_H}{\mathcal{R}_H} \propto \frac{\delta \mathcal{D}_2}{\mathcal{D}_2} \propto \frac{\sqrt{\sigma_z}}{\sqrt{V}} \quad (5.32)$$

In the next section we are going to discuss the scaling relations of the BAO scale and the Homogeneity scale coming from simulations of the potential surveys.

5.5.1 Simulating the Scaling Laws of Precision

In order to test the theoretical argument described in the previous section (section 5.5) we are going to simulate a realistic case. The fractional error of our current cosmological probes (the BAO scale \mathcal{R}_{BAO} and the Homogeneity scale \mathcal{R}_H) depends on the volume (V), the density (D) of the survey as well as the photometric-redshift error (σ_F) which we define as the survey variables $\mathbf{p}_{sv} = (V, n, \sigma_F)$.

We define the following scaling models, inspired from the results of the previous section, for the fractional error of each cosmological probe ($\mathcal{R}_X = \{\mathcal{R}_H, r_s\}$) with survey-variables dependence. Therefore we take the error of the BAO as defined by Eq. 5.24 and 5.26 and we rewrite it as:

$$\frac{\delta R}{R} = \frac{\sqrt{\sigma_z}}{\sqrt{V}} \left(1 + \frac{1}{nP} \right) \quad (5.33)$$

then we expanded in first order for the $\mathbf{p}_{sv} = (V, n, \sigma_F)$ each with a corresponding variation, for the volume ϵ_V , the density ϵ_n and the photometric error, ϵ_{σ_z} . We find that:

$$\frac{\delta R}{R} = \frac{\delta R}{R} [\sigma_{z0}, V_0, n_0] \left[1 + \alpha \epsilon_{\sigma_z} + \beta \epsilon_V + \gamma \frac{\epsilon_n}{nP} \right] \quad (5.34)$$

around the initial values, σ_{z0}, V_0, n_0 and $p = (\alpha, \beta, \gamma)$ are the scaling parameters and $nP = 3$ for the BOSS survey[?]. Thus, we are going to vary the scaling parameters individual for each different probe. We expect for the BAO scale that $\alpha = 1/2$, $\beta = -1/2$ and $\gamma = -1$.

We use the 100 QPM mock catalogues to study the fractional precision of the cosmological probes with the scaling of the 3 survey-variables. We define 15 configurations by varying the survey variables at each QPM mock catalogue as:

$$V = (90, 80, 70, 60, 20) \% \quad (5.35)$$

$$n = (90, 80, 70, 60, 20) \% \quad (5.36)$$

$$\sigma_F = (1, 2, 5, 10, 15) \times 10^{-3} \quad (5.37)$$

where the percentage is the percentage ratio in respect of the original volume or density of the CMASS galaxy sample at redshift $0.538 \leq z \leq 0.592$, on the North Galactic Cap. We decrease the volume by 5.35. Similarly, we decrease the density of each catalogue by shot-noising the galaxy sample by the percentage rate described by equation 5.36. Furthermore, we add photometric error on each QPM catalogue by the "Reshuffle" method as explained in chapter 5.5.2 using the different parameters defined by Eq. 5.37. Then, we measure the BAO scale with the method 1 of chapter 5.1.1 and the homogeneity scale with the spline method for the galaxy and matter distribution, as explained in chapter 4, at each configuration on each individual mock catalogue. Finally, we use the mean and one standard deviation of the precision computed of the 100QPM mock catalogues to obtain the scaling parameters $\mathbf{p}_{scaling}$.

5.5.2 Reshuffle Method

In order to mimic a photometric survey, we use the method called *Reshuffle*. Briefly in this method, we use the spectroscopic redshift information for each galaxy of our CMASS catalogue from our survey discussed in section chapter 3.4.1. Then, we assign a gaussian probability on each galaxy of the catalogue with mean its redshift and dispersion given by $\sigma_z = \sigma_F(1 + z)$. Then, we draw randomly from this probability a new redshift and we replace the old one. Finally, we have a new galaxy catalogue with "reshuffled" or "distorted" redshifts mimicking a photometric survey.

By assigning the gaussian probability at each galaxy, there are galaxies that lie in the region out of the redshift slice. This result in an increase of the volume since it increases the redshift space from the reshuffling method. In order to keep the volume constant, we perform the reshuffle method in the whole CMASS survey area and we then discard the galaxies that fall out of the z -slice.

5.5.3 Results of scaling laws

In this section, we keep investigate the fractional error of the two probes, the BAO scale and the Homogeneity scale as a function of the volume density and photometric redshift error as defined by the model Eq. 5.34 by estimating the scaling parameters $p = (\alpha, \beta, \gamma)$. The measurements are performed on the galaxy distribution using the method described in section 4.2.1. We present the results in diagrams of normalized fractional error as a function of the normalized volume or density or photometric redshift error. The normalization was done against our baseline measurement which is the measurement on the BOSS redshift slice $0.538 \leq z \leq 0.592$.

We first examine the fractional error of the two probes as a function of the volume. The results are shown in Fig. 5.7. With blue data points, we have the simulated data for the normalized fractional error for the BAO scale. Notice that the errorbars are so small that are not visible on the plot. With blue continuous line we display the fitted model for $\alpha = \gamma = 0$. The fitted value parameter is $\beta = -0.48 \pm 0.12$ which is consistent with the expected value $-1/2$ at 3σ . With green data we display the normalized fractional error of the homogeneity scale. With green continuous line, we display the model as defined finding $\beta = -0.55 \pm 0.16$ which is consistent with the expected value $-1/2$ at 2σ .

We then examine the fractional error of the two probes as a function of the density. The results are shown in Fig. 5.8. With blue data points, we have the simulated data for the normalized fractional error for the BAO scale. Note here that the errorbars are so small that are not visible on the plot. With blue continuous line, we display the fitted model for $\alpha = \beta = 0$. The fitted value parameter is $\gamma = -1.33 \pm 0.35$ which is consistent with the expected value -1 at 1σ . With green data, we display the normalized fractional error of the homogeneity scale. With green continuous line, we display the model finding $\gamma = -0.16 \pm 0.17$ which is not consistent with the expected value -1 . This tells us that the homogeneity scale has no dependence with the density.

We finally examine the fractional error of the two probes as a function of the photometric redshift error. The results are shown in Fig. 5.9. With blue data points, we have the simulated data for the normalized fractional error for the BAO scale. With blue continuous line, we display the fitted model for $\beta = \gamma = 0$. The fitted value parameter is $\alpha = 0.52 \pm 0.04$ which is consistent with the expected value $1/2$ at 1σ verifying our model. With green data, we display the normalized fractional error of the homogeneity scale. With green continuous line, we display the model finding $\alpha = -0.01 \pm 0.03$ which is consistent with 0 at 1σ . This tells us that the homogeneity scale has no dependence with the photometric redshift error.

Evidently, the homogeneity scale varies only with the volume of the survey. On the other hand, the BAO scale varies with the precision of the photometric redshift, and the density of the emitting source in addition with the volume. For these reasons, for our calculations, we use the precision of the BAO scale as defined by:

Notice that we have verified the precision law for the BAO scale, given by Eq. 5.33.

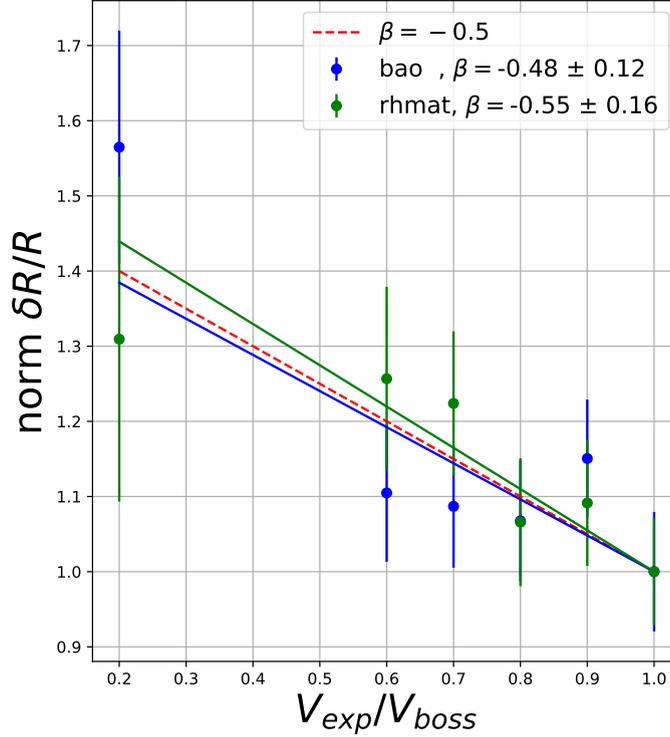


FIGURE 5.7: Normalized fractional error $\delta R/R$ as a function of volume, V , for the bao scale (blue) and the homogeneity scale (green) for the the galaxy distribution. [See text for details.]

While for the Homogeneity scale for the galaxy distribution, we use:

$$\frac{\delta R_H}{R_H} \propto \frac{1}{\sqrt{V}} \quad (5.38)$$

Here, it is important to stress out that we observe that the precision of \mathcal{R}_H does not vary with the redshift errors, while the BAO one is strongly degraded for large redshift errors. Therefore, the homogeneity scale is a better probe for the photometric surveys rather than the BAO scale.

5.5.4 Results with the Global Fitting Formula

We simulate a range of current future surveys as described in chapter 5.4. We measure for each simulation the precision of the homogeneity scale and the BAO scale as described in chapter 5.5.1. In this study, we consider the realistic configurations of the different experiments with different field of view, $d\Omega$, and cosmic depth $\Delta z = z_{max} - z_{min}$. We list them in table 5.3.

Instead of simulating the number density for BOSS, DESI and EUCLID projects, we use more realistic simulations as used by Font-Ribera et al. [198]. For the WFIRST and LSST, we use the aforementioned simulated methods described in section 5.4. Notice that we have consider an optimistic scenario for the EUCLID experiment. A pessimistic scenario would be with $\sigma_F = 0.01$. Furthermore, we calculate the

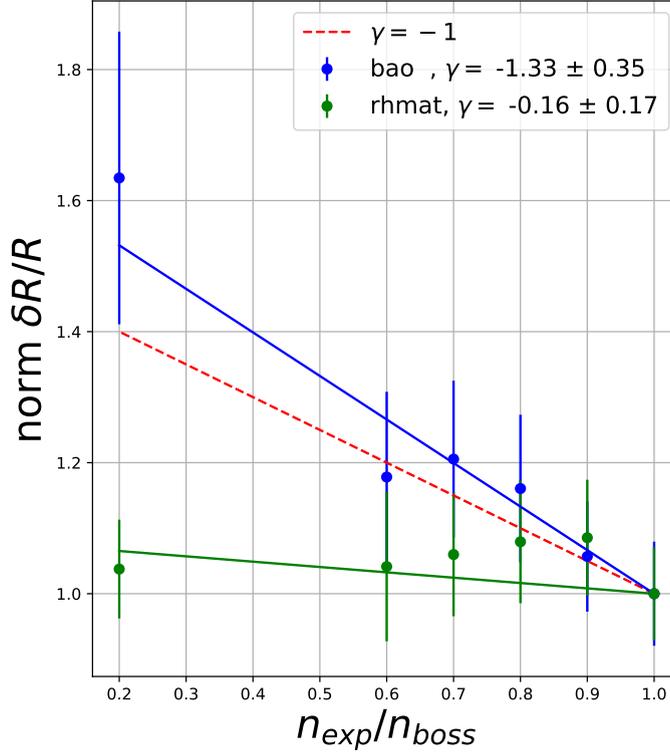


FIGURE 5.8: Normalized fractional error $\delta R/R$ as a function of density, n , for the BAO scale (blue) and the homogeneity scale (green) for the galaxy distribution. [See text for details.]

Experiment	BOSS	eBOSS	DESI	EUCLID	LSST	WFIRST
z_0	0.5	1.2	1.0	0.65	0.51	0.60
a	2.0	2.0	2.0	1.3	1.27	1.27
b	8.0	2.0	8.0	1.5	1.02	1.02
z_{min}	0.05	0.55	0.15	0.65	0.1	1.05
z_{max}	0.75	2.15	1.85	2.05	4.9	2.75
σ_F	0.001	0.001	0.0005	[0.001 – 0.01]	0.003	0.003
$d\Omega$	10, 400	2, 400	14, 000	15, 000	18, 000	2, 200

TABLE 5.3: Configuration of the simulation of the redshift density profile and survey area of the different experiments.

characteristic comoving volume of each experiment using:

$$V_{exp}(z; p_{cosmo}) = \frac{c}{H_0} \frac{(1+z)^2 d_A^2(z; p_{cosmo})}{E(z; p_{cosmo})} \Delta z_{exp} d\Omega_{exp} \quad (5.39)$$

where $E(z; p_{cosmo})$ is given by the ratio of Eq. 1.27 and the Hubble constant, H_0 , while $d_A(z; p_{cosmo})$ is the comoving angular diameter distance for the fiducially parametrized cosmology, p_{cosmo} .

For the BAO probe, we observe that b is consistent with 0 at 1σ C.L. and it is also consistent with both the volume and photometric redshift error scaling laws in the literature [189, 197]. For its density dependence, we are consistent with the literature as well finding we find $b = -0.657 \pm 1.121$ expected with the 0 value. This shows that the scaling with the density, for the BAO probe, has the same dependence

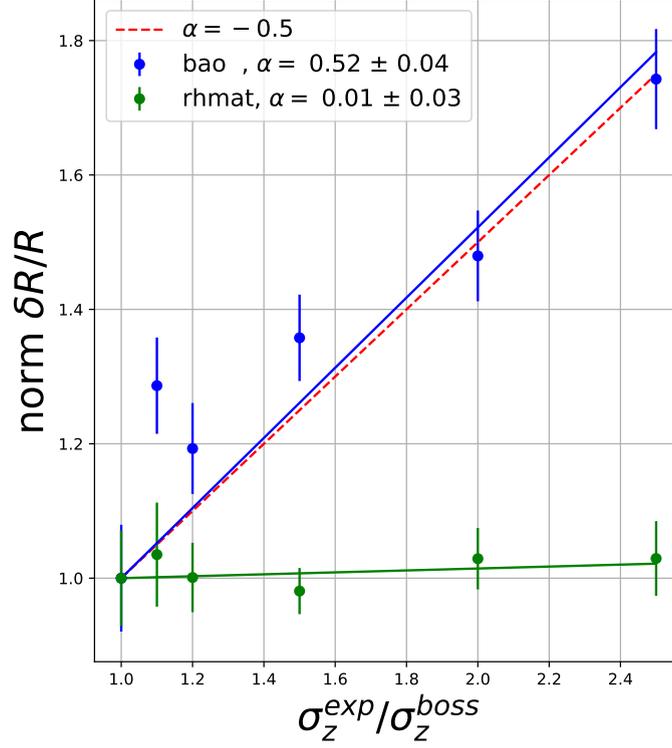


FIGURE 5.9: Normalized fractional error $\delta R/R$ as a function of photometric redshift error, σ_z , for the bao scale (blue) and the homogeneity scale (green) for the galaxy distribution. [See text for details.]

than the one obtained by Blake et al. Fitting Formula. Therefore, we follow Albrecht et al. [189] and we model empirically the precision of the BAO scale as:

$$\delta \ln \mathcal{R}_{BAO}(\sigma_F, V, D; A_{BAO}) = A_{BAO} \sqrt{\frac{\sigma_F(1+z)}{\sigma_{F,0}} \frac{V_0}{V}} \left(1 + \frac{1}{3n}\right) \quad (5.40)$$

where A_{BAO} is a free parameter that we fit with a global fit. Furthermore, $\sigma_{F,0} = 34h^{-1}\text{Mpc}/E(z; p_{\text{cosmo}})$ is the characteristic photometric redshift error obtained by Albrecht et al. [189] and $V_0 = 2.16h^{-3}\text{Gpc}^3$ is the characteristic volume of the precision studies [199]. Notice also that we have followed Albrecht et al. [189] and we have normalized the density dependence with the *ratio of the true clustering power to that from the shotnoise of our survey*, i.e. $nP = 3$. The global fit give us a value of $A_{BAO} = 0.0107 \pm 0.0029$ for the perpendicular to the line of sight which is within 2σ with the literature value $A_{BAO}^\perp = 0.0123$ or $A_{BAO}^\parallel = 0.085$ or 0.0145 for the transverse or perpendicular to the line of sight as measured by Blake et al. [197].

Since the homogeneity scale depends only with the volume, we model empirically the precision of the galaxy homogeneity scale as:

$$\delta \ln \mathcal{R}_{HOMO,G}(\sigma_F, V, D; A_{\mathcal{R}_H^M}) = A_{\mathcal{R}_H^M} \sqrt{\frac{V_0}{V}} \quad (5.41)$$

where $A_{\mathcal{R}_H^G}$ is a free parameter that we fit using our measurement and we find $A_{\mathcal{R}_H^G} = 0.0063 \pm 0.0021$.

Then we calculate the precision of the homogeneity scale and the BAO scale of each experiment, which are displayed as a function of redshift in Fig. 5.10.

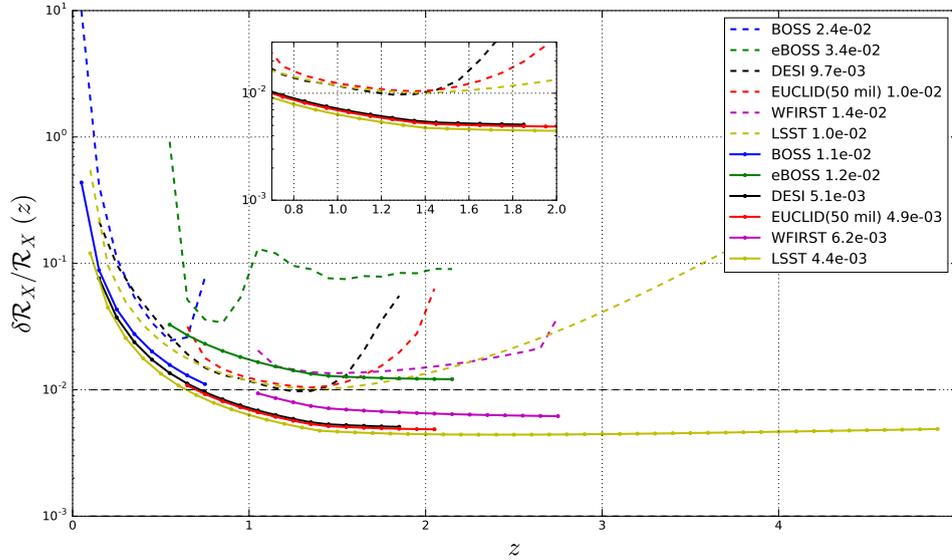


FIGURE 5.10: Precision of the scale, \mathcal{R}_X versus the redshift, z for the different experiments which are color-coded. The dash lines correspond to the BAO scale while the continuous lines correspond to the homogeneity scale for the matter distribution. See text for details

With dash lines, we plot the precision for the BAO scale as a function of redshift and they are color-coded with the different experiments. We display with continuous lines the precision for the homogeneity scale under the same color-code. Firstly, we can observe that for the BAO scale the precision is consistent with the one obtained by the literature [198]. Furthermore, we observe that the precision of the homogeneity scale does not depend on the density of the sources that we have as a function of redshift. This can be seen from the fact that the precision is smoothed out for all experiments in respect to the peculiar behaviour of the precision of the BAO, which is a result of the density dependence. We zoom in the region with the best precision that we obtained for the leading future experiments DESI, EUCLID and LSST.

In particular, we observe that the BAO scale, using the DESI experiment (black dash line), achieves better maximum precision 0.97% than using the EUCLID (red dash line) or LSST (yellow dash line) that both follow similar precisions, 1.1% and 1.0% respectively. This tells us that DESI will lead the BAO oscillation precision measurements on future large scale structure surveys. While for the homogeneity scale precision of the LSST experiment (continuous yellow) reaches a maximum precision of 0.44% leading the precision measurements of the homogeneity scale. Notice that this calculation assumes that we measure the same bias for all our tracers. In a more realistic case the bias varies with the different tracer and the picture becomes more complicated. Furthermore, we observe that overall precisions of the BAO scale is 0.5% while for the homogeneity scale is 0.24% rendering the homogeneity scale more precise measurement. Notice that this precision is somehow a density precision and if someone needs to find the total precision needs to integrate over redshift the inverse of this quantity [198]. Notice also that this result is due to the fact that the measurement of the homogeneity scale for the matter distribution results in scales less than the BAO scale where we are shot-noise limited. These results render the homogeneity scale a better probe of cosmology than the BAO scale for photometric surveys. However, it should also

be taken into account that the method, used here, is an approximate solution and an extrapolation was performed for the the variable regime of the future experiments.

5.6 Conclusion

We have compared the measurement of the transition to cosmic homogeneity scale, \mathcal{R}_H with the BAO scale, \mathcal{R}_{BAO} , using the CMASS galaxy sample. We found that there is insignificant correlation between the two probes, $r \sim 20\%$, rendering the the transition to cosmic homogeneity scale a complementary cosmological probe. In order to quantify the additional information of the homogeneity scale, we have provided the Fisher Matrix of each theoretical prediction against a basic set of cosmological parameters. We found that the homogeneity scale result into comparable information to that of the BAO scale using large scale structure to the measurement of this set of cosmological parameters. In particular, we found that the homogeneity scale is sensitive to the Hubble constant, H_0 , the total matter ratio density, Ω_m , the Dark energy ratio density, Ω_Λ , and the equation of state as a function of time or redshift, i.e. parametrized with two parameters (w_0, w_a) . Using an MCMC exploration algorithm and CMB measurement as prior knowledge, we have found for the basic parameters comparable constrains than the one obtain by Planck measurement[191]. In particular, we have found that the hubble constant is $H_0 = 67.6 \pm 0.8 \text{ km/s/Mpc}$, the total matter ratio density is $\Omega_m = 0.308 \pm 0.009$. The Dark energy ratio density is $\Omega_\Lambda = 0.727 \pm 0.0015$ which in tension with the Planck value $\Omega_\Lambda^{Planck} = 0.681 \pm 0.0016$. Finally, we have explored the sensitivity of the homogeneity scale for future experiments that will map the large scale structures and we have compared it to the sensitivity of the BAO scale. In particular, we have found that the overall precision of the homogeneity scale for six projects equal to 0.24%. The leading experiment will be the ground based experiment LSST achieving a maximum sensitivity of 0.44%. On the other hand, the BAO scale will be lead by the ground based experiment DESI providing a maximum density of 0.97%. The overall density of all experiments for the BAO studies will be 0.5%. These results render the homogeneity scale a better probe than the BAO scale for photometric surveys. However, bare in mind that the comparison method is a first order approximation and an extrapolation on the variable regime of the future experiments, thus a more detailed study is required for a quality test of this result.

Chapter 6

Cosmic Homogeneity with Cosmic Clocks and type Ia SN

In this chapter, we discuss the validation the Cosmological Principle in a model independent manner. We are using the redshift information through a model independent transformation to infer distance between us and the galaxies. This work was performed in collaboration with James Rich.

6.1 Introduction

So far we have explored a model dependent measurement of the homogeneity scale. We have use the astrometry of the SDSS telescope to obtain the angular positions of our galaxy sample and the spectroscopic information to obtain the redshift measurement of each galaxy. Then, by assuming the cosmological principle, we have used the FLRW metric to infer the radial distances from corresponding redshifts of each galaxy. We then have used the fractal correlation dimension as an indicator of homogeneity and we have found that it becomes homogeneous in a model dependent way.

However, in order to validate the cosmological principle in a model-independent way, we need to use the redshift information of each galaxy source through a model independent transformation to infer distances between us and the emitting sources as well as among the emitting sources (galaxies). Jimenez and Loeb [200] have measured the cosmological parameters using the relative galaxy ages. This method assumes only astrophysical assumptions in order to determine the relative redshift against age of the different galaxies that are close in a small redshift bin, Δz . Furthermore, in order not to assume any curvature of the cosmic spacetime, we are going to use the luminosity distance relation to extract the transverse distances among the emitting sources, which assumes only astrophysical models. Therefore in our study, we are going to use a model independent relation to transform redshifts into distances, rendering our study free of assuming the Cosmological Principle. As a result, in this chapter, we are going to assume only a minimal set of information about our universe. These assumptions are summarised as:

1. Metric Theory of Gravity
2. Stellar Population Synthesis model

3. Supernova Astrophysical Luminosity model
4. Smooth expansion history of the universe
5. Universe looks the same outside of our current survey geometry
6. Flatness of the universe

Notice that the first assumption is necessary since we are observers that we measure distances within a complex gravitational field generated by the different gravitational bodies existing in the universe. The second one is an assumption of constructing the CMASS sample that assumes that the properties of the galaxies do not change with time, i.e. rendering this sample volume limited. Furthermore, the stellar population synthesis model will help as on the determination of the hubble expansion rate. The third one is necessary since we need to have the information of the amount of luminosity at different wavelengths that each galaxy emits so that not to assume that the universe has an arbitrary curvature of spacetime. The fourth one is another cosmological fact and therefore is evidently taken into account but it does not assume a homogeneous or isotropic universe. Notice that this assumption assumes neither homogeneity nor isotropy. The final assumption is another rephrasing of the fact that we are limited to only a small sample of our universe in respect of the whole one, which is impossible to test with our current observational devices. Notice that we do not restrict ourselves in a particular metric for our measurements, rendering our framework free of homogeneity and isotropy assumption.

6.2 Methodology

We are using the str-estimator, i.e. using Eq. 4.13, to estimate the fractal correlation dimension¹, in order not to assume homogeneity as an assumption. The weakness of this estimator is that it is not as precise as the cor-estimator but the precision is not relevant for this study. Furthermore, as explained in chapter 4.1.1, we are using a random catalogue to correct for the geometrical effects of our survey. Additionally, to correct for the efficiency of our telescope, i.e. the amount of galaxies per redshift bin, we impose that the random catalogue has the same redshift distribution as the galaxy catalogue. Therefore we are not able to work around this restriction. A possible solution would be to use surveys that have galaxy densities that do not vary with redshift. However, this is not the case for our current surveys. We also use the Eq. 4.39 to estimate the homogeneity scale from our dataset, i.e. we use the minimum scale at which the fractal dimension reaches reaches the homogeneity value, 3, within 1σ error.

We use the redshift slice $0.538 < z < 0.592$ on the North Galactic Cap (NGC) to perform this measurement. We weight the galaxies as explained in chapter 3.4.2. This weighting scheme does not introduce any extra assumptions on the construction of the sample. Furthermore, we show in chapter 4.3.11 this introduces no significant deviations for the measurement of the homogeneity scale.

Throughout this study, in order to compare our model independent results with the model dependent (MD) ones that we obtained in the previous chapters, we use a flat Λ CDM cosmology defined via:

$$\mathbf{p}_{\Lambda\text{CDM}} = (\Omega_m, \Omega_\Lambda, h) = (0.274, 0.726, 0.70) \quad (6.1)$$

where Ω_m is the total matter energy density ratio in the universe, Ω_Λ is the Dark Energy density ratio in the universe and $h \equiv \frac{H_0}{100 \text{ km s}^{-1} \text{ Mpc}^{-1}}$ is the hubble dimensionless parameter. Moreover, in order to determine the galaxy survey uncertainties, we use 20 QPM[150] mock catalogues constructed via a flat

¹ We have used a modified version of the publicly available code <https://github.com/damonge/CUTE/tree/master/CUTE>. For more information look at [201].

Λ CDM cosmology defined via Eq. 6.1. The number of mock catalogues is small but sufficient for the sake of this measurement. In the next section, we describe how to reconstruct a (z, θ, ϕ) -distribution into a 3Dimensional cartesian one using cosmic clocks and SuperNovae data only. Therefore, the main difference, in respect of the previous chapters, is that we are going to use a model independent (MI) way to reconstruct our distances as we explain in the next section.

6.3 3D Space reconstruction

In order to reconstruct our 3D galaxy distribution assuming no homogeneity or isotropy, we need to measure distances without assuming a cosmological model. Firstly, we are considering the radial dependence on redshift. We then consider the transverse distance (or motion distance) between two objects at a certain redshift.

We substitute the usual comoving redshift-distance relation of Λ CDM model with the new relation of cosmic chronometers. The Λ CDM predictions of the comoving distance is given by Eq. 1.31. In the new method, we reconstruct distances using cosmic chronometers[202] to define the redshift-distance relation. Cosmic chronometers are used to measure the Hubble expansion rate, $H_{clock}(z)$. Thus, we can measure the radial comoving distance of each galaxy via:

$$d_{clock}(z) = c \int_0^z \frac{dz'}{H_{clock}(z')} \quad (6.2)$$

and $H_{clock}(z)$ will be a linear model as:

$$H_{clock}(z) = Az + B \quad (6.3)$$

where A and B are two independent parameters with the same units as the hubble rate $H(z)$ that are fitted to the measurement of Moresco et al. [202] for the two possible assumptions of the stellar population synthesis algorithm as explained in chapter 6.4.1.

Secondly, we are considering that the transverse distances between the galaxies. To measure the transverse comoving distance without assuming any cosmology we need a relation that gives transverse comoving distances. The transverse comoving distance (or motion distance) is related to the luminosity distance as

$$d_M(z) = \frac{1}{1+z} d_L(z) . \quad (6.4)$$

Therefore instead of assuming a cosmological model for the transverse comoving distance and in extend the homogeneity assumption or a curvature about the universe we use an astrophysical model for the definition of the luminosity distance. The motivation behind the use of the transverse distance is that we use this distance to infer the transverse size of objects via: $d_M(z)d\theta = d_A(z)d\theta$. Riess et al. [203] have shown that one can obtained the luminosity distance through the distance module, $\mu(z)$, as:

$$d_L(z) = 10^{[\mu(z)-25]/5} \quad (6.5)$$

Their measurement of $\mu(z)$ was performed under the astrophysical assumption that the apparent magnitude has been corrected from the interstellar absorption. Therefore, we can use their measurement and we can simply find a linear model for determining the distance module, $\mu = \mu(z; A_\mu, B_\mu)$ where A_μ and B_μ are two parameters that capture the astrophysical assumptions. We determine these parameters in

section 6.4.2. Then the luminosity distance is given by:

$$d_L(z; A_\mu, B_\mu) = 10^{[\mu(z; A_\mu, B_\mu) - 25]/5} \quad (6.6)$$

The transverse comoving distance is only a function of the redshift and the astrophysical parameters (A_μ, B_μ) defined as:

$$d_M(z; A_\mu, B_\mu) = \frac{1}{1+z} 10^{[\mu(z; A_\mu, B_\mu) - 25]/5} \quad (6.7)$$

Now, we reconstruct the astrophysical coordinates of our galaxy catalogue (z, θ, ϕ) into cartesian coordinates (x_c, y_c, z_c) via the transformation:

$$x_c = d_M(z; A_\mu, B_\mu) \sin \theta \cos \phi \quad (6.8)$$

$$y_c = d_M(z; A_\mu, B_\mu) \sin \theta \sin \phi \quad (6.9)$$

$$z_c = d_{clock}(z; A, B) \cos \theta \quad (6.10)$$

So using this transformation, we assume only an astrophysical model which is quantified by the parameters (A, B, A_μ, B_μ) . Therefore we are free of assuming homogeneity or isotropy of our galaxy catalogue and in extend the universe.

6.4 Results

6.4.1 Hubble Rate calibration with Cosmic Chronometers

Since Moresco et al. [202] used 2 different stellar population models to measure the hubble expansion rate, namely MaSTro library (HM) and BC03 (HB), we use both of them to measure the $H_{clock}(z)$, in order to cross validate the result. As shown in the left panel of Fig. 6.1, we have two simple solutions, namely HM (green line) and HB (blue line), of Eq. 6.3 by fitting Moresco et al. [202] data. HM correspond to $(A, B) = (58, 63)$, while HB corresponds to $(A, B) = (36, 75)$ solutions of Eq. 6.3. As a result, we have measured the Hubble expansion rate using the cosmic chronometers method by assuming only a stellar population synthesis model, i.e. assuming no homogeneity neither Isotropy of the galaxy distribution.

For comparison, we display with red line the Λ CDM-model. We observe that our simplistic models and the Λ CDM-model, all agree with both datasets.

6.4.2 Luminosity Distance Calibration using SN

We use the data of type Ia supernova of the measured apparent magnitude at redshift $0.30 \leq z \leq 0.98$ given in table 6 of Riess et al. [204]. By using a simple model for the apparent magnitude, i.e.:

$$\mu(z; A_\mu, B_\mu) = A_\mu \log_{10}(z) + B_\mu \quad (6.11)$$

where A_μ, B_μ are the nuisance parameters that contain all the astrophysical information we constrain the apparent magnitude, i.e. we measure that $(A_\mu, B_\mu) = (5.72 \pm 0.48, 44.32 \pm 0.15)$ for $\chi^2 = 7.11$ per 8 degrees of freedom. We present the result of our model in Fig. 6.2.

Therefore, we are able calibrate the luminosity distance relation using Eq. 6.6 at the redshift of our interest, $z = 0.538 - 0.592$. This enable us to determine the transverse comoving distance among the emitting sources, d_M , without assuming a curvature for the spacetime.

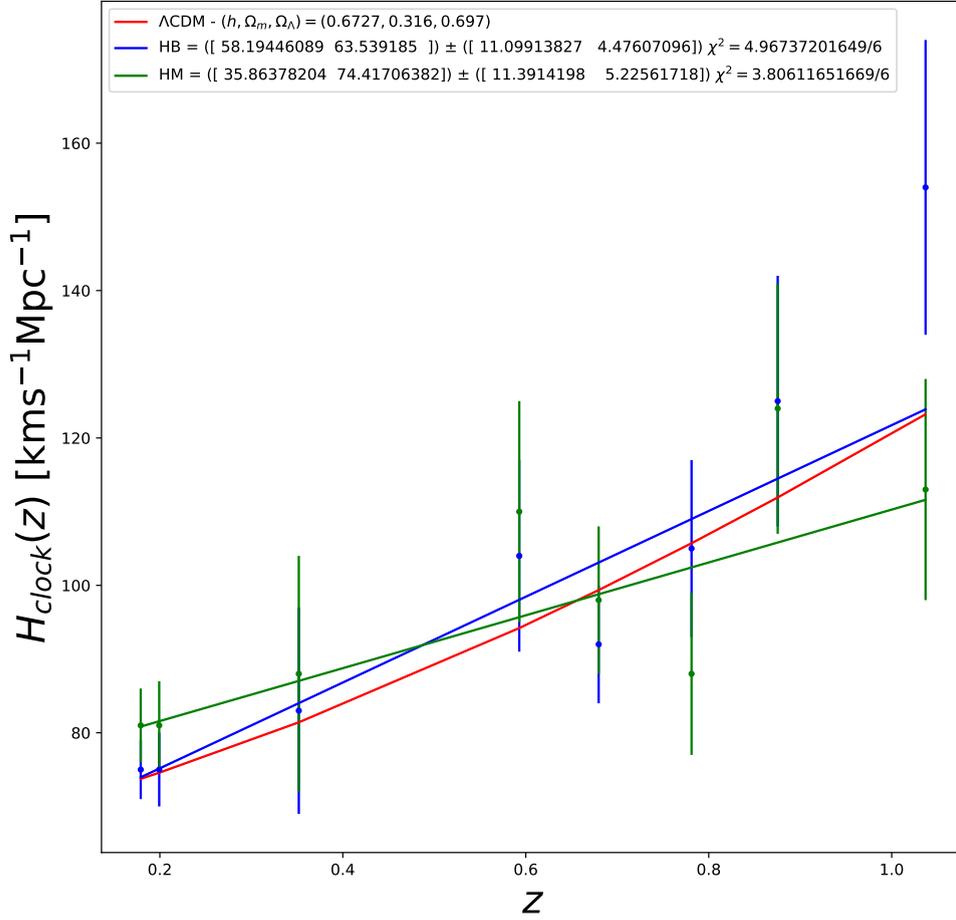


FIGURE 6.1: *Left:* Calibrating the Hubble expansion rate using cosmic chronometers. [See text for details]

6.4.3 Model independent Cosmological distances

We summarise all the different distances as described by Eq. 1.31, Eq. 6.7 and Eq. 6.2 in Fig. 6.3. We observe that the cosmic clock distance, d_{clock} , is compatible with the comoving distance, d_C , at less than 5% throughout the redshift slice that we are interested in. On the other hand, the transverse comoving distance, d_M is compatible with the comoving distance at less than 13%.

6.4.4 Fractal Correlation Dimension with Cosmic Chronometers and SN

We measure the fractal correlation dimension, \mathcal{D}_2 , as a function of scales, r , which is displayed in the right panel of Fig. 6.4, for the 3 different ways of reconstructing distances from redshift, in 50 bins in the range $r = [0, 1500]$ Mpc. With red, we plot the Fractal Correlation dimension using the Λ CDM redshift-distance relation. With blue(green), we display the fractal correlation dimension measurement for the

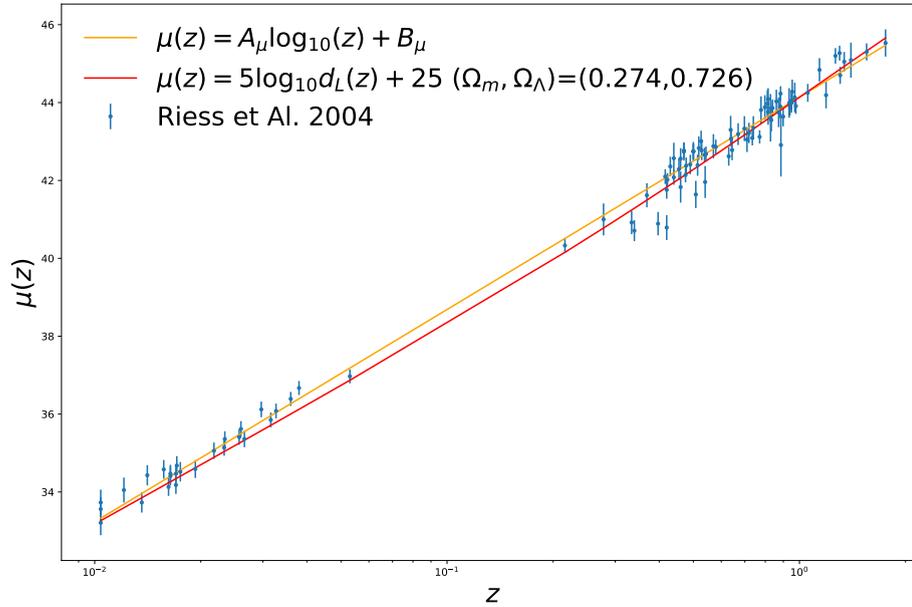


FIGURE 6.2: Supernovae distance module, μ , as a function of redshift, z . Data points are denoted with blue, red line is the function of Λ CDM-model and orange is a simple relation between the distance module and the redshift.

HM(HB) cosmic chronometers calibration of the redshift distance relation and the luminosity distance calibration using SN data, namely HMF(HBF) respectively. With grey color, we plot the 20 QPM mock catalogue that we used in this analysis to estimate our uncertainties. We observe that the fractal dimension, of the CMASS galaxy sample, is an increasing function of the separations, r , for all different methods of estimations, and asymptotically reaches the value of 3 for all three different methods of estimation.

This shows that the CMASS galaxy sample becomes homogeneous on scales $\mathcal{R}_H \simeq 150\text{Mpc}$ as defined by Eq. 4.39, in a model independent way. We can, also, observe that the estimation of the fractal dimension using Λ CDM-method is compatible with the HMF and the HBF methods at all scales. We also find that the galaxy distribution becomes homogeneous at scales larger than 400Mpc for the HMF method and Λ CDMmethod. For the HBF method the galaxy distribution becomes homogenous at scales larger than 300Mpc. This might be a statistical fluke since we are using only 20 QPM mock catalogues for the determination of our variance, which is a poor statistical choice. However, the fractal correlation dimension for the galaxy distribution becomes homogeneous at scales larger than 400Mpc, confirming the homogeneity assumption for the galaxy distribution in a model independent way.

For the matter distribution we need to assume a bias between the galaxy distribution and the matter distribution. This means that still we need to assume the homogeneity assumption to confirm the cosmological principle which is still a model dependent way.

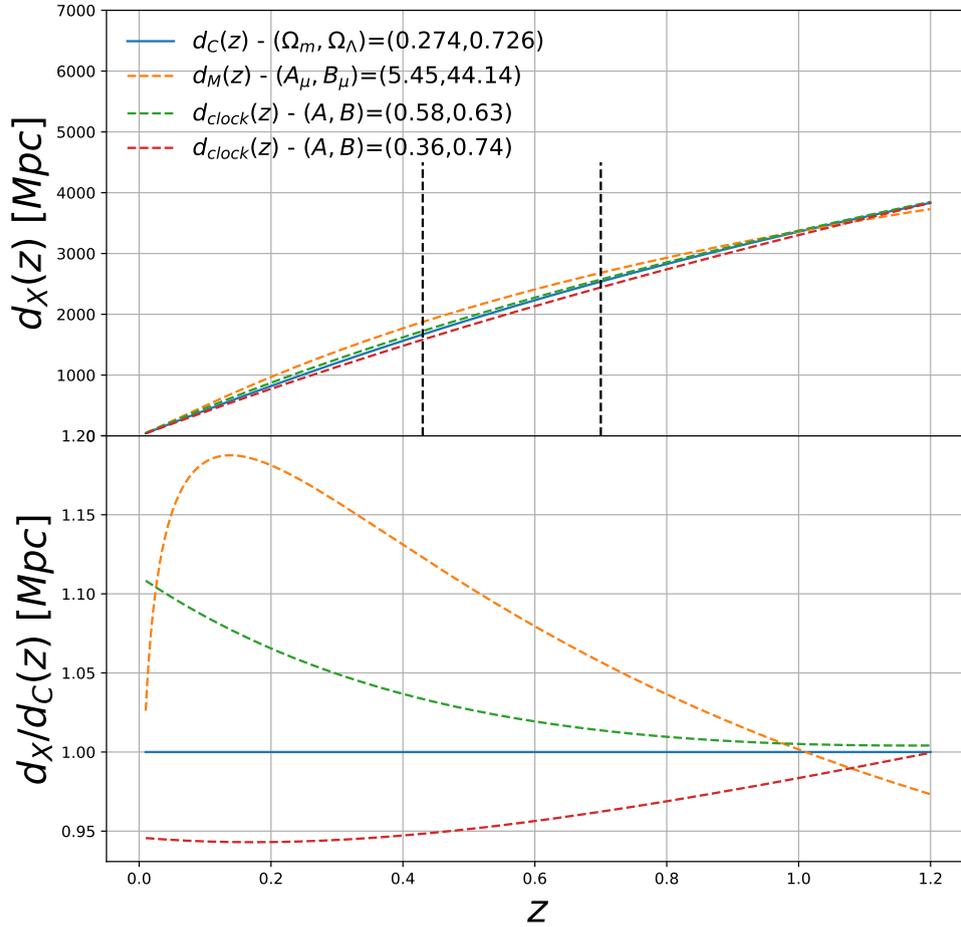


FIGURE 6.3: (Top) Cosmological distance versus redshift for the FLRW comoving distance, d_C (blue), model-independent transverse comoving distance, d_M (orange), and the two model independent cosmic clock comoving distances, d_{clock} (green and red). The black vertical lines indicate the redshift cuts of our survey. (Bottom) Different cosmological distances over comoving distance versus redshift.

6.5 Conclusion

We have studied the transition to cosmic homogeneity in a model independent fashion using the 12th data release of BOSS CMASS galaxy sample. We have shown, that using an alternative estimation of the radial distance from redshift data, i.e. free of assumptions of Homogeneity and isotropy, the fractal correlation dimension still reaches the homogeneity value, 3, asymptotically. In particular we have used two cosmic chronometers methods to calibrate the hubble expansion rate, which is a model independent estimation of the radial comoving distance. Additionally, we have used the high- z and low- z Supernovae Data to calibrate the comoving luminosity distance in a cosmological model independent way and then measure the transverse comoving distance. Furthermore, we show that the fractal correlation dimension for the galaxy distribution becomes 3 at scales larger than 400Mpc for both model independent methods which is compatible with the fiducial Λ CDM-method.

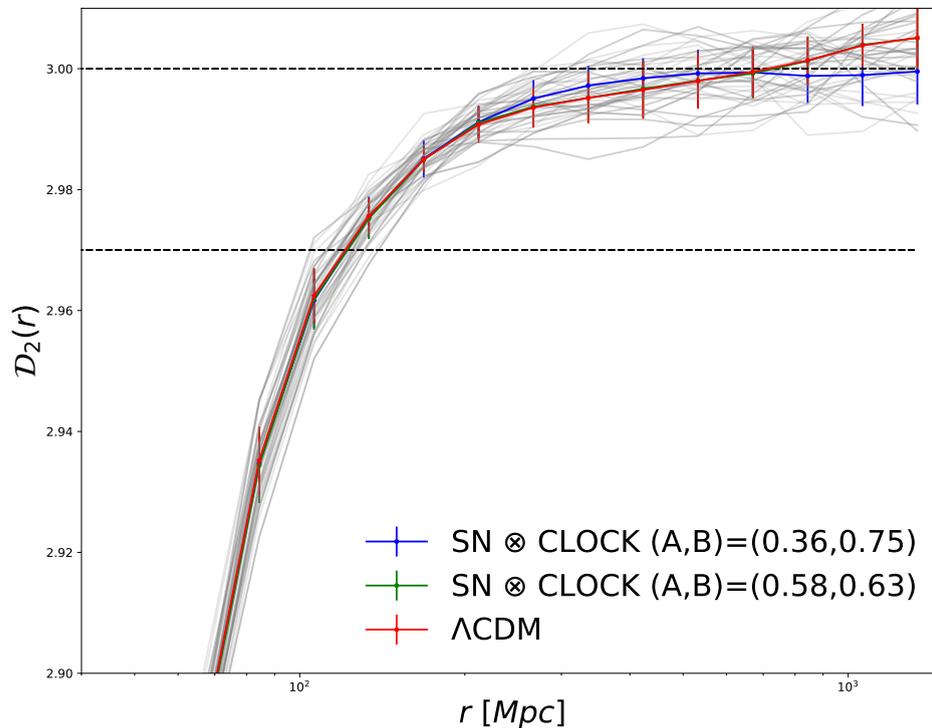


FIGURE 6.4: Fractal Correlation Dimension as a function of scales, $D_2(r)$, for the CMASS galaxy sample using a model independent distance measure and confirming the *Cosmological Principle*. [See text for details]

However, we are still insensitive to a possible redshift evolution of the galaxy sample since in order to correct for geometric effects, we construct a random catalogue that follows the same redshift distribution as the galaxy sample. Therefore in principle one should use a galaxy sample that has a number density which is constant with redshift. This issue is left to be corrected in future work.

Furthermore, this measurement is a rough estimate to the actual measurement. One should in principle measure the age difference of pairs of galaxies of the sample from their spectrum and furthermore measure the magnitude of the supernova located on the surrounding environment of each galaxy to acquire a more accurate estimate of the individual magnitude and therefore a more accurate measurement of the transverse comoving distance. This is left to future work where the ideal sample of galaxies would have contain galaxies which have at least one identified Supernovae on their surroundings.

Chapter 7

Conclusion

In the framework of modern cosmology, we have investigated its main pillars, the Cosmological Principle. This principle states that the universe is homogeneous and isotropic at enough large scales. The Cosmological Principle, not physical motivated due to the large inhomogeneity that we observe on structures, it was first adopted to simplify our knowledge and find simple solution for the dynamics of the universe. In this study, we have used the data release 12 of BOSS CMASS galaxy sample of the to study the transition to cosmic homogeneity over a volume of $5.1 h^{-3} \text{ Gpc}^3$. We use the counts-in-spheres, $N(< r)$, i.e. the average number of objects around a given object, and its logarithmic derivative, the fractal correlation dimension, $D_2(r)$. For a homogeneous sample, $N(< r) \propto r^3$ and $D_2 = 3$. We define a characteristic homogeneity scale, \mathcal{R}_H , as the value for which D_2 reaches the homogeneous value within 1%, i.e. $D_2(\mathcal{R}_H)=2.97$.

For the galaxy distribution, we get $3 - \langle D_2 \rangle = (0.6 \pm 1.3) \times 10^{-3}$ at 1σ over the range $300 h^{-1} \leq r \leq 1300 h^{-1} \text{ Mpc}$, which is consistent with homogeneity assumption of the universe and a transition to homogeneity at a characteristic scale, $\mathcal{R}_H^{(G)} = 114.2 \pm 5.8 h^{-1} \text{ Mpc}$.

Using an alternative definition for the homogeneity scale, i.e. the minimum scale at which the fractal correlation dimension is consistent with the homogeneity value within 1σ error, we estimate a range for the homogeneity scale, $150 \leq \mathcal{R}_H^E \leq 400 h^{-1} \text{ Mpc}$, using 1000 mocks for all redshift bins and galactic caps. This further confirms that the universe transits to homogeneity state in both matter and galaxy distributions. However this estimator has 30% precision which is lower to ours which it is in %-level.

Furthermore, we perform the measurement on the matter distribution in a model dependent way. In particular, we fit the two point correlation function on the CMASS galaxy sample in the range $1 < r < 40 h^{-1} \text{ Mpc}$ to obtain the bias relative to the ΛCDM prediction for the matter correlation function as long as the peculiar velocities to correct for *Redshift Space Distortions*. We correct our measurement of $N(< r)$ for this bias and the peculiar velocities, in order to get the result for the matter distribution, that we finally compare to the ΛCDM prediction.

For the distribution of matter, we achieved a measurement of $3 - \langle \mathcal{D}_2(r > 300 h^{-1} \text{ Mpc}) \rangle_z = (1.7 \pm 1.0) \times 10^{-4}$ at 1σ . We obtained a characteristic scale $\mathcal{R}_H^{(M)} = 61.9 \pm 0.8 h^{-1} \text{ Mpc}$ at an average $z = 0.538 - 0.592$ for the transition to homogeneity. This measurement of \mathcal{R}_H is more precise than the previous measurements of Scrimgeour et al. [93] Collaboration by a factor 5. While Sarkar et al. [178], only give a qualitative measurement of \mathcal{R}_H .

We also investigate the redshift evolution of our observables. We find that the homogeneity scale for the matter distribution, $\mathcal{R}_H^{(M)}$, is decreasing with time as expected if clustering is increasing with time. We find accordance with Λ CDM model with a reduced $\chi^2 = 0.89(0.61)$ for the North(South) Galactic Cap for the 6 redshift bins, i.e. accounting also the measurement of Laurent et al. [166] at redshifts $2.2 \leq z \leq 2.8$. However, the characteristic scale of the galaxy distribution, $\mathcal{R}_H^{(G)}$, remains almost constant with time or redshift, z . The homogeneity scale is an intrinsic property of the galaxy sample. Therefore, this result is consistent with the fact that the CMASS sample has properties that remain constant with time.

Moreover, we performed several test to ensure that the measurement is unbiased against several choices on the performance of the measurement. Firstly, we checked for a potential systematic against alternative modelling of *Redshift Space Distortions*. Furthermore, we ensure that measuring the fractal correlation dimension on several fractal distributions with the survey mask does not bias the intrinsic properties of these distributions, i.e. the fractal dimension of the distribution. Moreover, we perform the basis measurement of the homogeneity scale with different combination of the weighting scheme, observing no significant systematic errors. Furthermore, we check for alternative estimators of the homogeneity scale and we found that the optimum one (most accurate and the 2nd least precise) is the one we developed in this thesis. In a further study we found out that the bias and peculiar velocities, used to measure the homogeneity on the matter distribution are slightly correlated with the characteristic scale, i.e. $r \leq 20\%$ for all possible combination between of the quantities $(\mathcal{R}_H^{(G)}, \mathcal{R}_H^{(M)}, b, \sigma_p)$ at all redshift bins, rendering their correlation negligible. Additionally, we find that using a different definition for the homogeneity scale, i.e. the minimum scale at which the fractal correlation dimension becomes consistent with the nominal value within 0.1%, $\mathcal{D}_2(\mathcal{R}_H^{\mathcal{D}_2=2.997}) = 2.997$, namely per-Mille definition, gives scales which are more than twice the scales obtained with the adopted definition of our basis analysis, $\mathcal{D}_2(\mathcal{R}_H^{\mathcal{D}_2=2.97}) = 2.97$.

Furthermore, we find that the precision of the observables of the fractal correlation dimension and the scaled-counts-in-spheres for the matter distribution gains about 1 order of magnitude in respect of the same observable for the galaxy distribution, due to the bias correction, which is not the case for the 2-point-correlation function that it is being used in regular large scale structure analysis. This renders our observables, counts-in-spheres and fractal correlation dimension a potential precision observables of cosmological inferences for large scale structure cosmography.

We also conclude that this measurement of the homogeneity scale is only a validation of the Cosmological Principle. This is true since we cannot infer distances from observational data, i.e. redshift, without the assuming in prior a Homogeneous universe with an FLRW metric. Therefore future projects should relax the homogeneity assumption and use the same observable and estimators for a model independent measurement of the characteristic scale of the transition to cosmic homogeneity for a confirmation test.

Additionally, another conclusion we draw is that we are unable to perform measurements outside the observable universe. Therefore, we rely on the assumption that the observable universe is a fair sample, i.e. the properties of the observable universe are the same as the unobserved one. However, since the measurements done by Hogg et al. [83] in 2005 have a homogeneous universe at redshift, $z \simeq 0.3$. While in this study we have shown that the galaxies even beyond the observable universe, of 2005, i.e. the observable universe 2017 at redhif $z \simeq 0.56$, the universe has also a homogeneous behaviour. Therefore by induction, we conclude that the unobserved universe will look uniform as well, with a model dependent measurements of course.

We stress that another issue one need to resolve is the fact in these studies, the only tool that we have in our disposal are the galaxy catalogues coming from vast surveys of technologically improved telescopes.

But even with this amazing technological power we still end up with a galaxy catalogue survey, which have holes in them and of course geometrical boundaries. In order to remove the geometrical boundaries we throw in the survey region simulated random points and we compare the galaxy counts to the one with the random catalogue. The unfortunate thing is that we have no tool to mimic independently randomly the redshift distribution, so we end up using random catalogues that their redshift distribution matches the galaxy catalogue. This is due to the fact that we need to mimic target detection efficiency of the telescope (magnitude cuts) which end up being a peculiar function of the radial(redshift) density ¹. A possible solution might come with a survey that manages to measure a galaxy number density that is constant with redshift. This is left to be determined by future work.

Furthermore, the information of redshift is the information we obtain from the photons that travel from each individual source to the observers. This means that each galaxy is at different redshift from one another. Which means they live in different times epochs. So in a sense, one does not have access to the actual spatial distribution of galaxies in the same epoch but on the contrary he has access to a distribution that lies on the past lightcone, a distribution of objects in different times. Therefore, we cannot study directly the 3D comoving spatial distribution, but we rely always on our modelling of the universe. However, passively evolving galaxy distribution (the CMASS sample we discussed in chapter 3.4), gives you access to galaxies that do not evolve in time. This means that their properties do not change in time and they are all of the same type. Then this means that we are using a distribution that is conserve in time. Then ones has access to the actual spatial distribution of galaxies. However, to construct this sample you rely on astrophysical models but these models are not in the same paradigm as Λ CDM-model. Therefore this sample of galaxies is constructed with a model free of the homogeneity assumption. This renders the CMASS sample the best homogeneity assumption free sample that we can use, to date.

Moreover, we have developed a new complementary cosmological probe, i.e. we have used the transition to cosmic homogeneity scale as standard rule to measure cosmological parameters. In particular, we have shown that the measurement of the transition to cosmic homogeneity scale, \mathcal{R}_H is comparable to the BAO scale, \mathcal{R}_{BAO} , using the CMASS galaxy sample with insignificant correlation, $r \sim 20$. We have found that our probe is sensitive to the cosmological parameters, the Hubble constant, H_0 , the total matter ratio density, Ω_m , the Dark energy ratio density, Ω_Λ , and the equation of state as a function of time or redshift, $w(z) = w_0 + w_a \frac{1}{1+z}$, using a Fisher Matrix formalism. An MCMC exploration algorithm of the homogeneity scale probe let us find comparable constrains using the prior knowledge of CMB measurements[6]. In particular, we find that the hubble constant is $H_0 = 68.4 \pm 0.8 \text{ km/s/Mpc}$, the total matter ration density is $\Omega_m = 0.294 \pm 0.009$ and the Dark energy ratio density is $\Omega_\Lambda = 0.709 \pm 0.0010$, while for the equation of state we find, $w_0 = -1.047 \pm 0.035$. Finally, we show that the sensitivity of the homogeneity scale is only a function of the volume of the large scale structures that will be mapped by future experiments. We find that the overall precision of the homogeneity scale for the future projects will be 0.24% with leading experiment the ground based experiment LSST achieving a maximum sensitivity of 0.44% (compared to the BAO scale sensitivity of 0.5%). These results render the homogeneity scale a better probe of cosmology than the BAO scale for photometric surveys.

We have also studied the transition to cosmic homogeneity in a model independent fashion. We have shown, that using an alternative estimation of the radial distance from redshift data, i.e. free of assumptions of homogeneity or isotropy for the distribution in study, the fractal correlation dimension for the galaxy distribution still reaches the homogeneity value, 3, asymptotically. Furthermore, we show that the fractal correlation dimension is compatible with 3 for the galaxy distribution at scales larger than 400Mpc model independently.

¹ See for example the $\bar{n}(z)$ -diagram by Reid et al. [139] (figure 11).

Furthermore, for the matter distribution we need to assume and FLRW model to infer the bias between the galaxy and the matter distribution. Therefore we cannot confirm the Cosmological Principle in a model Independent way. Additionally, using this model independent way for homogeneity, we are still insensitive to a possible redshift evolution of the galaxy sample since in order to correct for geometric effects with the same ways as we do for the model dependent study. Therefore in principle one should use a galaxy sample that has a number density which is constant with redshift. These issue are left to be corrected in future work. To correct for these issues one need to use as surveys whose the number density of the acquired galaxy samples is constant with redshift and the inference of the bias between luminous and matter distribution can be achieve model independently.

Appendix A

Cosmo-Tips

A.1 Cosmography

Hubble distance:

$$d_H = \frac{c}{H_0} \simeq \frac{3 \times 10^8 \text{ m/s}}{7 \times 10^4 \text{ m/s/Mpc}} \simeq 5\text{Gpc} \quad (\text{A.1})$$

Hubble Volume:

$$V_H = d_H^3 \simeq 3\text{Gpc}^3 \quad (\text{A.2})$$

Comoving Distance:

$$d_C(z) = d_H \int_0^z \frac{dz'}{E(z')} \quad (\text{A.3})$$

where $E(z; \vec{\Omega}) = H(z)/H_0$ and $H(z)$ is given by Eq. 1.27.

Transverse Comoving Distance(Motion)

$$d_M(z) = \begin{cases} \frac{d_H}{\sqrt{\Omega_k}} \sinh\left(\sqrt{\Omega_k} \frac{d_C(z)}{d_H}\right), & \Omega_k > 0 \\ d_C(z), & \Omega_k = 0 \\ \frac{d_H}{\sqrt{\Omega_k}} \sin\left(\sqrt{\Omega_k} \frac{d_C(z)}{d_H}\right), & \Omega_k < 0 \end{cases}$$

Angular Diameter Distance $d_A(z) = \frac{s(z)}{\theta} = \frac{\text{size}_{\text{object}}(z)}{(\text{Angular apperture of object})}$:

$$d_A(z) = \frac{d_M(z)}{1+z} \quad (\text{A.4})$$

Luminosity Distance:

$$d_L(z) = (1+z)d_M(z) \quad (\text{A.5})$$

Light-travel Distance:

$$d_T(z) = d_H \int_0^z \frac{dz'}{(1+z')E(z')} \quad (\text{A.6})$$

Age of the Universe:

$$\lim_{z \rightarrow \infty} t(z) = \lim_{z \rightarrow \infty} d_T(z)/c \quad (\text{A.7})$$

Comoving Volume:

$$dV_C = d_H \frac{(1+z)^2 d_A^2(z)}{E(z)} d\Omega dz \quad (\text{A.8})$$

Comoving Volume distance:

$$d_V(z; \mathbf{p}) = [czH^{-1}(z; \mathbf{p})(1+z)^2 d_A^2(z; \mathbf{p})]^{1/3} \quad (\text{A.9})$$

A.1.1 Luminosity Distance revisited

SN can be used as standard candles. By measuring their

- m : color magnitude
- M : assuming an absolute magnitude (which maybe using a bayesian analysis you can assume a gaussian around $M \pm \Delta M$, [205])
- z : redshift (can be infered also from their Host Galaxy)

For object in the neaby universe ($z \ll 1$ in practice $z < 0.02$)

$$d_L(m) = 10^{\frac{m-M}{5}+1} \quad (\text{A.10})$$

We can measure $d_L(m)$ only by astrophysical assumptions about the emitted flux.

For objects that are not nearby then on need to calculate:

$$d_L(z) = \sqrt{\frac{L}{4\pi F}} \quad (\text{A.11})$$

If we are interested in differential Flux, F_ν , and Luminosity, L_ν rather than the bolometric ones the above formula need some corrections due to redshift since the source emits at a different band than the one we observe. Therefore, if we are working with frequencies, ν the correction gives us:

$$d_L(z) = \sqrt{(1+z) \frac{L_{(1+z)\nu}}{L_\nu} \frac{L_\nu}{4\pi F_\nu}} \quad (\text{A.12})$$

while if we are working with wavelengths, λ we use:

$$d_L(z) = \sqrt{\frac{1}{1+z} \frac{L_{(1+z)\lambda}}{L_\lambda} \frac{L_\lambda}{4\pi F_\nu}} \quad (\text{A.13})$$

A.2 Topological Restrictions

A universe may be described by its topological properties and its energy content. Here we are going to describe first the topological properties. Any kind of universe is described by a general metric $g_{\mu\nu}(x)$, defined via:

$$ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu \quad (\text{A.14})$$

In the context of topology, we can assign labels to each comoving observer and use spacetime comoving coordinates on a (1,3)D manifold described by equation:

$$x^\mu = (t, x, y, z) \quad (\text{A.15})$$

The general metric of a comoving observer takes the form of:

$$ds^2 = -c(x)^2 dt^2 + 2g_{0i}(x) dx^i dt + g_{ij}(x) dx^i dx^j \quad (\text{A.16})$$

In Concordance Cosmology, we need to respect the geometrical properties arised from the construction of the Cosmological Model which is the Cosmological Principle. The properties, which arised from this construction, give the following topological restrictions on the construction of the metric: e.i. spatial stationary, non-accelerating comoving observer and isotropical local expansion.

A.2.1 Spatial stationary comoving observer

By assuming isotropy the comoving observer's 4-velocity u must be perpendicular on the hypersurface Σ_t . Since the position of the observer (x^i) is fixed, we have:

$$u^\alpha = (c^{-1}[x], 0, 0, 0) \quad (\text{A.17})$$

which is the first component of normalisation. The 4-velocity of the observer is perpendicular to every 4-vector v , which is tangent to Σ_t , e.i. $v^0 = 0$. The generalized dot product, reduces to:

$$u \cdot v = g_{\alpha\beta} u^\alpha v^\beta = g_{0i} v^i c^{-1}(x) \quad (\text{A.18})$$

Since this product should be 0 for every tangent vector v , we have that $g_{0i} = g^{0i} = 0$ and g^{ij} is the inverse of g_{ij} ¹.

A.2.2 Spatial non-accelerating comoving observer

The spatial component of the acceleration of the comoving observer ($a^\mu = u^\nu \nabla_\nu u^\mu$) can be written as:

$$a_i = \partial_i \ln c(x) \quad (\text{A.19})$$

Since the 4-acceleration a is measurable from the comoving observer, $a_i = 0$ (no preferred spatial direction). From Eq. A.19, we infer:

$$\partial_i c(x) = 0 \rightarrow c = c(t) \quad (\text{A.20})$$

We may redefine the time coordinate as:

$$t'(t) = c \int c(t) dt \quad (\text{A.21})$$

The new metric becomes:

$$ds^2 = -c^2 dt'^2 + g_{ij}(x) dx^i dx^j \quad (\text{A.22})$$

For simplicity, we through the prime on our calculations. In these coordinates $u^\alpha = (1, 0, 0, 0)$ and $\mathbf{a} = 0$, which implies spacetime-non-accelerating observer.

A.2.3 Isotropical local expansion

For any purely spatial vector v ² that is normalised ($v^\nu v_\nu = 1$), we may construct the following quantity with a general dependence on time, location and direction:

$$H(t, x^i, v) = v^\nu v_\mu \nabla_\nu u^\mu \quad (\text{A.23})$$

¹N.B. $c(x)=0$ would imply a singular coordinate system.

²spatial in the comoving frame, i.e. $u \cdot v = 0$

which is a component of the gradient of the 4-velocity. This function can only depend on time (t). Since by the previous restriction $v^0 = v_0 = 0$, we have:

$$H(t, x^i, v) = v^i v_j g^{jk} \frac{1}{2} \partial_t g_{ik} = \frac{1}{2} v^i v^k \partial_t g_{ik} \quad (\text{A.24})$$

If this quantity satisfy the normalization condition $v^i v^k g_{ik} = 1$, $H(t, x^i, v)$ is independent of the direction (v). Thus, the 3×3 symmetric matrix $\partial_t g_{ik}$ must be a scalar multiple of g_{ik} :

$$\partial_t g_{ik} = I g_{ik} . \quad (\text{A.25})$$

Inspection gives $H = I/2$ and:

$$\partial_t g_{ik} = 2H(t) g_{ik} \quad (\text{A.26})$$

We now may define the *scale factor* via:

$$a(t) = \exp \left[\int H(t) dt \right] \quad (\text{A.27})$$

and by integrating Eq. A.26:

$$g_{ik}(t, x^j) = \gamma_{ik}(x^j) \exp \left[2 \int H(t) dt \right] = \gamma_{ik}(x^j) a^2(t) \quad (\text{A.28})$$

we may reform the metric defined by Eq. A.16 to the Robertson-Walker (RW) metric:

$$ds^2 = -c^2 dt^2 + a^2(t) \gamma_{ij}(\vec{x}) dx^i dx^j \quad (\text{A.29})$$

A.2.4 Selection of the 3D spatial metric

Topologically, there are several options for a homogeneous isotropic 3-manifold defined via the 3-metric $\gamma_{ij}(\vec{x})$. Thus generally, we may describe an isotropic 3-manifold by:

$$ds_3^2 = d\chi^2 + f(\chi) d\Omega \quad (\text{A.30})$$

where $d\Omega = d\theta^2 + \sin^2\theta d\phi^2$ reflects the isotropy of the spatial comoving space. All possible homogeneous 3 manifolds may be parametrize the functional $f(\chi)$, via the k -parameter which is an intrinsic geometrical property³, and allow for a possible curvature-variation of the 3D-spatial-space as it is represented in Fig. 1.1.

Thus, assuming the Cosmological Principle and a possible 3D-spatial-space curvature, Eq. A.30 takes the form of an Friedmann Lemaître Robertson Walker (FLRW) metric:

$$ds^2 = -c^2 dt^2 + a^2(t) \left[\frac{1}{1 - kr^2} dr^2 + r^2 d\Omega \right] \quad (\text{A.31})$$

Note, that this is a construction of a comoving coordinate system only by considering geometrical properties.

³Intrinsic geometrical properties are geometrical properties that can be measured by measuring quantities inside the same dimensional geometrical space

A.3 General n-point Correlation function

Large scale structures are currently being investigated with the statistical tool dedicated to study the correlation of the structures which is a manifestation of number density. If one would like to study the correlations of a spatial distribution, independent of any assumption of homogeneity and isotropy or a model describing the distribution, he could build estimators based on the definition of the correlation function described by equation n-point-correlation function.

In liquid theory[206], it became clear that *the presence of one molecule in a region increased the probability of finding another nearby*. To express this, consider a volume of space containing N galaxies (or molecules or, generally, particles) labelled each by $1, 2, 3, \dots, N$. Consider the probability that galaxy 1 is in a specified small region $d\vec{r}_1$ and galaxy 2 in $d\vec{r}_2$ and so on for n of the N galaxies by $P^{(n)}(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n) d\vec{r}_1 d\vec{r}_2 \dots d\vec{r}_n$ and whatever the configuration of the remaining $N - n$ galaxies. Since all n galaxies must be somewhere in the total volume V , the normalization for each $P^{(n)}$ is:

$$\int_{V_1} \int_{V_2} \dots \int_{V_n} P^{(n)}(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n) d\vec{r}_1 d\vec{r}_2 \dots d\vec{r}_n = 1 \quad (\text{A.32})$$

Next, suppose we wish to designate the probability that there are n galaxies in the volume elements $d\vec{r}_1 d\vec{r}_2 \dots d\vec{r}_n$ without specifying which galaxy is in which volume element. Then any of the N galaxies could be in $d\vec{r}_1$, any of $N - 1$ in $d\vec{r}_2$, and any of $N - n + 1$ in $d\vec{r}_n$, giving a total of

$$N(N - 1) \dots (N - n + 1) = \frac{N!}{(N - n)!} \quad (\text{A.33})$$

possibilities and the general distribution function is:

$$\rho^{(n)}(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n) = \frac{N!}{(N - n)!} P^{(n)}(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n) \quad (\text{A.34})$$

whose normalisation turns out to be:

$$\int_V \dots \int_V \rho^{(n)}(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n) d\vec{r}_1 d\vec{r}_2 \dots d\vec{r}_n = \frac{N!}{(N - n)!} \quad (\text{A.35})$$

from equation (A.32).

To define the correlation function, we note that if the distribution is completely uncorrelated (analogous to a perfect gas), the probabilities of finding galaxies in different volume elements are independent and therefore the multiplication is giving:

$$\rho_{uncorr}^{(n)}(\vec{r}_1, \dots, \vec{r}_n) = \rho^{(1)}(\vec{r}_1) \cdot \dots \cdot \rho^{(1)}(\vec{r}_n) \quad (\text{A.36})$$

where $\rho^{(1)}(\vec{r}_i)$ is just the probability that some galaxy is in $d\vec{r}_i$ at \vec{r}_i . But any correlations will introduce modifications, which we can represent by writing more generally, as:

$$\rho_{corr}^{(n)}(\vec{r}_1, \dots, \vec{r}_n) = \rho^{(1)}(\vec{r}_1) \cdot \dots \cdot \rho^{(1)}(\vec{r}_n) \left[1 + \xi^{(n)}(\vec{r}_1, \dots, \vec{r}_n) \right] \quad (\text{A.37})$$

Up to this point no assumption about the homogeneity or isotropy was made [157]. So the n-point correlation function is defined via:

$$dP^{(n)}(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n) = \rho^{(1)}(\vec{r}_1) \dots \rho^{(1)}(\vec{r}_n) \left[1 + \xi^{(n)}(\vec{r}_1, \dots, \vec{r}_n) \right] d\vec{r}_1 d\vec{r}_2 \dots d\vec{r}_n \quad (\text{A.38})$$

A.4 CMASS Effective Volume Calculation

For a survey at redshift z with volume $V(z)$ the effective volume is given by:

$$V_{\text{eff}} = \sum_z \Delta V(z) \left(\frac{n(z)P}{1 + n(z)P} \right)^2 \quad (\text{A.39})$$

where $n(z)$ is the sample density at redshift z and P is the power spectrum calibration value of the survey. For the CMASS galaxy sample the *ratio of true clustering power to that from shot-noise* is given by $n(z_{\text{cmass}})P = 3$ [139]. This quantity is calculated in the following manner. We have: Total CMASS galaxies = NGC + SGC = (571379 + 215571)gal = 786950gal within a volume of $V_{\text{survey}} = 5.1h^{-3}\text{Gpc}^3$, assuming Planck-2015[6]. This gives a density of $n(z_{\text{cmass}}) = 154303h^3\text{Gpc}^{-3} \approx 1.5 \times 10^5 \times 10^{-9}h^3\text{Mpc}^{-3} \approx 1.5 \times 10^{-4}h^3\text{Mpc}^{-3}$. The calibration Power Spectrum value is $P = 20000h^3\text{Mpc}^{-3}$ corresponding to the observed power spectrum at $k \approx 0.15h\text{Mpc}^{-1}$ given at 6.7 section of [139]. Therefore:

$$V_{\text{eff}}^{\text{cmass}} = 5.1h^{-3}\text{Gpc}^3 * \left(\frac{3}{1+3} \right)^2 = 2.9h^{-3}\text{Gpc}^3 \quad (\text{A.40})$$

A.5 Cosmic Variance vs Shot Noise Domination

A shot-noise dominated sample is a sample in which enough modes have enter the survey volume, but they are purely measured since we have very low density of objects. While in the case of cosmic variance dominated sample because the density of objects is large enough, we can measure the modes. The test to define such a sample is given through the equation:

$$nP = \left\{ \begin{array}{l} \simeq 1 \text{ Cosmic Variance dominated} \\ \ll 1 \text{ Shot noise dominated} \end{array} \right\}$$

A.6 Systematic Error from the underlying Modelling

Example: Imagine we measure the Homogeneity Scale with 2 different Methods. First method is by reconstructing the \mathcal{D}_2 observable measuring the bias by only applying the Kaiser (K) at scales $r = [15, 40] h^{-1}\text{Mpc}$ and then by applying the Kaiser plus Finger of God ($K + F$) models at scales $r = [1, 40] h^{-1}\text{Mpc}$, by marginalising over the rest of the parameters of the model of course. The output measurements are following values:

- $R_H^K = 63.2 \pm 0.8h^{-1}\text{Mpc}$
- $R_H^{K+F} = 62.6 \pm 0.8h^{-1}\text{Mpc}$

then we have statistical error (random error) $\sigma_{R_H}^{\text{stat}} = 0.8h^{-1}\text{Mpc} = \delta R^K = \delta R^{K+F}$. However the systematic error arise from the fact that there is a bias between the 2 models. So the systematic error is about $\sigma_{R_H}^{\text{sys}} = |R_H^K - R_H^{K+F}| = 0.6h^{-1}\text{Mpc}$. Then we say is that *the contribution from the Finger of God modelling is less than the statistical error*. So there is a systematic difference between the two methods which is less than the statistical error. So we dominated by statistics. It is a systematic since we use different scales to apply the 2 models.

From the experimental point of view, the systematic error is, generally, thought of as the error generated from the instrument or detection methods malfunctions. Usually, they are treated with calibrations to known values of the parameter space. These results into weights that you need to take into account on to your observable.

A.7 $\mathcal{D}_2(r)$ - Error Propagation from s.s.s. to l.s.s.

While Linear theory describes very accurately the large scale structures (*l.s.s.* : $r > 40h^{-1}\text{Mpc}$), it fails to describe the small scale structures (*s.s.s.* : $r < 40h^{-1}\text{Mpc}$). This means that the estimation of the 2-point correlation function $\xi(r)$ is not accurately described at *s.s.s.*. Our Fractal Correlation Dimension estimator of the structures of the universe $\mathcal{D}_2(r)$ is an radius-derivative of an integral over all the possible angles of the 2-point correlation function. This means that before integration we multiply the 2-point correlation function with the function s^2 . This will provide a small weighing at small scales and a large weigh on larger scale. As a result, the alteration of the description of structures through the 2-point correlation function estimator at *s.s.s.* will not alter the description of structures using the Fractal Correlation dimension estimator at *l.s.s.*

A.8 *bias*, a gain for \mathcal{D}_2^M

Assuming a scale independent bias, we have the relationship of the correlation function between the matter distribution and the galaxy distribution at any scale r as:

$$\xi^G(r) = b^2 \xi^M(r) \quad (\text{A.41})$$

The variance of the correlation functions are related through error propagation as:

$$\sigma_{\xi^G}^2 = (b^2 \sigma_{\xi^M})^2 + (\xi^M - 2 b \sigma_b)^2 \quad (\text{A.42})$$

Since $\xi^M(r) \simeq 0$ at large r , we have that the variances are related by:

$$\sigma_{\xi^G}(r) \simeq b^2 \sigma_{\xi^M}(r) \quad (\text{A.43})$$

Then the ξ -precision is similar for matter and galaxy distribution at any scale:

$$\frac{\sigma_{\xi^G}}{\xi^G}(r) \simeq \frac{\sigma_{\xi^M}}{\xi^M}(r) \quad (\text{A.44})$$

This means that the precision on the measurement of the correlation function on the matter distribution is not affected by the bias. While for the scaled counts in spheres and Fractal correlation dimension, we have a different picture on the relation of the precision. For the scaled count in spheres from:

$$\mathcal{N}(< r) = 1 + \frac{3}{r^3} \int_0^r \xi(s) s^2 ds \quad (\text{A.45})$$

we can have:

$$\mathcal{N}^M(< r) = \frac{\mathcal{N}^G(< r) - 1}{b^2} + 1 \quad (\text{A.46})$$

With simple algebra this relation can be rewritten for any scale r :

$$\mathcal{N}^M = \frac{1}{b^2} \mathcal{N}^G + \left(1 - \frac{1}{b^2}\right) \quad (\text{A.47})$$

With error propagation from Eq. A.46 we get:

$$\sigma_{\mathcal{N}^M}^2 = \left(\frac{1}{b^2} \sigma_{\mathcal{N}^G} \right)^2 + \left(-2 \frac{\mathcal{N}^G - 1}{b^3} \sigma_b \right)^2 \quad (\text{A.48})$$

Since $\mathcal{N}^M(< r) - 1 \simeq 0$ we have:

$$\sigma_{\mathcal{N}^M} \simeq \frac{1}{b^2} \sigma_{\mathcal{N}^G} \quad (\text{A.49})$$

Then by dividing Eq. A.49 with Eq. A.47 we have that the precision is written:

$$\frac{\sigma_{\mathcal{N}^M}}{\mathcal{N}^M} = \frac{\frac{1}{b^2} \sigma_{\mathcal{N}^G}}{\frac{1}{b^2} \mathcal{N}^G + \frac{b^2-1}{b^2}} = \frac{\sigma_{\mathcal{N}^G}}{\mathcal{N}^G + (b^2 - 1)} \quad (\text{A.50})$$

This means that the \mathcal{N} -precision between galaxy and matter distribution for large scales is given by:

$$\frac{\sigma_{\mathcal{N}^M}}{\mathcal{N}^M}(r) \simeq \frac{\sigma_{\mathcal{N}^G}(r)}{\mathcal{N}^G(r) + (b^2 - 1)} \quad (\text{A.51})$$

There is a gain for a biased tracer, while for an unbiased tracer there is no gain ($b^2 = 1$).

Similar we have for the Fractal Dimension Precision! We consider that:

$$\mathcal{D}_2(r) = 3 + \frac{\partial \ln}{\partial \ln r} \left[1 + \frac{3}{r^3} \int_0^r \xi(s) s^2 ds \right] \quad (\text{A.52})$$

Or simply:

$$\mathcal{D}_2(r) - 3 = \frac{\partial \ln \mathcal{N}(< r)}{\partial \ln r} \quad (\text{A.53})$$

It follows that the relationship between matter and galaxy, at any scale r , can be given by:

$$\mathcal{D}_2^M - 3 = \frac{\partial \ln}{\partial \ln r} [\mathcal{N}^M - 1] = \frac{\partial \ln}{\partial \ln r} \left[\frac{1}{b^2} \mathcal{N}^G - 1 \right] = \frac{1}{b^2} \frac{\partial \ln}{\partial \ln r} \mathcal{N}^M \quad (\text{A.54})$$

which leads to:

$$\mathcal{D}_2^M - 3 = \frac{1}{b^2} (\mathcal{D}_2^G - 3) \quad (\text{A.55})$$

This then can be rewritten with simple algebra as:

$$\mathcal{D}_2^M(r) = \frac{1}{b^2} \mathcal{D}_2^G(r) + \frac{3(b^2 - 1)}{b^2} \quad (\text{A.56})$$

Simple error propagation from Eq. A.55 leads to:

$$\sigma_{\mathcal{D}_2^M}^2 = \left(\frac{1}{b^2} \sigma_{\mathcal{D}_2^G} \right)^2 + \left(\frac{-2}{b^3} (\mathcal{D}_2^G - 3) \sigma_b \right)^2 \quad (\text{A.57})$$

Then since $\mathcal{D}_2^G(r) - 3 \simeq 0$ for large r only the first term survives and we get:

$$\sigma_{\mathcal{D}_2^M} = \frac{1}{b^2} \sigma_{\mathcal{D}_2^G} \quad (\text{A.58})$$

By dividing Eq. A.59 with Eq. A.47 we end up with:

$$\frac{\sigma_{\mathcal{D}_2^M}}{\mathcal{D}_2^M} = \frac{\frac{1}{b^2} \sigma_{\mathcal{D}_2^G}}{\frac{1}{b^2} \mathcal{D}_2^G + 3 \frac{(b^2-1)}{b^2}} \quad (\text{A.59})$$

Then the precision for the Fractal Correlation dimension will be given by:

$$\frac{\sigma_{\mathcal{D}_2^M}}{\mathcal{D}_2^M}(r) = \frac{\sigma_{\mathcal{D}_2^G}(r)}{\mathcal{D}_2^G(r) + 3(b^2 - 1)} \quad (\text{A.60})$$

The Eq. A.60 and Eq. A.51 tell us that the precision of measuring the observable from a biased tracer is gained from the individual bias.

If we measure the bias from the same data then that is a recursive technic. Although if we measure the bias from low scale data and use it to reconstruct our observable on higher scales this is independent measure of bias and we can reconstruct our observable to benefit from its gain. Furthermore, if we could predict the bias from the theory then the use of bias would reconstruct the observable and make a gain as well. Finally, since for the measurement of the homogeneity scale \mathcal{R}_H^O at 1% we propagate the error from the observables \mathcal{D}_2 and \mathcal{N} then the Homogeneity scale is benefit by the same gain from the bias reconstruction. However, we find slightly different precision relation for the homogeneity scale since the $\mathcal{R}_H^{M, \mathcal{D}_2=2.97}$ and $\mathcal{R}_H^{G, \mathcal{D}_2=2.97}$ measurements as they are performed at different scales.

A.9 Alternative homogeneity observables

Several authors have proposed to study the underlying matter distribution of galaxy samples with different observables, but we do not investigated them in this thesis. In particular, Fractal Correlation dimension $\mathcal{D}_2(r)$ is a special case of the generalised fractal dimension. Thus there are alternative observables to use to study a distribution of galaxies such as the generalised fractal correlation dimension [207]:

$$D_q(r) = \frac{1}{q-1} \frac{d \log}{d \log r} \left[\frac{1}{MN} \sum_{i=1}^M [n_i(< r)]^{q-1} \right] \quad (\text{A.61})$$

where $n_i(< r) = \sum_{j=1}^N \Theta(r - |\vec{x}_i - \vec{x}_j|)$ for each i -galaxy as a center is the number density defined via the heaviside function $\Theta(X)$. Another special case of the generalised fractal correlation dimension is the shannon-entropy observable (Information Dimension (D_1)) :

$$H(X) = - \sum_{i=1}^n p(x_i) \log p(x_i) \quad (\text{A.62})$$

where $p(x_i)$ is the probability distribution of the random variable X . The estimators of this observable gives consistent result with Λ CDM N-body simulations [208]. Furthermore, another observable can be defined using the anomalous diffusion (a.d.) of particles [209], which can be defined via:

$$\sqrt{\langle r^2 \rangle} \propto t^{\alpha/2} \quad (\text{A.63})$$

where t is the number of steps of a random-walker walking on the centers of the galaxy sample and $\alpha^{a.d.} = 1$ for an normal diffusion. Therefore, a thorough investigation is needed to study which is the optimal estimator for the measurement of the homogeneity scale measurement. However, as we discussed in chapter 4, we observed that among the estimators that are related to the fractal correlation dimension the most optimum and simplest one is the cor-estimator, i.e. $\mathcal{D}_2(r) = 3 + \frac{d}{d-1} \frac{\ln}{r} \left(1 + \frac{1}{r^3} \int_0^r \xi_{ls}(s) s^2 ds \right)$, where ξ_{ls} is the usual *Landy & Szalay* estimator of the two point correlation function.

A.10 Gravitational Waves

The detection of the primordial black holes at the very late universe ($z = 0.09_{-0.04}^{+0.03}$) was performed for the first time from our best current Gravitational Wave Observatory LIGO [210].

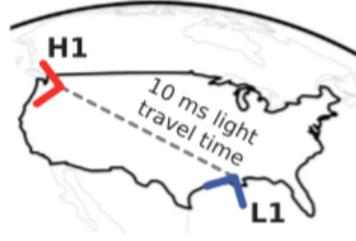
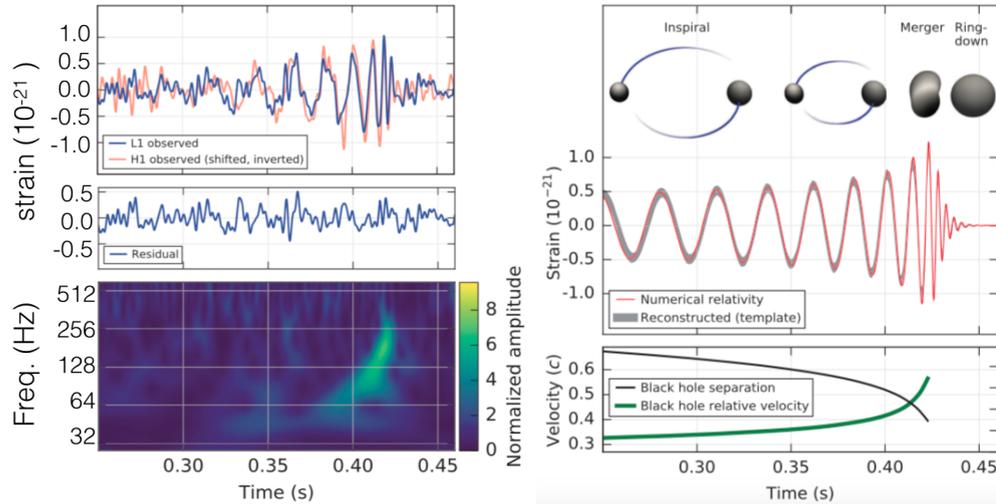


FIGURE A.1: (LIGO) Laser Interferometer Gravitational Wave Observatory.

The detection was achieved with a modified Michelson interferometer where the Gravitational Wave (GW) strain was measured as a difference in lengths of each orthogonal arms. The two arms are formed by two mirrors, acting as test masses, separated by $L_x = L_y = L = 4\text{km}$. Then a passing GW can alter the arm lengths by a distance $\Delta L(t) = \delta L_x - \delta L_y = \eta(t)L$, where $\eta(t)$ is the GW strain amplitude as a function of time t projected onto the detector. The detection was achieved in both detectors, located in Hanford Washington (H1) and Livingston Louisiana (L1) as shown in figure Fig. A.1 on September 14 2015 at 09:50:45 UTC, in less than 10 *msec* apart.

The left part of figure Fig. A.2 represents the detected signal. This is the strain amplitude as a function of time. One can observe that the signal becomes more apparent a few parts of seconds after the triggering mechanism where the residuals are less apparent than the signal. After 0.35 seconds of the triggering time, the signal-to-noise ratio is increasing significantly ~ 2 . The bottom of figure Fig. A.2, shows the time-frequency representation of the strain data, expressing the signal frequency increasing over time. That was the famous detection of the GW!

FIGURE A.2: *Left:* Gravitational Wave (GW) Signal as a function of time *Right:* Physical interpretation of the GW signal, which correspond to a coalescence of two Black Holes

This current observational picture matches the predicted signal from numerical relativity which is shown in the left part of figure Fig. A.2. This figure shows the phases of the two black hole coalescence effect of the two massive black holes of masses $\vec{M} = (36^{+5}_{-4}, 29^{+4}_{-4})M_{\odot}$ as a function of time at 1.5 ± 0.6 billion light years away from earth (which is the most precise but yet high fairly uncertainty). In the coalescence effect there is the inspiral phase where the binary system performs an inspiral trajectory for about 0.15 seconds. Then for about 0.02 seconds, there is the merger phase where the GW emission peaks. Finally the newborn system of $62M_{\odot}$ mass performs a ringdown phase for about 0.01 seconds where there exists still some GW emission until the final stable phase. The bottom panel quantifies the extreme velocities

(0.5c) that the binary system exhibits during the coalescence phenomenon. Credits should be given to both LIGO and VIRGO [211] collaborations for the tremendous effort that they put to analyse these peculiar and beautiful detection pipeline.

A.11 Fisherology

Assuming that we have a measurement $y(b)$ of an observable $o(b; \vec{p})$ (where b are the bins and \vec{p} are the dependence of the parameters) which is random Gaussian variable centred at each true value:

$$y(b) = o(b; \vec{p}) + n(b) \quad (\text{A.64})$$

with

$$\begin{aligned} \langle y(b) \rangle &= o(b; \vec{p}) \\ \langle n(b) \rangle &= 0 \\ \langle y^2(b) \rangle &= \sigma_{o(\vec{p})}^2(b) \\ \langle n^2(b) \rangle &= \sigma_{o(\vec{p})}^2(b) \end{aligned}$$

Then the propability of the measurement is:

$$P(\vec{n}|\vec{p}) \propto \prod_b \exp\left(-\frac{n^2(b)}{2\sigma_{o(\vec{p})}^2(b)}\right) \quad (\text{A.65})$$

which is rewritten as:

$$P(\vec{n}|\vec{p}) \propto \exp\left(-\frac{1}{2} \sum_b \left(\frac{y(b) - o(b; \vec{p})}{\sigma_{o(\vec{p})}(b)}\right)^2\right) \quad (\text{A.66})$$

We can maximize the value:

$$\chi^2(\vec{p}) = \sum_b \left(\frac{y(b) - o(b; \vec{p})}{\sigma_{o(\vec{p})}(b)}\right)^2 \quad (\text{A.67})$$

However, using Taylor theorem we can expand χ^2 around the true values, $\vec{p} = \vec{p}_0 + \delta\vec{p}$, and average over realisations of the data we have:

$$\langle \chi^2(\vec{p}) \rangle = \chi^2(\vec{p}_0) + \sum_i \delta p_i \langle \frac{\partial \chi^2}{\partial p_i} \rangle + \frac{1}{2} \sum_{i,j} \delta p_i \delta p_j \langle \frac{\partial^2 \chi^2}{\partial p_i \partial p_j} \rangle + \mathcal{O}(4) \quad (\text{A.68})$$

if we substitute Eq. A.67 into the taylor expansion we have for the distribution of the error, $n(b)$ in the measured parameters is thus in the limit of high statistics (i.e. $\langle \frac{\partial \chi^2}{\partial p_i} \rangle = 0$):

$$\begin{aligned} P(p_i, p_j) &\propto \exp\left(-\frac{1}{2} \langle \chi^2(\vec{p}) \rangle\right) \\ &\propto \exp\left(-\frac{1}{4} \delta p_i \delta p_j \langle \frac{\partial^2 \chi^2}{\partial p_i \partial p_j} \rangle\right) \\ &\propto \exp\left(-\frac{1}{4} \delta p_i \delta p_j F_{ij}\right) \end{aligned}$$

where we have defined the *Fisher matrix*, F_{ij} , as:

$$F_{ij} = \sum_b \frac{1}{\sigma^2(b)} \frac{\partial o(b)}{\partial p_i} \frac{\partial o(b)}{\partial p_j} \quad (\text{A.69})$$

Note that we have assumed that $\sigma_o(b; \vec{p}) = \sigma_o(b)$. The inverse of the fisher matrix gives us the covariance matrix of the parameters

$$C_{ij} = \langle \delta p_i \delta p_j \rangle = (F^{-1})_{ij} \quad (\text{A.70})$$

Therefore, the Fisher Matrix quantifies the Maximum Information that an observable can give with estimated error $\sigma(b)$. Notice also, that this definition of the Fisher Matrix is only valid for errors that do not depend on the parameters. In the more general case:

$$F_{ij} = \sum_{bb'} \frac{\partial o_b}{\partial p_i} C_{bb'}^{-1} \frac{\partial o_{b'}}{\partial p_j} + \frac{1}{2} \text{trace} \left(C_{bb'}^{-1} \frac{\partial C_{bb'}^{-1}}{\partial p_i} C_{bb'}^{-1} \frac{\partial C_{bb'}^{-1}}{\partial p_j} \right) \quad (\text{A.71})$$

Furthermore, If one can construct a probability of the parameters of the model given a set of observational data, $P(\vec{p}|y(b))$, e.g. with a Bayesian method, then the Fisher Matrix Components are given by:

$$F_{ij} = - \left\langle \frac{\partial^2 \ln P(\vec{p}|y(b))}{\partial p_i \partial p_j} \right\rangle \quad (\text{A.72})$$

The *Cramer-Rao bound theorem* suggest that the any unbiased estimator for the parameters delivers a covariance matrix of parameter that is no better than the inverse of the Fisher Matrix, thus Fisher Matrix has the ability to estimate parameters given a set of observations[189]. Fisher Matrix has some useful properties:

1. *Easy parameter-space change*: Under param-change: $\vec{q} = M\vec{p}$ we have simply $F_q = M^T F_p M$, where M is the transformation matrix.
2. *Additive property* of 2 different datasets A and B: $F_{A+B} = F_A + F_B$
3. *Easy marginalisation*: Remove the non-interesting ("nuisance") parameters by removing the corresponding columns of F^{-1} and reinverse to obtain the reduced Fisher matrix.

A.12 Numerical Error propagation on interpolated point

Consider that we estimate a function ($y_d(x_d)$ at bin d through a polynomial function(or spline function[212] or other):

$$y_{pol} = \sum_{i=0}^{N-1} a_i x^i = F(x; \vec{a}) \quad (\text{A.73})$$

This results in the estimated parameters \vec{a} and the covariance, $C(a_i, a_j)$. Then at a point which is between the bins $x_h \neq x_d$ we have that:

$$y(x_h) \equiv y_h \simeq a_0 x^0 + a_1 x^1 + a_2 x^2 + a_3 x^3 + \dots \quad (\text{A.74})$$

Then perturbing the y in respect of each parameters a_i at position $x_d = x_h$ we have:

$$y_{h,0} = (a_0 + \delta a_0) + a_1 x_h^1 + a_2 x_h^2 + \dots \quad (\text{A.75})$$

$$y_{h,1} = a_0 + (a_1 + \delta a_1) x_h^1 + a_2 x_h^2 + \dots \quad (\text{A.76})$$

$$\dots = \dots \quad (\text{A.77})$$

$$y_{h,N} = a_0 + a_1 x_h^1 + a_2 x_h^2 + \dots + (a_N + \delta a_N) x_h^N \quad (\text{A.78})$$

This allows us to define the partial derivative at position x_h as:

$$\frac{\partial y_h}{\partial a_0} \Big|_{x_d=x_h} = \frac{y_{h,0} - y_h}{\delta a_0} = \frac{\delta a_0}{\delta a_0} = 1 \quad (\text{A.79})$$

$$\frac{\partial y_h}{\partial a_1} \Big|_{x_d=x_h} = \frac{y_{h,1} - y_h}{\delta a_1} = \frac{\delta a_1 x_h^1}{\delta a_1} = x_h \quad (\text{A.80})$$

$$\dots = \dots \quad (\text{A.81})$$

$$\frac{\partial y_h}{\partial a_N} \Big|_{x_d=x_h} = \frac{y_{h,N} - y_h}{\delta a_N} = \frac{\delta a_N x_h^N}{\delta a_N} = x_h^N \quad (\text{A.82})$$

In compact form this is written as:

$$\frac{\partial y_h}{\partial \vec{a}} \Big|_{x_d=x_h} = \left(\frac{\partial y_h}{\partial a_0}, \dots, \frac{\partial y_h}{\partial a_N} \right) \quad (\text{A.83})$$

Therefor the error of y at the intermidiate position $x_d = x_h$ is given by:

$$\sigma_y^2(x_d) = \sum_{i,j=1}^N \frac{\partial y_h}{\partial a_i} \Big|_{x_d=x_h} \cdot C_{a_i, a_j} \cdot \frac{\partial y_h}{\partial a_j} \Big|_{x_d=x_h} \quad (\text{A.84})$$

If we need to transfer the function of the y -axis to the x -axis at x_h , we only need to estimate the slope of the function at position x_h with the estimated parameters, $\tilde{\vec{a}}$:

$$\frac{dy_h}{dx} \Big|_{x_d=x_h} = \frac{d}{dx} \left[F(x, \tilde{\vec{a}}) \right] \Big|_{x=x_d} \quad (\text{A.85})$$

Then the error on the x axis at $x_d = x_h$ is given by:

$$\delta x_h = \frac{\sigma_y}{\frac{dy_h}{dx} \Big|_{x_d=x_h}} \quad (\text{A.86})$$

In our analysis, we consider a small perturbation at all parameters of the spline, i.e. $a_i = 0.01$ for every $i \in \mathbb{N}$.

A.13 Why locally we observe Minkowski space while globally we feel expansion?

Any observer can sense global expansion since he/she reads:

$$ds^2 = -dt^2 + a^2(t)dx^2 \quad (\text{A.87})$$

But due to the fact that Locally we have spacetime curvature through:

$$G_{\mu\nu} \propto T_{\mu\nu}(\Lambda, m) \quad (\text{A.88})$$

one can make a change of variables:

$$ds^2 = -dt'^2 + dx'^2 \quad (\text{A.89})$$

which looks Minkowski. So 2 objects that are locally bound by the gravitational force seem to stay together (because we read from the transformed spacetime (Minkowski)), but 2 objects that are far are always going away from each other as can be seen by Doux [213].

A.14 Can we use ratios as an \mathcal{N} -Estimator?: No

Why don't we use :

$$\mathcal{N}(< r) = \sum_{i=i}^{N_R=NG} \frac{DD_i \frac{1}{N_G(N_G-1)}}{DR_i \frac{1}{N_R(N_R-1)}} \quad (\text{A.90})$$

or the Variance of the ratios DD_i/DR_i . Consider the following example from Ogiore et al. [214].

When we have two Random Variables X, Y : and we want to estimate the $R = \frac{\bar{Y}}{\bar{X}}$ there are 2 possible easy ways⁴:

$$r_1 = \frac{\frac{1}{n} \sum_{i=0}^{\infty} y_i}{\frac{1}{n} \sum_{i=0}^{\infty} x_i} \simeq R \left[1 + \frac{1}{n\bar{X}} + \frac{1}{n^2\bar{X}^2} \right] \quad (\text{A.91})$$

or

$$r_2 = \frac{1}{n} \sum_{i=0}^{\infty} \frac{y_i}{x_i} \simeq R \left[1 + \frac{1}{\bar{X}} + \frac{1}{\bar{X}^2} - \frac{3Cov[X, Y]}{\bar{X}\bar{Y}} \right] \sim R \left[1 + \frac{1}{\bar{X}} + \frac{1}{\bar{X}^2} \right] \quad (\text{A.92})$$

. Note that r_1 is n times less biased than the r_2 in first order. The variance of r_1 is given by:

$$V[r_1] = E[(r_1 - E[r_1])] \simeq R^2 \left[\frac{1}{n} \left(\frac{1}{\bar{X}} + \frac{1}{\bar{Y}} \right) + \frac{1}{n^2} \left(\frac{6}{\bar{X}^2} + \frac{3}{\bar{X}\bar{Y}} \right) \right] \quad (\text{A.93})$$

while:

$$V[r_2] = V \left[\frac{\bar{y}}{\bar{x}} \right] = \frac{1}{n} V \left[\frac{y_j}{x_j} \right] \simeq \frac{R^2}{n} \left[\frac{1}{\bar{X}} + \frac{1}{\bar{Y}} + \frac{6}{\bar{X}^2} + \frac{3}{\bar{X}\bar{Y}} \right] \quad (\text{A.94})$$

Lets consider the simple case where we have 1 dimension. Python Exercise:

Then if X, Y are both 2 random(uniform=homogeneous) distributions between the value $[0,1]$. Then if we measure:

- $\text{mean}(X/Y) = 6$, while $\text{mean}(X)/\text{mean}(Y) = 0.99$

Then if X, Y are both 2 gaussian distributions with the same mean=10 and variance=1 then:

- $\text{mean}(X/Y) - \text{mean}(X)/\text{mean}(Y) = 0.01$ always.

this means that $\text{mean}(X)/\text{mean}(Y)$ is a less biased estimator of 2 random variables.

A.14.1 Hubble Distance Measure

If one would like to prove that we live in an expanding isotropic and Homogeneous universe, he would prefer, ideally, to use probes that are model independent in order to have an observational fact independent from a theoretical framework. As we discussed thoroughly in chapter 1 and chapter 2, part of the assumptions of cosmological models are well tested from other studies, we can focus on the Homogeneity part of the set of cosmological assumptions. So lets address the question: *Can we observe Cosmic Homogeneity in a model independent way?*

Since resolution of the redshift information on large scale structure is not important (actually it is), we could use the low redshift galaxies samples. At these redshift regimes the Hubble Law is reduced from Eq. 1.31 to the simple form $cz = H_0 d$. The difference between the Hubble distance and Λ CDM-distance as a function of redshift is described in figure in Fig. A.3. We could use this law to define distances from our spectroscopic data and have a model independent measurement of galaxy-counts. The distance

⁴There are also other possibilities according to Ogiore et al. [214].

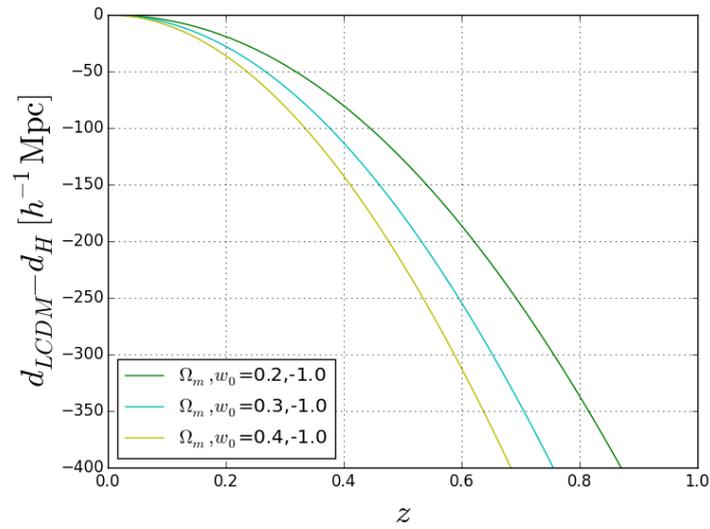


FIGURE A.3: Λ CDM-model Distance minus hubble distance as a function of redshift, z , for different fiducial cosmologies.

error between the model dependent distance measure $d_{\Lambda\text{CDM}}(z) = d_{\text{FLRW}}(z; \Omega_m)$ and hubble distance $d_H(z)$ is about $250h^{-1}\text{Mpc}$ at $z = 0.6$. Since precision is not an important issue we can use d_H which is model independent relationship. However, the reconstructed 3D distribution of galaxies from d_H is completely different from the model dependent reconstructed distribution from d_{FLRW} at the epoch of interests, $z \simeq 0.5$, as shown in Fig. A.3. Therefore, this is not good idea, because one infers wrong distances that are approximately true only on late times, $z \ll 0.1$.

Bibliography

- [1] Emile Faguet. *Initiation Into Philosophy*. BookRix, 2016.
- [2] Tatsumi Aoyama, Masashi Hayakawa, Toichiro Kinoshita, and Makiko Nio. Tenth-order qed contribution to the electron $g-2$ and an improved value of the fine structure constant. *Physical Review Letters*, 109(11):111807, 2012.
- [3] "metre". <https://en.wikipedia.org/wiki/Metre>. Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc. 22 July 2004. Web. 10 Aug. 2004.
- [4] Elias Kiritsis. *String theory in a nutshell*. Princeton University Press, 2011.
- [5] Y Jack Ng. Quantum foam and quantum gravity phenomenology. In *Planck Scale Effects in Astrophysics and Cosmology*, pages 321–349. Springer, 2005.
- [6] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. xiii. cosmological parameters. 594:A13, September 2016.
- [7] Albert Einstein. Kosmologische und relativitätstheorie. *SPA der Wissenschaften*, 142, 1917.
- [8] Mauro Pieroni. Classification of inflationary models and constraints on fundamental physics. *arXiv preprint arXiv:1611.03732*, 2016.
- [9] DJ Fixsen. The temperature of the cosmic microwave background. *The Astrophysical Journal*, 707(2):916, 2009.
- [10] QR Ahmad, RC Allen, TC Andersen, JD Anglin, JC Barton, EW Beier, M Bercovitch, J Bigu, SD Biller, RA Black, et al. Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory. *Physical review letters*, 89(1):011301, 2002.
- [11] Anton Koekemoer. The legacy of the hubble space telescope spectrographs. In *Hubble after SM4. Preparing JWST*, volume 1, page 13, 2010.
- [12] Bodifee, gerard. "catalogue of one thousand named galaxies". retrieved 21 may 2017.
- [13] F Aharonian, AG Akhperjanian, U Barres De Almeida, AR Bazer-Bachi, Yvonne Becherini, B Behera, W Benbow, K Bernlöhr, C Boisson, A Bochow, et al. Energy spectrum of cosmic-ray electrons at tev energies. *Physical Review Letters*, 101(26):261104, 2008.
- [14] Yury Kudenko. Neutrino detectors for oscillation experiments. *arXiv preprint arXiv:1705.06059*, 2017.
- [15] M Fukugita, CJ Hogan, and PJE Peebles. The cosmic baryon budget. *The Astrophysical Journal*, 503(2):518, 1998.

- [16] Fritz Zwicky. On the masses of nebulae and of clusters of nebulae. *The Astrophysical Journal*, 86: 217, 1937.
- [17] Douglas Clowe, Maruša Bradač, Anthony H Gonzalez, Maxim Markevitch, Scott W Randall, Christine Jones, and Dennis Zaritsky. A direct empirical proof of the existence of dark matter. *The Astrophysical Journal Letters*, 648(2):L109, 2006.
- [18] Vera C Rubin and W Kent Ford Jr. Rotation of the andromeda nebula from a spectroscopic survey of emission regions. *The Astrophysical Journal*, 159:379, 1970.
- [19] Florent Leclercq, Guilhem Lavaux, Jens Jasche, and Benjamin Wandelt. Comparing cosmic web classifiers using information theory. *Journal of Cosmology and Astroparticle Physics*, 2016(08): 027, 2016.
- [20] Orfeu Bertolami, Catarina Cosme, and João G Rosa. Scalar field dark matter and the higgs field. *Physics Letters B*, 759:1–8, 2016.
- [21] Shadab Alam, Metin Ata, Stephen Bailey, Florian Beutler, Dmitry Bizyaev, Jonathan A Blazek, Adam S Bolton, Joel R Brownstein, Angela Burden, Chia-Hsun Chuang, et al. The clustering of galaxies in the completed sdss-iii baryon oscillation spectroscopic survey: cosmological analysis of the dr12 galaxy sample. *arXiv preprint arXiv:1607.03155*, 2016.
- [22] R Adam, PAR Ade, N Aghanim, Y Akrami, MIR Alves, M Arnaud, F Arroja, J Aumont, C Baccigalupi, M Ballardini, et al. Planck 2015 results. i. overview of products and scientific results. 2015.
- [23] Scott Dodelson. *Modern cosmology*. Academic press, 2003.
- [24] Daniel Baumann. Cosmology, part iii mathematical tripos. *University lecture notes*, 2014.
- [25] Edwin Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the National Academy of Sciences*, 15(3):168–173, 1929.
- [26] Éric Aubourg, Stephen Bailey, Julian E Bautista, Florian Beutler, Vaishali Bhardwaj, Dmitry Bizyaev, Michael Blanton, Michael Blomqvist, Adam S Bolton, Jo Bovy, et al. Cosmological implications of baryon acoustic oscillation measurements. *Physical Review D*, 92(12):123516, 2015.
- [27] Max Tegmark, Michael A Strauss, Michael R Blanton, Kevork Abazajian, Scott Dodelson, Havard Sandvik, Xiaomin Wang, David H Weinberg, Idit Zehavi, Neta A Bahcall, et al. Cosmological parameters from sdss and wmap. *Physical Review D*, 69(10):103501, 2004.
- [28] Adam G Riess, Alexei V Filippenko, Peter Challis, Alejandro Clocchiatti, Alan Diercks, Peter M Garnavich, Ron L Gilliland, Craig J Hogan, Saurabh Jha, Robert P Kirshner, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009, 1998.
- [29] N Suzuki, D Rubin, C Lidman, G Aldering, R Amanullah, K Barbary, LF Barrientos, J Botyanszki, M Brodwin, N Connolly, et al. The hubble space telescope cluster supernova survey. v. improving the dark-energy constraints above $z_i > 1$ and building an early-type-hosted supernova sample. *The Astrophysical Journal*, 746(1):85, 2012.
- [30] Saul Perlmutter. The evolving search for the nature of dark energy. <http://newscenter.lbl.gov/2009/10/27/evolving-dark-energy/>.

- [31] Charles Alcock and Bohdan Paczynski. An evolution free test for non-zero cosmological constant. *Nature*, 281:358, 1979.
- [32] Friedmann-lemaitre-walker-robertson metric. https://en.wikipedia.org/wiki/Friedmann-Lemaître-Robertson-Walker_metric.
- [33] Julian E Bautista, Nicolás G Busca, Julien Guy, James Rich, Michael Blomqvist, Hélión du Mas des Bourboux, Matthew M Pieri, Andreu Font-Ribera, Stephen Bailey, Timothée Delubac, et al. Measurement of bao correlations at $z = 2.3$ with sdss dr12 lya-forests. *arXiv preprint arXiv:1702.00176*, 2017. .
- [34] Alan H Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, 23(2):347, 1981.
- [35] Andrei D Linde. A new inflationary universe scenario: a possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B*, 108(6):389–393, 1982.
- [36] Andreas Albrecht and Paul J Steinhardt. Cosmology for grand unified theories with radiatively induced symmetry breaking. *Physical Review Letters*, 48(17):1220, 1982.
- [37] Daniel Baumann. Cosmological inflation: Theory and observations. *Advanced Science Letters*, 2(2):105–120, 2009.
- [38] Arno A Penzias and Robert Woodrow Wilson. A measurement of excess antenna temperature at 4080 mc/s. *The Astrophysical Journal*, 142:419–421, 1965.
- [39] Robert H Dicke, P James E Peebles, Peter G Roll, and David T Wilkinson. Cosmic black-body radiation. *The Astrophysical Journal*, 142:414–419, 1965.
- [40] Ralph A Alpher, Hans Bethe, and George Gamow. The origin of chemical elements. *Physical Review*, 73(7):803, 1948.
- [41] DJ Fixsen, ES Cheng, JM Gales, John C Mather, RA Shafer, and EL Wright. The cosmic microwave background spectrum from the full coBE* FIRAS data set. *The Astrophysical Journal*, 473(2):576, 1996. <https://arxiv.org/abs/astro-ph/9605054>.
- [42] R Adam, PAR Ade, N Aghanim, Y Akrami, MIR Alves, F Argüeso, M Arnaud, F Arroja, M Ashdown, J Aumont, et al. Planck 2015 results-i. overview of products and scientific results. *Astronomy & Astrophysics*, 594:A1, 2016.
- [43] Peter AR Ade, N Aghanim, MIR Alves, C Armitage-Caplan, M Arnaud, M Ashdown, F Atrio-Barandela, J Aumont, H Aussel, C Baccigalupi, et al. Planck 2013 results. i. overview of products and scientific results. *Astronomy & Astrophysics*, 571:A1, 2014.
- [44] George F Smoot, Mark V Gorenstein, and Richard A Muller. Detection of anisotropy in the cosmic blackbody radiation. *Physical Review Letters*, 39(14):898, 1977.
- [45] A Kogut, C Lineweaver, George F Smoot, Charles L Bennett, and A Banday. Dipole anisotropy in the coBE dmr first-year sky maps. *arXiv preprint astro-ph/9312056*, 1993.
- [46]

- [47] CL Bennett, D Larson, JL Weiland, N Jarosik, G Hinshaw, N Odegard, KM Smith, RS Hill, B Gold, M Halpern, et al. Nine-year wilkinson microwave anisotropy probe (wmap) observations: final maps and results. *The Astrophysical Journal Supplement Series*, 208(2):20, 2013.
- [48] Gary Hinshaw, D Larson, Eiichiro Komatsu, DN Spergel, CL Bennett, J Dunkley, MR Nolta, M Halpern, RS Hill, N Odegard, et al. Nine-year wilkinson microwave anisotropy probe (wmap) observations: cosmological parameter results. *The Astrophysical Journal Supplement Series*, 208(2):19, 2013.
- [49] CGT Haslam, CJ Salter, H Stoffel, and WEz Wilson. A 408 mhz all-sky continuum survey. ii-the atlas of contour maps. *Astronomy and Astrophysics Supplement Series*, 47:1, 1982.
- [50] Sudeep Das, Thibaut Louis, Michael R Nolta, Graeme E Addison, Elia S Battistelli, J Richard Bond, Erminia Calabrese, Devin Crichton, Mark J Devlin, Simon Dicker, et al. The atacama cosmology telescope: temperature and gravitational lensing power spectrum measurements from three seasons of data. *Journal of Cosmology and Astroparticle Physics*, 2014(04):014, 2014.
- [51] R Keisler, CL Reichardt, KA Aird, BA Benson, LE Bleem, JE Carlstrom, CL Chang, HM Cho, TM Crawford, AT Crites, et al. A measurement of the damping tail of the cosmic microwave background power spectrum with the south pole telescope. *The Astrophysical Journal*, 743(1):28, 2011.
- [52] LE Bleem, A van Engelen, GP Holder, KA Aird, R Armstrong, MLN Ashby, MR Becker, BA Benson, T Biesiadzinski, M Brodwin, et al. A measurement of the correlation of galaxy surveys with cmb lensing convergence maps from the south pole telescope. *The Astrophysical Journal Letters*, 753(1):L9, 2012.
- [53] Daniel J Eisenstein, Idit Zehavi, David W Hogg, Roman Scoccimarro, Michael R Blanton, Robert C Nichol, Ryan Scranton, Hee-Jong Seo, Max Tegmark, Zheng Zheng, et al. Detection of the baryon acoustic peak in the large-scale correlation function of sdss luminous red galaxies. *The Astrophysical Journal*, 633(2):560, 2005.
- [54] Shaun Cole, Will J Percival, John A Peacock, Peder Norberg, Carlton M Baugh, Carlos S Frenk, Ivan Baldry, Joss Bland-Hawthorn, Terry Bridges, Russell Cannon, et al. The 2df galaxy redshift survey: power-spectrum analysis of the final data set and cosmological implications. *Monthly Notices of the Royal Astronomical Society*, 362(2):505–534, 2005.
- [55] Simeon Bird, Ilias Cholis, Julian B Muñoz, Yacine Ali-Haïmoud, Marc Kamionkowski, Ely D Kovetz, Alvise Raccanelli, and Adam G Riess. Did ligo detect dark matter? *Physical review letters*, 116(20):201301, 2016.
- [56] Donghui Jeong and Fabian Schmidt. Large-scale structure with gravitational waves. i. galaxy clustering. *Physical Review D*, 86(8):083512, 2012.
- [57] J-Ch Hamilton. What have we learned from observational cosmology? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 46:70–85, 2014.
- [58] DJ Fixsen, ES Cheng, JM Gales, John C Mather, RA Shafer, and EL Wright. The Cosmic Microwave Background Spectrum from the Full COBE FIRAS Data Set. *The Astrophysical Journal*, 473(2):576, 1996.
- [59] Stephen Appleby and Arman Shafieloo. Testing isotropy in the local universe. *Journal of Cosmology and Astroparticle Physics*, 2014(10):070, 2014.

- [60] Dhiraj Kumar Hazra and Arman Shafieloo. Search for a direction in the forest of lyman- α . *Journal of Cosmology and Astroparticle Physics*, 2015(11):012, 2015.
- [61] CA Schärf, MA Treyer, K Jahoda, T Piran, O Lahav, and EA Boldt. The 2-10 keV xrb dipole and its cosmological implications. 1999.
- [62] Chris Blake and Jasper Wall. Detection of the velocity dipole in the radio galaxies of the nrao vla sky survey. 2002.
- [63] "cosmological principle". https://en.wikipedia.org/wiki/Cosmological_principle. Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc. 22 July 2004. Web. 10 Aug. 2004.
- [64] George FR Ellis. Cosmology and verifiability. *Quarterly Journal of the Royal Astronomical Society*, 16:245–264, 1975.
- [65] Roy Maartens. Is the universe homogeneous? *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 369(1957):5115–5137, 2011.
- [66] Chris Clarkson. Establishing homogeneity of the universe in the shadow of dark energy. *Comptes Rendus Physique*, 13(6):682–718, 2012.
- [67] Norbert Straumann. Minimal assumptions leading to a robertson-walker model of the universe. *Helvetica Physica Acta*, 47(3):379–383, 1974.
- [68] Ruth Durrer, J-P Eckmann, F Sylos Labini, et al. Angular projections of fractal sets. *EPL (Europhysics Letters)*, 40(5):491, 1997.
- [69] Ben Hoyle, Rita Tojeiro, Raul Jimenez, et al. Testing homogeneity with galaxy star formation histories. *The Astrophysical Journal Letters*, 762(1):L9, 2012. <https://arxiv.org/abs/1209.6181>.
- [70] James Zibin, Adam Moss, et al. Nowhere to hide: closing in on cosmological homogeneity. 2014.
- [71] R Brent Tully, Helene Courtois, Yehuda Hoffman, and Daniel Pomarède. The laniakea supercluster of galaxies. *Nature*, 513(7516):71–73, 2014. <https://arxiv.org/pdf/1409.0880.pdf>.
- [72] RG Clowes. Large quasar groups—a short review. In *The New Era of Wide Field Astronomy*, volume 232, page 108, 2001. <http://adsabs.harvard.edu/full/2001ASPC...232..108C>.
- [73] R Brent Tully. Alignment of clusters and galaxies on scales up to 0.1 c. *The Astrophysical Journal*, 303:25–38, 1986. Pisces-Cetus SuperCluster Complex (PCSC), <http://adsabs.harvard.edu/abs/1986ApJ...303...25T>.
- [74] Margaret J Geller and John P Huchra. Mapping the universe. *Science*, 246(4932):897, 1989. Great Wall.
- [75] AA Klypin and Ali Kopylov. The spatial covariance function for rich clusters of galaxies. *Soviet Astronomy Letters*, 9:41–44, 1983. Giand Void, <http://adsabs.harvard.edu/full/1983SvAL...9...41K>.
- [76] J Richard Gott III, Mario Jurić, David Schlegel, Fiona Hoyle, Michael Vogeley, Max Tegmark, Neta Bahcall, and Jon Brinkmann. A map of the universe. *The Astrophysical Journal*, 624(2):463, 2005. <http://iopscience.iop.org/article/10.1086/428890/meta>.

- [77] Yuichi Matsuda, Toru Yamada, Tomoki Hayashino, Hajime Tamura, Ryosuke Yamauchi, Takashi Murayama, Tohru Nagao, Kouji Ohta, Sadanori Okamura, Masami Ouchi, et al. Large-scale filamentary structure around the protocluster at redshift $z=3.1$. *The Astrophysical Journal Letters*, 634(2):L125, 2005. <https://arxiv.org/abs/astro-ph/0510762>.
- [78] Lawrence Rudnick, Shea Brown, and Liliya R Williams. Extragalactic radio sources and the wmap cold spot. *The Astrophysical Journal*, 671(1):40, 2007. <https://arxiv.org/abs/0704.0908v2>.
- [79] Roger G Clowes, Kathryn A Harris, Srinivasan Raghunathan, Luis E Campusano, Ilona K Söchtig, and Matthew J Graham. A structure in the early universe at $z \sim 1.3$ that exceeds the homogeneity scale of the Λ CDM concordance cosmology. *Monthly Notices of the Royal Astronomical Society*, 429(4):2910–2916, 2013. <https://arxiv.org/pdf/1211.6256.pdf>.
- [80] I Horvath, J Hakkila, and Z Bagoly. The largest structure of the universe, defined by gamma-ray bursts. *arXiv preprint arXiv:1311.1104*, 2013. <https://arxiv.org/abs/1311.1104>.
- [81] Kolmogorov–smirnov test. *Encyclopedia of Mathematics*.
- [82] H Lietzen, E Tempel, LJ Liivamägi, A Montero-Dorta, M Einasto, A Streblyanska, C Maraston, JA Rubiño-Martín, and E Saar. Discovery of a massive supercluster system at $z \sim 0.47$. *Astronomy & Astrophysics*, 588:L4, 2016. <https://arxiv.org/abs/1602.08498>.
- [83] D. W. Hogg, D. J. Eisenstein, M. R. Blanton, et al. Cosmic homogeneity demonstrated with luminous red galaxies. *The Astrophysical Journal*, 624(1):54, 2005.
- [84] Jaswant Yadav, Somnath Bharadwaj, Biswajit Pandey, and TR Seshadri. Testing homogeneity on large scales in the sloan digital sky survey data release one. *Monthly Notices of the Royal Astronomical Society*, 364(2):601–606, 2005.
- [85] Prakash Sarkar, Jaswant Yadav, Biswajit Pandey, and Somnath Bharadwaj. The scale of homogeneity of the galaxy distribution in sdss dr6. *Monthly Notices of the Royal Astronomical Society: Letters*, 399(1):L128–L131, 2009.
- [86] Vicent J Martínez and Peter Coles. Correlations and scaling in the qdot redshift survey. *The Astrophysical Journal*, 437:550–555, 1994.
- [87] Luigi Guzzo. Is the universe homogeneous?(on large scales). *New Astronomy*, 2(6):517–532, 1997.
- [88] Vicent J Martínez, Maria-Jesus Pons-Borderia, Rana A Moeed, and Matthew J Graham. Searching for the scale of homogeneity. *Monthly Notices of the Royal Astronomical Society*, 298(4):1212–1222, 1998.
- [89] R Scaramella, L Guzzo, G Zamorani, E Zucca, C Balkowski, A Blanchard, A Cappi, V Cayatte, G Chincarini, C Collins, et al. The eso slice project [esp] galaxy redshift survey: V. evidence for a $d=3$ sample dimensionality. 1998.
- [90] Luca Amendola and Emilia Palladino. The scale of homogeneity in the las campanas redshift survey. *The Astrophysical Journal Letters*, 514(1):L1, 1999.
- [91] Jun Pan and Peter Coles. Large-scale cosmic homogeneity from a multifractal analysis of the pscz catalogue. *Monthly Notices of the Royal Astronomical Society*, 318(4):L51–L54, 2000.
- [92] T Kurokawa, M Morikawa, and H Mouri. Scaling analysis of galaxy distribution in the las campanas redshift survey data. *Astronomy & Astrophysics*, 370(2):358–364, 2001.

- [93] M. I. Scrimgeour, T. Davis, C. Blake, J. B. James, et al. The wiggles dark energy survey: the transition to large-scale cosmic homogeneity. *Monthly Notices of the Royal Astronomical Society*, 425(1):116–134, 2012.
- [94] Paul H Coleman and Luciano Pietronero. The fractal structure of the universe. *Physics Reports*, 213(6):311–389, 1992.
- [95] N Turok. Critical dialogues in cosmology. proceedings. *Critical dialogues in cosmology. Proceedings.*, by Turok, N.. World Scientific, Singapore (Singapore), 1997, X+ 614 p., ISBN 981-02-2859-7, ISBN 981-02-2860-0 (paper)., 1, 1997. <https://arxiv.org/abs/astro-ph/9611197>.
- [96] F Sylos Labini, M Montuori, and L Pietronero. Comment on the paper by l. guzzo” is the universe homogeneous?”. 1998.
- [97] F Sylos Labini, Marco Montuori, and Luciano Pietronero. Scale-invariance of galaxy clustering. *Physics Reports*, 293(2):61–226, 1998.
- [98] Michael Joyce, Marco Montuori, and F Sylos Labini. Fractal correlations in the cfa2-south redshift survey. *The Astrophysical Journal Letters*, 514(1):L5, 1999.
- [99] Cosmologist blog: Stolpovskiy misha. <http://stolpovskiy.weebly.com>.
- [100] ML McClure and CC Dyer. Anisotropy in the hubble constant as observed in the hst extragalactic distance scale key project results. *New Astronomy*, 12(7):533–543, 2007.
- [101] RR Caldwell and A Stebbins. A test of the copernican principle. *Physical review letters*, 100(19):191302, 2008.
- [102] András Kovács and Juan García-Bellido. Cosmic troublemakers: the cold spot, the eridanus supervoid, and the great walls. *Monthly Notices of the Royal Astronomical Society*, 462(2):1882–1893, 2016.
- [103] Michael R Blanton, Matthew A Bershady, Bela Abolfathi, Franco D Albareti, Carlos Allende Prieto, Andres Almeida, Javier Alonso-García, Friedrich Anders, Scott F Anderson, Brett Andrews, et al. Sloan digital sky survey iv: Mapping the milky way, nearby galaxies and the distant universe. *arXiv preprint arXiv:1703.00052*, 2017.
- [104] Gábor Rácz, László Dobos, Róbert Beck, István Szapudi, and István Csabai. Concordance cosmology without dark energy. *Monthly Notices of the Royal Astronomical Society: Letters*, 469(1):L1–L5, 2017.
- [105] Daniela Saadeh, Stephen M Feeney, Andrew Pontzen, Hiranya V Peiris, and Jason D McEwen. How isotropic is the universe? *Physical Review Letters*, 117(13):131302, 2016.
- [106] Dominik J Schwarz. Accelerated expansion without dark energy. *arXiv preprint astro-ph/0209584*, 2002.
- [107] Thomas Buchert. A cosmic equation of state for the inhomogeneous universe: can a global far-from-equilibrium state explain dark energy? *Classical and quantum gravity*, 22(19):L113, 2005.
- [108] Nan Li. Cosmological backreaction: from the local hubble expansion rate to dark energy. 2008.
- [109] Jean-Christophe HAMILTON. *Recherche automatisée de supernovae à des distances intermédiaires et analyse photométrique de leurs courbes de lumière*. Theses, Université Paris Sud - Paris XI, April 1999. URL <https://tel.archives-ouvertes.fr/tel-00000941>. Cosmologie observationnelle.

- [110] Pilar Ruiz-Lapuenteft. Cosmological insights from supernovae. In *Dark Matter in Astrophysics and Particle Physics 1998: Proceedings of the Second International Conference on Dark Matter in Astro and Particle Physics, held in Heidelberg, Germany, 20-25 July 1998*, page 197. CRC Press, 1999. <https://arxiv.org/pdf/astro-ph/9810423.pdf>.
- [111] Julian Adamek, Chris Clarkson, David Daverio, Ruth Durrer, and Martin Kunz. Safely smoothing spacetime: backreaction in relativistic cosmological simulations. *arXiv preprint arXiv:1706.09309*, 2017.
- [112] M Rasse and E Ahmed. On fractal cosmology. *Astrophysical Letters and Communications*, 35: 311, 1996.
- [113] Kock snowflake.
- [114] Nick Kaiser. Clustering in real space and in redshift space. *Monthly Notices of the Royal Astronomical Society*, 227(1):1–21, 1987.
- [115] Eric V Linder and Robert N Cahn. Parameterized beyond-einstein growth. *Astroparticle Physics*, 28(4):481–488, 2007.
- [116] AJS Hamilton. Measuring omega and the real correlation function from the redshift correlation function. *The Astrophysical Journal*, 385:L5–L8, 1992.
- [117] Nathalie Palanque-Delabrouille, Christophe Yèche, Julien Baur, Christophe Magneville, Graziano Rossi, Julien Lesgourgues, Arnaud Borde, Etienne Burtin, Jean-Marc LeGoff, James Rich, et al. Neutrino masses and cosmology with lyman-alpha forest power spectrum. *Journal of Cosmology and Astroparticle Physics*, 2015(11):011, 2015.
- [118] Stacy S McGaugh. The baryonic tully-fisher relation of gas-rich galaxies as a test of λ cdm and mond. *The Astronomical Journal*, 143(2):40, 2012.
- [119] James E Gunn and Bruce A Peterson. On the density of neutral hydrogen in intergalactic space. *The Astrophysical Journal*, 142:1633–1641, 1965.
- [120] Michael J Drinkwater, Russell J Jurek, Chris Blake, David Woods, Kevin A Pimblet, Karl Glazebrook, Rob Sharp, Michael B Pracy, Sarah Brough, Matthew Colless, et al. The wigglez dark energy survey: survey design and first data release. *Monthly Notices of the Royal Astronomical Society*, 401(3):1429–1452, 2010.
- [121] Amir Aghamousa, Jessica Aguilar, Steve Ahlen, Shadab Alam, Lori E Allen, Carlos Allende Prieto, James Annis, Stephen Bailey, Christophe Balland, Otger Ballester, et al. The desi experiment part i: Science, targeting, and survey design. *arXiv preprint arXiv:1611.00036*, 2016.
- [122] LSST Dark Energy Science Collaboration et al. Large synoptic survey telescope: dark energy science collaboration. *arXiv preprint arXiv:1211.0310*, 2012.
- [123] Y Mellier. Euclid: Mapping the geometry of the dark universe. In *Science from the Next Generation Imaging and Spectroscopic Surveys*, volume 1, page 3, 2012.
- [124] Dietrich Korsch. Closed form solution for three-mirror telescopes, corrected for spherical aberration, coma, astigmatism, and field curvature. *Applied Optics*, 11(12):2986–2987, 1972.
- [125] Donald G York, J Adelman, John E Anderson Jr, Scott F Anderson, James Annis, Neta A Bahcall, JA Bakken, Robert Barkhouser, Steven Bastian, Eileen Berman, et al. The sloan digital sky survey: Technical summary. *The Astronomical Journal*, 120(3):1579, 2000.

- [126] James E Gunn, Walter A Siegmund, Edward J Mannery, Russell E Owen, Charles L Hull, R French Leger, Larry N Carey, Gillian R Knapp, Donald G York, William N Boroski, et al. The 2.5 m telescope of the sloan digital sky survey. *The Astronomical Journal*, 131(4):2332, 2006.
- [127] Daniel J Eisenstein, David H Weinberg, Eric Agol, Hiroaki Aihara, Carlos Allende Prieto, Scott F Anderson, James A Arns, Éric Aubourg, Stephen Bailey, Eduardo Balbinot, et al. Sdss-iii: Massive spectroscopic surveys of the distant universe, the milky way, and extra-solar planetary systems. *The Astronomical Journal*, 142(3):72, 2011.
- [128] Steven R Majewski, MF Skrutskie, RP Schiavon, JC Wilson, RW O’Connell, VV Smith, M Shetrone, K Cunha, PM Frinchaboy, IN Reid, et al. The apache point observatory galactic evolution experiment (apogee). In *Bulletin of the American Astronomical Society*, volume 39, page 962, 2007.
- [129] Kevin Bundy, Richard S Ellis, Christopher J Conselice, James E Taylor, Michael C Cooper, Christopher NA Willmer, Benjamin J Weiner, Alison L Coil, Kai G Noeske, and Peter RM Eisenhardt. The mass assembly history of field galaxies: detection of an evolving mass limit for star-forming galaxies. *The Astrophysical Journal*, 651(1):120, 2006.
- [130] A Merloni, P Predehl, W Becker, H Böhringer, Th Boller, H Brunner, M Brusa, K Dennerl, M Freyberg, P Friedrich, et al. erosita science book: mapping the structure of the energetic universe. *arXiv preprint arXiv:1209.3114*, 2012.
- [131] H. A. Flewelling, E. A. Magnier, K. C. Chambers, J. N. Heasley, C. Holmberg, M. E. Huber, W. Sweeney, C. Z. Waters, X. Chen, D. Farrow, G. Hasinger, R. Henderson, K. S. Long, N. Metcalfe, M. A. Nieto-Santisteban, P. Norberg, R. P. Saglia, A. Szalay, A. Rest, A. R. Thakar, J. L. Tonry, J. Valenti, S. Werner, R. White, L. Denneau, P. W. Draper, K. W. Hodapp, R. Jedicke, N. Kaiser, R.-P. Kudritzki, P. A. Price, R. J. Wainscoat, S. Chastel, B. McClean, M. Postman, and B. Shiao. The Pan-STARRS1 Database and Data Products. *ArXiv e-prints*.
- [132] Kyle S Dawson, Jean-Paul Kneib, Will J Percival, Shadab Alam, Franco D Albareti, Scott F Anderson, Eric Armengaud, Éric Aubourg, Stephen Bailey, Julian E Bautista, et al. The sdss-iv extended baryon oscillation spectroscopic survey: overview and early data. *The Astronomical Journal*, 151(2):44, 2016.
- [133] Gong-Bo Zhao, Yuting Wang, Ashley J Ross, Sarah Shandera, Will J Percival, Kyle S Dawson, Jean-Paul Kneib, Adam D Myers, Joel R Brownstein, Johan Comparat, et al. The extended baryon oscillation spectroscopic survey: a cosmological forecast. *Monthly Notices of the Royal Astronomical Society*, 457(3):2377–2390, 2016.
- [134] James E Gunn, Walter A Siegmund, Edward J Mannery, Russell E Owen, Charles L Hull, R French Leger, Larry N Carey, Gillian R Knapp, Donald G York, William N Boroski, et al. The 2.5 m telescope of the sloan digital sky survey. *The Astronomical Journal*, 131(4):2332, 2006.
- [135] JE Gunn, M Carr, C Rockosi, M Sekiguchi, K Berry, B Elms, E De Haas, Ž Ivezić, G Knapp, R Lupton, et al. The sloan digital sky survey photometric camera. *The Astronomical Journal*, 116(6):3040, 1998.
- [136] Stephen A Smee, James E Gunn, Alan Uomoto, Natalie Roe, David Schlegel, Constance M Rockosi, Michael A Carr, French Leger, Kyle S Dawson, Matthew D Olmstead, et al. The multi-object, fiber-fed spectrographs for the sloan digital sky survey and the baryon oscillation spectroscopic survey. *The Astronomical Journal*, 146(2):32, 2013.

- [137] Daniel J Eisenstein, James Annis, James E Gunn, Alexander S Szalay, Andrew J Connolly, RC Nichol, Neta A Bahcall, Mariangela Bernardi, Scott Burles, Francisco J Castander, et al. Spectroscopic target selection for the sloan digital sky survey: The luminous red galaxy sample. *The Astronomical Journal*, 122(5):2267, 2001.
- [138] Claudia Maraston, G Strömbäck, Daniel Thomas, et al. Modelling the colour evolution of luminous red galaxies—improvements with empirical stellar spectra. *Monthly Notices of the Royal Astronomical Society: Letters*, 394(1):L107–L111, 2009.
- [139] Beth Reid, Shirley Ho, Nikhil Padmanabhan, Will J Percival, Jeremy Tinker, Rita Tojeiro, Martin White, Daniel J Eisenstein, Claudia Maraston, Ashley J Ross, et al. Sdss-iii baryon oscillation spectroscopic survey data release 12: galaxy target selection and large-scale structure catalogues. *Monthly Notices of the Royal Astronomical Society*, 455(2):1553–1573, 2016.
- [140] Stephen A Smee, James E Gunn, Alan Uomoto, Natalie Roe, David Schlegel, Constance M Rockosi, Michael A Carr, French Leger, Kyle S Dawson, Matthew D Olmstead, et al. The multi-object, fiber-fed spectrographs for the sloan digital sky survey and the baryon oscillation spectroscopic survey. *The Astronomical Journal*, 146(2):32, 2013.
- [141] Adam S Bolton, David J Schlegel, Éric Aubourg, Stephen Bailey, Vaishali Bhardwaj, Joel R Brownstein, Scott Burles, Yan-Mei Chen, Kyle Dawson, Daniel J Eisenstein, et al. Spectral classification and redshift measurement for the sdss-iii baryon oscillation spectroscopic survey. *The Astronomical Journal*, 144(5):144, 2012.
- [142] Timothy A Hutchinson, Adam S Bolton, Kyle S Dawson, Carlos Allende Prieto, Stephen Bailey, Julian E Bautista, Joel R Brownstein, Charlie Conroy, Julien Guy, Adam D Myers, et al. Redshift measurement and spectral classification for eboss galaxies with the redmonster software. *The Astronomical Journal*, 152(6):205, 2016.
- [143] Héctor Gil-Marín, Will J Percival, Antonio J Cuesta, Joel R Brownstein, Chia-Hsun Chuang, Shirley Ho, Francisco-Shu Kitaura, Claudia Maraston, Francisco Prada, Sergio Rodríguez-Torres, et al. The clustering of galaxies in the sdss-iii baryon oscillation spectroscopic survey: Bao measurement from the los-dependent power spectrum of dr12 boss galaxies. *Monthly Notices of the Royal Astronomical Society*, 460(4):4210–4219, 2016.
- [144] Pierros Ntelis, Jean-Christophe Hamilton, Nicolas Guillermo Busca, Eric Aubourg, Pierre Laurent, Jean-Marc Le Goff, James Rich, Etienne Burtin, Hélión du Mas des Bourboux, Nathalie Palanque Delabrouille, et al. Exploring cosmic homogeneity with the boss dr12 galaxy sample. *arXiv preprint arXiv:1702.02159*, 2017.
- [145] Éric Aubourg, Stephen Bailey, Florian Beutler, et al. Cosmological implications of baryon acoustic oscillation measurements. *Physical Review D*, 92(12):123516, 2015.
- [146] Hume A Feldman, Nick Kaiser, and John A Peacock. Power spectrum analysis of three-dimensional redshift surveys. *arXiv preprint astro-ph/9304022*, 1993.
- [147] Suketu P Bhavsar. Bootstrap, data permuting and extreme value distributions: Getting the most out of small samples. *Errors, Bias and Uncertainties in Astronomy*, Cambridge University Press, Cambridge (UK), pages 107–122, 1990.
- [148] Andreas A Berlind, David H Weinberg, Andrew J Benson, Carlton M Baugh, Shaun Cole, Romeel Davé, Carlos S Frenk, Adrian Jenkins, Neal Katz, and Cedric G Lacey. The halo occupation distribution and the physics of galaxy formation. *The Astrophysical Journal*, 593(1):1, 2003.

- [149] Volker Springel, Simon DM White, Adrian Jenkins, Carlos S Frenk, Naoki Yoshida, Liang Gao, Julio Navarro, Robert Thacker, Darren Croton, John Helly, et al. Simulations of the formation, evolution and clustering of galaxies and quasars. *nature*, 435(7042):629–636, 2005. <https://wwwmpa.mpa-garching.mpg.de/galform/virgo/millennium/>.
- [150] Martin White, Jeremy L Tinker, and Cameron K McBride. Mock galaxy catalogues using the quick particle mesh method. *Monthly Notices of the Royal Astronomical Society*, 437(3):2594–2606, 2014.
- [151] Martin White. The zel’dovich approximation. *Monthly Notices of the Royal Astronomical Society*, 439(4):3630–3640, 2014.
- [152] Ya B Zel’Dovich. Gravitational instability: An approximate theory for large density perturbations. *Astronomy and astrophysics*, 5:84–89, 1970.
- [153] Marc Manera, Lado Samushia, Rita Tojeiro, Cullan Howlett, Ashley J Ross, Will J Percival, Hector Gil-Marín, Joel R Brownstein, Angela Burden, and Francesco Montesano. The clustering of galaxies in the sdss-iii baryon oscillation spectroscopic survey: mock galaxy catalogues for the low-redshift sample. *Monthly Notices of the Royal Astronomical Society*, 447(1):437–445, 2015.
- [154] Svetlin Tassev, Matias Zaldarriaga, and Daniel J Eisenstein. Solving large scale structure in ten easy steps with cola. *Journal of Cosmology and Astroparticle Physics*, 2013(06):036, 2013.
- [155] Cullan Howlett, Marc Manera, and Will J Percival. L-picola: A parallel code for fast dark matter simulation. *Astronomy and Computing*, 12:109–126, 2015.
- [156] Chia-Hsun Chuang, Francisco-Shu Kitaura, Francisco Prada, Cheng Zhao, and Gustavo Yepes. Ezmocks: extending the zel’dovich approximation to generate mock galaxy catalogues with accurate clustering statistics. *Monthly Notices of the Royal Astronomical Society*, 446(3):2621–2628, 2015.
- [157] PJE Peebles. The large-scale structure of the universe (research supported by the national science foundation. princeton, nj, princeton university press, 435 p., 1980). In *ADS*, volume 25, page 82, 1980. http://ned.ipac.caltech.edu/level5/Sept02/Saslaw/Saslaw1_6.html.
- [158] Jaswant K Yadav, JS Bagla, and Nishikanta Khandai. Fractal dimension as a measure of the scale of homogeneity. *Monthly Notices of the Royal Astronomical Society*, 405(3):2009–2015, 2010.
- [159] M Vargas-Magaña, Julian E Bautista, J-Ch Hamilton, É Aubourg, A Labatie, J-M Le Goff, S Escoffier, Marc Manera, CK McBride, DP Schneider, et al. An optimized correlation function estimator for galaxy surveys. *Astronomy & Astrophysics*, 554:A131, 2013.
- [160] Stephen D Landy and Alexander S Szalay. Bias and variance of angular correlation functions. *The Astrophysical Journal*, 412:64–71, 1993.
- [161] Phillip James Edwin Peebles and Michael G Hauser. Statistical analysis of catalogs of extragalactic objects. iii. the shane-wirtanen and zwicky catalogs. *The Astrophysical Journal Supplement Series*, 28:19, 1974.
- [162] Paul C Hewett. The estimation of galaxy angular correlation functions. *Monthly Notices of the Royal Astronomical Society*, 201(4):867–883, 1982.
- [163] Marc Davis and PJE Peebles. A survey of galaxy redshifts. v-the two-point position and velocity correlations. *The Astrophysical Journal*, 267:465–482, 1983.

- [164] AJS Hamilton. Toward better ways to measure the galaxy correlation function. *The Astrophysical Journal*, 417:19, 1993.
- [165] Phillip James Edwin Peebles and Michael G Hauser. Statistical analysis of catalogs of extragalactic objects. iii. the shane-wirtanen and zwicky catalogs. *The Astrophysical Journal Supplement Series*, 28:19, 1974.
- [166] P. Laurent, J.-M. Le Goff, E. Burtin, J.-C. Hamilton, D. W. Hogg, A. Myers, P. Ntelis, I. Pâris, J. Rich, E. Aubourg, J. Bautista, T. Delubac, H. du Mas des Bourboux, S. Eftekharzadeh, N. Palanque Delabrouille, P. Petitjean, G. Rossi, D. P. Schneider, and C. Yeche. A $14 \text{ h}^{-3} \text{ Gpc}^3$ study of cosmic homogeneity using BOSS DR12 quasar sample. 11:060, November 2016. doi: 10.1088/1475-7516/2016/11/060.
- [167] Trapezoidal rule.
- [168] Finite difference.
- [169] Diego Blas, Julien Lesgourgues, and Thomas Tram. The cosmic linear anisotropy solving system (class). part ii: approximation schemes. *Journal of Cosmology and Astroparticle Physics*, 2011 (07):034, 2011.
- [170] Antony Lewis and Anthony Challinor. Camb: Code for anisotropies in the microwave background. *Astrophysics Source Code Library*, 2011.
- [171] Daniel Thomas, Oliver Steele, Claudia Maraston, J Johansson, Alessandra Beifiori, Janine Pforr, G Strömbäck, CA Tremonti, D Wake, D Bizyaev, et al. Stellar velocity dispersions and emission line properties of sdss-iii/boss galaxies. *Monthly Notices of the Royal Astronomical Society*, page stt261, 2013.
- [172] WE Ballinger, JA Peacock, and AF Heavens. Measuring the cosmological constant with redshift surveys. *arXiv preprint astro-ph/9605017*, 1996.
- [173] Martin White, M Blanton, A Bolton, D Schlegel, J Tinker, A Berlind, L Da Costa, E Kazin, Y-T Lin, M Maia, et al. The clustering of massive galaxies at $z \approx 0.5$ from the first semester of boss data. *The Astrophysical Journal*, 728(2):126, 2011.
- [174] Sebastián E Nuza, Ariel G Sánchez, Francisco Prada, et al. The clustering of galaxies at $z \approx 0.5$ in the sdss-iii data release 9 boss-cmass sample: a test for the Λ cdm cosmology. *Monthly Notices of the Royal Astronomical Society*, 432(1):743–760, 2013.
- [175] Shirley Ho, Antonio Cuesta, Hee-Jong Seo, Roland De Putter, Ashley J Ross, Martin White, Nikhil Padmanabhan, Shun Saito, David J Schlegel, Eddie Schlafly, et al. Clustering of sloan digital sky survey iii photometric luminous galaxies: the measurement, systematics, and cosmological implications. *The Astrophysical Journal*, 761(1):14, 2012.
- [176] Johan Comparat, Eric Jullo, Jean-Paul Kneib, Carlo Schimd, HuanYuan Shan, Thomas Erben, Olivier Ilbert, Joel Brownstein, Anne Ealet, Stephanie Escoffier, et al. Stochastic bias of colour-selected bao tracers by joint clustering–weak lensing analysis. *Monthly Notices of the Royal Astronomical Society*, 433(2):1146–1160, 2013.
- [177] Héctor Gil-Marín, Jorge Noreña, Licia Verde, Will J Percival, Christian Wagner, Marc Manera, and Donald P Schneider. The power spectrum and bispectrum of sdss dr11 boss galaxies–i. bias and gravity. *Monthly Notices of the Royal Astronomical Society*, 451(1):5058–5099, 2015.

- [178] Prakash Sarkar, Subhabrata Majumdar, Biswajit Pandey, et al. The many scales to cosmic homogeneity: Use of multiple tracers from the sdss. 2016.
- [179] DJ Bacon, AN Taylor, ML Brown, et al. Evolution of the dark matter distribution with three-dimensional weak lensing. *Monthly Notices of the Royal Astronomical Society*, 363(3):723–733, 2005.
- [180] Andy Taylor, Benjamin Joachimi, and Thomas Kitching. Putting the precision in precision cosmology: How accurate should your data covariance matrix be? *Monthly Notices of the Royal Astronomical Society*, 432(3):1928–1946, 2013.
- [181] C Castagnoli and Astron Provenzale. From small-scale fractality to large-scale homogeneity—a family of cascading models for the distribution of galaxies. *Astronomy and Astrophysics*, 246: 634–643, 1991.
- [182] G Lemaître. The expanding universe, 1997. *Gen. Rel. Grav*, 29:641.
- [183] Richard C Tolman. Effect of inhomogeneity on cosmological models. *Proceedings of the National Academy of Sciences*, 20(3):169–176, 1934.
- [184] Hermann Bondi. Spherically symmetrical models in general relativity. *Monthly Notices of the Royal Astronomical Society*, 107(5-6):410–425, 1947.
- [185] Nazeem Mustapha, Charles Hellaby, and GFR Ellis. Large-scale inhomogeneity versus source evolution: can we distinguish them observationally? *Monthly Notices of the Royal Astronomical Society*, 292(4):817–830, 1997.
- [186] Lauren Anderson, Eric Aubourg, Stephen Bailey, Dmitry Bizyaev, Michael Blanton, Adam S Bolton, J Brinkmann, Joel R Brownstein, Angela Burden, Antonio J Cuesta, et al. The clustering of galaxies in the sdss-iii baryon oscillation spectroscopic survey: baryon acoustic oscillations in the data release 9 spectroscopic galaxy sample. *Monthly Notices of the Royal Astronomical Society*, 427(4):3435–3467, 2012.
- [187] Eusebio Sánchez, Aurelio Carnero, Juan García-Bellido, Enrique Gaztañaga, F De Simoni, Martín Crocce, Anna Cabré, Pablo Fosalba, and David Alonso. Tracing the sound horizon scale with photometric redshift surveys. *Monthly Notices of the Royal Astronomical Society*, 411(1):277–288, 2011.
- [188] Shadab Alam, Metin Ata, Stephen Bailey, Florian Beutler, Dmitry Bizyaev, Jonathan A Blazek, Adam S Bolton, Joel R Brownstein, Angela Burden, Chia-Hsun Chuang, et al. The clustering of galaxies in the completed sdss-iii baryon oscillation spectroscopic survey: cosmological analysis of the dr12 galaxy sample. *Monthly Notices of the Royal Astronomical Society*, page stx721, 2017.
- [189] Andreas Albrecht, Gary Bernstein, Robert Cahn, Wendy L Freedman, Jacqueline Hewitt, Wayne Hu, John Huth, Marc Kamionkowski, Edward W Kolb, Lloyd Knox, et al. Report of the dark energy task force. *arXiv preprint astro-ph/0609591*, 2006.
- [190] Fisher matrix.
- [191] Planck Collaboration et al. Planck 2013 results. xvi. cosmological parameters. 2014.
- [192] Pymc user’s guide. <https://pymc-devs.github.io/pymc/>.

- [193] Pierre Laurent. L'univers aux grandes échelles : études de l'homogénéité cosmique et de l'énergie noire à partir des relevés de quasars boss et eboss. <https://tel.archives-ouvertes.fr/tel-01392968>.
- [194] Zhaoming Ma, Wayne Hu, and Dragan Huterer. Effects of photometric redshift uncertainties on weak-lensing tomography. *The Astrophysical Journal*, 636(1):21, 2006.
- [195] Emmanuel Schaan, Elisabeth Krause, Tim Eifler, Olivier Doré, Hironao Miyatake, Jason Rhodes, and David N Spergel. Looking through the same lens: Shear calibration for lsst, euclid, and wfirst with stage 4 cmb lensing. *Physical Review D*, 95(12):123512, 2017.
- [196] M. Tegmark. Measuring Cosmological Parameters with Galaxy Surveys. *Physical Review Letters*, 79:3806–3809, November 1997. doi: 10.1103/PhysRevLett.79.3806.
- [197] C. Blake, D. Parkinson, B. Bassett, K. Glazebrook, M. Kunz, and R.C. Nichol. Universal fitting formulae for baryon oscillation surveys. *mnras*, 365:255–264, January 2006. doi: 10.1111/j.1365-2966.2005.09706.x.
- [198] Andreu Font-Ribera, Patrick McDonald, Nick Mostek, Beth A Reid, Hee-Jong Seo, and Anže Slosar. Desi and other dark energy experiments in the era of neutrino mass measurements. *Journal of Cosmology and Astroparticle Physics*, 2014(05):023, 2014.
- [199] Sean M Carroll, William H Press, and Edwin L Turner. The cosmological constant. *Annual review of astronomy and astrophysics*, 30(1):499–542, 1992.
- [200] Raul Jimenez and Abraham Loeb. Constraining cosmological parameters based on relative galaxy ages. *The Astrophysical Journal*, 573(1):37, 2002.
- [201] David Alonso. Cute solutions for two-point correlation functions from large cosmological datasets. *arXiv preprint arXiv:1210.1833*, 2012.
- [202] M Moresco, A Cimatti, Raul Jimenez, L Pozzetti, G Zamorani, M Bolzonella, J Dunlop, F Lamareille, M Mignoli, H Pearce, et al. Improved constraints on the expansion rate of the universe up to $z \sim 1.1$ from the spectroscopic evolution of cosmic chronometers. *Journal of Cosmology and Astroparticle Physics*, 2012(08):006, 2012.
- [203] Adam G Riess, Alexei V Filippenko, Peter Challis, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009, 1998.
- [204] Adam G Riess, Louis-Gregory Strolger, John Tonry, Stefano Casertano, Henry C Ferguson, Bahram Mobasher, Peter Challis, Alexei V Filippenko, Saurabh Jha, Weidong Li, et al. Type Ia supernova discoveries at $z < 1$ from the hubble space telescope: evidence for past deceleration and constraints on dark energy evolution. *The Astrophysical Journal*, 607(2):665, 2004.
- [205] Alan Heavens, Raul Jimenez, and Licia Verde. Standard rulers, candles, and clocks from the low-redshift universe. *Physical review letters*, 113(24):241302, 2014.
- [206] John G Kirkwood. Statistical mechanics of fluid mixtures. *The Journal of Chemical Physics*, 3(5):300–313, 1935.
- [207] Jaswant Yadav, Somnath Bharadwaj, Biswajit Pandey, and TR Seshadri. Testing homogeneity on large scales in the sloan digital sky survey data release one. *Monthly Notices of the Royal Astronomical Society*, 364(2):601–606, 2005.

-
- [208] Biswajit Pandey. A method for testing the cosmic homogeneity with shannon entropy. *Monthly Notices of the Royal Astronomical Society*, page stt134, 2013.
- [209] David Kraljic. Characterizing cosmic inhomogeneity with anomalous diffusion. *Monthly Notices of the Royal Astronomical Society*, 451(4):3393–3399, 2015.
- [210] Benjamin P Abbott, Richard Abbott, TD Abbott, MR Abernathy, Fausto Acernese, Kendall Ackley, Carl Adams, Thomas Adams, Paolo Addesso, RX Adhikari, et al. Observation of gravitational waves from a binary black hole merger. *Physical review letters*, 116(6):061102, 2016.
- [211] BP Abbott, R Abbott, TD Abbott, MR Abernathy, F Acernese, K Ackley, C Adams, T Adams, P Addesso, RX Adhikari, et al. The basic physics of the binary black hole merger gw150914. *Annalen der Physik*, 2016.
- [212] Wikipedia. B-spline — wikipedia, the free encyclopedia, 2017. <https://en.wikipedia.org/wiki/B-spline>.
- [213] Cyrille Doux. Cross-correlation between cmb and lss, 2017. In preparation.
- [214] RC Ogliore, GR Huss, and K Nagashima. Ratio estimation in sims analysis. *Nuclear instruments and methods in physics research section B: beam interactions with materials and atoms*, 269(17):1910–1918, 2011. <https://arxiv.org/pdf/1106.0797.pdf> and https://en.wikipedia.org/wiki/Ratio_estimator.