NNT/NL : 2020AIXM0474/055ED352

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 16 décembre 2020 par

Marie Aubert

Contraintes cosmologiques avec les vides cosmiques dans eBOSS

Discipline
Physique et Sciences de la Matière

Spécialité
Astrophysique et Cosmologie

École doctorale 352
PHYSIQUE ET SCIENCES DE LA MATIERE

Laboratoire
Centre de Physique des Particules de Marseille

Composition du jury

Cristinel DIACONU AMU/CNRS	Président du jury
Nathalie PALANQUE-DELABROUILLE CEA-Saclay	Rapportrice
Jean-Paul KNEIB EPFL	Rapporteur
Hélène COURTOIS Université Claude Bernard Lyon I	Examinatrice
Pierre TAXIL AMU	Examinateur
Stéphanie ESCOFFIER AMU/CNRS	Directrice de thèse

Résumé

L'Univers est actuellement décrit par le modèle standard ΛCDM dont la contribution majeure vient de l'existence de l'énergie noire, un processus physique responsable de l'expansion accélérée de notre Univers. Un défi majeur de la cosmologie est de dévoiler la nature de cette substance inconnue en mesurant des effets subtils qui démontreraient une déviation du modèle standard établi. Dans ce but, les grandes structures de l'Univers sont d'un intérêt primordial pour la contrainte cosmologique, leur formation et leur croissance étant animées par l'expansion de l'Univers. En particulier, les vides cosmiques, des zones sous-denses étendues dans les structures à grande échelle, fournissent un environnement idéal pour étudier l'énergie noire, dont les effets devraient être dominants en leur sein. Cette thèse s'inscrit dans le cadre de l'étude des vides cosmiques comme une sonde prometteuse de l'énergie noire. Nous présenterons d'abord comment les vides sont détectés au sein des structures à grande échelle. Ensuite, nous extrairons les vides dans les dernières données du relevé spectroscopique SDSS-eBOSS et étudierons les distorsions de l'espace des redshifts autour de ceux-ci pour contraindre le taux de croissance des structures pour trois époques différentes. Enfin, nous aborderons l'application du test cosmologique Alcock-Paczynski en considérant les vides comme des sphères standards. Cette thèse met en avant le potentiel des vides comme une sonde discriminante de la cosmologie, tout en montrant les obstacles entravant l'étude de ces objets pour parvenir à une précision au niveau du pourcent dans les contraintes cosmologiques des futurs relevés.

Abstract

The Universe is currently described according to the standard ΛCDM model whose main contributions comes from the existence of dark energy, an unknown physical process responsible for the accelerated expansion of our Universe. The main challenge of cosmology today is to uncover the nature of this unknown substance by measuring the subtle effects that would evidence a deviation from the established standard model. To this end, the Large-Scale Structure is of prime interest as its formation and growth are driven by the expansion of the Universe and especially: cosmic voids. Cosmic voids are large under-dense zones within the Large-Scale Structure of the Universe. Being deprived of matter, they provide a perfect environment to study dark energy whose effects are expected to be dominant in their vicinity. This thesis is part of the study of cosmic voids as a promising dark energy probe. We will first present how cosmic voids are detected within large-scale structures. Then, we will focus on the extraction of voids in the latest eBOSSS data, part of SDSS, and the study of redshift-space distortions around them in order to constrain the growth rate of the structures at three different epochs. Finally, we will discuss the use of the voids as standard spheres that would, theoretically, directly constrain the energy content of the Universe through the application of the Alcock-Paczynski test. This thesis highlights the potential of voids as a discriminating probe of cosmology, while presenting the challenges that the study of these objects presents to achieve percent level precision in cosmological constraints in future surveys.

Remerciements

Avant toute chose, je tiens à remercier les membres de mon jury d'avoir accepté d'examiner ce manuscrit dans des temps assez mouvementés, notamment mes rapporteurs Nathalie Palanque-Delabrouille et Jean-Paul Kneib pour leurs retours et conseils sur le manuscrit ci-après.

Je remercie ma directrice de thèse, Stéphanie Escoffier, pour ses conseils scientifiques - mais pas que, pour les petits hacks de la vie d'un chercheur et pour toutes les opportunités qu'elle m'a offertes. Merci de m'avoir soutenue et supportée, d'avoir encouragé ma persévérance et d'avoir fait preuve de patience avec moi, qui suis relativement têtue.

Je tiens à souligner le support sans faille de Marie-Claude Cousinou durant toute cette entreprise, avec qui j'ai également travaillé en étroite collaboration. Je la remercie pour son aide précieuse et son enthousiasme à chaque avancée.

Je remercie également Alice et Adam pour leur bienveillance et leur entrain ainsi que pour nos échanges scientifiques stimulants.

Le groupe Renoir au sein du CPPM a été une super équipe d'accueil formée par des gens fort sympathiques dont je suis très heureuse d'avoir fait la connaissance. En particulier André Tilquin, dont j'ai partagé le bureau pendant quelques temps, que je remercie infiniment pour sa disponibilité et sa gentillesse. Je remercie également Smaïn pour la gazette CPPM, Julien pour ses 10 conseils à ne pas suivre **du tout** en rédaction de thèse (mais qui m'ont bien fait rire), Aurélia et Johanna mes collaboratrices pour l'organisation de chasses aux bonbons et autres animations. Merci à l'ensemble du groupe (anciens et nouveaux) pour les conversations, les discussions scientifiques, les blagues lors du café journalier après le déjeuner.

Je n'oublie pas tous les gens que j'ai croisés au détour d'un couloir, d'un café ou d'une pause clope et qui ont contribué à faire de ces trois années en leur compagnie un moment fort agréable. Et je leur présente toutes mes excuses de ne pas avoir pu fêter cette thèse autour d'un pot comme de coutume à cause des circonstances actuelles (j'en suis la première désolée).

Je ne pourrai jamais assez remercier mes chouchous, co-thésards et amis chers pour leur présence constante ces trois dernières années Philippe et Sylvain. Votre patience envers ma fâcheuse manie de poser des questions sans en écouter les réponses (oups) a été exemplaire.

Enfin, je remercie ma famille et les *copaings* pour leur confiance en moi dans mes moments de doutes, pour leurs encouragements et pour leur amour. Un énorme big up à mes colocs pour leur soutien ++ lors de ce sprint final.

Introduction

Cosmology, as we know, it today represents a recent development in modern physics. This field of physics, which originates from the beginnings of philosophy and its questioning of the laws of nature, have seen its renewal as a fundamental part of physics in the early 20th century. At this time, the formulation of the General Relativity by Albert Einstein provides a theoretical framework to describe the laws of our Universe, introducing a major paradigm shift in the consideration of the laws of Gravity. The massive content of the Universe bends the space-time metric. In tandem with the development of observations of the outskirts of our galaxy, the Universe is found expanding.

Cosmology has once again gone through a transformation in recent years. In the late 1990s is made the discovery that the Universe is currently undergoing an acceleration of its expansion. This is encoded in the measurement of a non-zero Λ cosmological constant contribution. In the decade that follows, the parameters that drive the expansion of the Universe are estimated with increasing precision, allowing to establish a standard flat Λ CDM model. These major discoveries, however, are shadowed by the fact that our Universe is dominated by two unknowns: dark matter and dark energy. The former has been confirmed since the 70's as a form of invisible matter that dominates the matter content of our Universe, the latter, is the name of the physical process encoded by the Λ constant.

While current researches regarding the nature of dark matter are slowly driven toward the domain of particle physics, dark energy remains, for now, a purely cosmological issue. Several scenarios have been advanced to explain the late-time cosmic acceleration which involve a modification of the laws of gravity on large scales or the existence of a fifth force, opposed to that of gravity, that would evolve slowly in time. These two scenarios present themselves in the form of very subtle observational effects, that would indicate a deviation from the standard ΛCDM model. Discerning such deviations from either General Relativity or ΛCDM have brought out a new era: that of precision cosmology. To satisfy the accuracy requirements on the cosmological constraints to be below the percent, cosmology has become a data-driven science that necessitates years-long surveys to acquire a large statistic. Another challenge undertook is the development of new probes that would enable cosmologist to remove the degeneracies between different observables and therefore tightening the constraints. In this framework, the Large-Scale Structure of the Universe represents an observable of choice to probe for dark energy and in particular, the cosmic voids within.

Cosmic voids are extended under-dense regions in the Large-Scale Structure of the Universe. Being devoid of matter, it is considered that their evolution and properties should be dominated by the action of dark energy, thus providing an environment of choice to research this quantity. With the advent of redshift surveys, the delineation of the Large-Scale Structure of the Universe has reached unprecedented finesse. The amount of statistic provided by the positions of galaxies has allowed the extraction of a significant number of voids to be used in cosmological analyses.

Cosmic voids have shown to be sensitive to dark energy or modified gravity through various aspects, such as their sizes, shapes or surrounding density field and as such, provide numerous probes of cosmology. However, while the advent of large-scale surveys has tremendously increased the statistics of detected cosmic voids, their identification is non-trivial nor are the physical processes that drive their growth. While being a high potential probe that could narrow down efficiently the cosmological constraint, their use as discriminating cosmological probes has to be investigated.

Two probes of cosmology, in particular, have been put forward in regard of the study of voids: the Alcock-Paczynski test, as a probe of dark energy and the study of redshift space distortions around voids as a probe of modified gravity. This manuscript investigates the use of cosmic voids in the context of precision cosmology, it is structured as follows: In the first chapter the theoretical context of modern cosmology is introduced, from the Einstein equations to the latest constraints on the cosmological parameters, as well as an overview of the linear depiction of the Large-Scale structure growth. The second chapter focuses on cosmic voids in the large-scale structures and details their identification, their properties and their potential in terms of cosmological probes. The third chapter addresses the constraint of the redshift-space distortions parameter in the final data release DR16 of the eBOSS survey with the void galaxy two-point cross-correlation. Finally, the last chapter investigates the use of the Alcock-Paczynski test on stacks of voids to recover the cosmological parameters.

Contents

1 Observational cosmology

Our current understanding of the Universe is based both on the theoretical modelling of large-scale gravity and on the observation of our Universe outside the boundary of our galaxy. Over the centuries, with the development of observation techniques, humanity has been able to probe farther and farther, within the appropriate spirit of: *the sky is the limit*. These observations, when compared to theory, have allowed us to discover some of the mysteries of the universe while opening doors to other unknown physical processes.

1.1 Hubble, Einstein and Friedmann: an expanding Universe

1.1.1 Foundations

1.1.1.1 Einstein Equations

At the heart of our current understanding of the observable Universe lies the gravitational interaction. This interaction, an attractive force between massive objects, has an infinite range affecting the most remote places of the Universe. The laws of gravitation were first developed in the 17^{th} century by Isaac Newton. They have proven to be a good description of the mechanics at work within our local reference frame that is the Earth and the solar system and have been at the heart of our understanding of the motion of objects for centuries. With his work on Special Relativity, Einstein helped to understand the importance of the reference frame, the coordinate system, when considering physical laws. A consequence is the introduction of a relative, rather than absolute, definition of time, which depends on the frame of reference adopted. One can no longer consider physical interaction in a spatial frame only but, has to consider a space-time reference frame.

A few years later, Einstein devised a general formulation of the laws of gravity: General Relativity. It corresponds to a formulation of the laws of gravity in regard to any reference frame encoded in what are now known as Einstein Field Equations:

$$G^{\mu\nu} := R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = 8\pi G T^{\mu\nu} + \Lambda g^{\mu\nu}. \tag{1.1}$$

The $G^{\mu\nu}$ tensor, which corresponds to the right-hand side of the equation (RHS) is itself a combination of the Ricci tensor $R^{\mu\nu}$ also known as the space-time curvature tensor, the scalar curvature R and the space-time metric tensor $g^{\mu\nu}$ which defines the reference frame adopted. The left-hand side is composed of the energy-momentum tensor $T^{\mu\nu}$, which relates to the energetic content of the Universe, Λ, an arbitrary constant called the cosmological constant and G, the gravitational constant.

The consequence of these equations is that the very definition of the reference frame considered and the energetic content of the Universe affects one another. The energetic content affects the general geometry of the Universe, curving the space-time metric. A second consequence of this formulation is that it implies a dynamic universe, an evolving Universe. The arbitrary Λ constant was at first added by Einstein to the equations as a way to counteract this implied evolution of the Universe to depict a static Universe. This addition was later on qualified by himself as *'its biggest blunder'*.

I leave the Λ constant in the formulation of these equations as a free parameter. If positive, this constant acts as a repulsive force, opposite to that of gravity, if it is negative, it becomes an additional contribution to the gravitational interaction. If null, then, the evolution of the Universe is solely governed by the energetic content carried by the $T_{\mu\nu}$ tensor.

This general formulation, *a priori* quite simple provides a theoretical framework to model the laws governing the Universe. In order to solve these equations, however, a metric $g_{\mu\nu}$ has to be defined. This definition is derived from general assumptions and observations of our Universe.

1.1.1.2 Cosmological principle

A key principle adopted to depict our Universe is that it obeys the *cosmological principle*. The cosmological principle states that the Universe is both homogeneous and isotropic. The homogeneity assumption is an extension of the Copernican principle, which simply says that Earth is not the centre of the Universe, to all of the objects found in the Universe. *There are no privileged positions in the Universe*, which means that the positions in the Universe are invariant by translation.

In contrast, the isotropy assumption was derived from the observation of the night sky. In any direction an observer can look, the distribution of stars and objects doesn't seem to be different. This observation led astronomers to postulate isotropy that is, *there are no privileged directions in the Universe*, inferring that positions are also invariant by rotation.

1.1.1.3 Receding galaxies

While Einstein imagined a static Universe and therefore added a cosmological constant to impose his view, a fundamental observation questioned and buried the idea of a static Universe. The observation of nebulae in the night sky was the subject of two major discoveries. Firstly, it was evidenced that nebulae were extra-galactic objects similar to our own Milky Way (Slipher, 1917). Secondly, further observations of these objects allowed to estimate both their distance and dynamics (Slipher, 1913) uncovering a linear relation between the distances of the galaxies and their velocities (Hubble, 1929; Lemaître, 1927; Slipher, 1917). The farther the galaxy, the faster it seemed to recede from us. This relation has come to be known as the Hubble-Lemaitre law:

$$v = H_0 d, \tag{1.2}$$

where v is the recession velocity of the galaxy, d is its distance and H_0 is the Hubble constant which corresponds to the rate at which the galaxy recedes from us. This relation consisted in the evidence of an expanding Universe. The idea of a static Universe and the associated non-zero cosmological constant were abandoned. The interpretation of the H_0 constant changed to depict the rate at which the Universe is expanding today.

As the extra-galactic objects are observed thanks to their luminous nature, it results that the wavelengths of the emitted photons are influenced by the recession velocities of the galaxies. These introduce a form of Doppler shift toward higher and redder wavelength: *the redshift*. This redshift can be quantified by comparing the observed wavelength λ_{obs} to the corresponding original emitted wavelength λ_{em}:

$$1 + z = \frac{\lambda_{\mathsf{obs}}}{\lambda_{\mathsf{em}}} \tag{1.3}$$

1.1.1.4 Friedman-Lemaitre-Roberston-Walker Metric

Solving the Einstein equations requires a definition of the reference frame $g^{\mu\nu}$. The metric has to be formulated so according to our knowledge of the Universe. The expansion of the Universe is already encoded in the Einstein equations, therefore the metric has to comply to the cosmological principle. The general form of the metric respecting these prescriptions is then the Friedman-Lemaître-Roberston-Walker metric:

$$ds^2 = dt^2 - a(t)^2 [\frac{d\chi^2}{1 - k\chi^2} + \chi^2 d\theta^2 + \chi^2 \sin^2\theta d\phi^2], \tag{1.4}$$

where $a(t)$ is the scale factor of the Universe and k is indicative of the curvature of space. Homogeneity of the Universe leads to the introduction of $a(t)$ which accounts for the change of the physical distance due to the expansion of the Universe. By convention, the scale factor at present day t_0 is set to unity.

Considering the expansion today and at an arbitrary time t, we can now properly define the cosmological redshift:

$$1 + z = \frac{a(t_0)}{a(t)} = \frac{1}{a(t)} \tag{1.5}$$

As such, the cosmological redshift is an observable that can be measured thanks to the relation given by Eq. 1.3.

The introduction of the curvature of space k in the spatial coordinates comes from the isotropy requirement. Only three values of k comply to the cosmological principle:

$$k = 1, \qquad \text{spherical} \quad \text{Closed Universe} \tag{1.6}$$
$$k = 0, \qquad \text{flat} \quad \text{Open Universe} \tag{1.7}$$
$$k = -1, \qquad \text{hyperbolic} \quad \text{Open Universe} \tag{1.8}$$

As such, the curvature defines the geometry of the Universe which can be either Spherical, Flat or Hyperbolic and therefore depict either a finite universe (closed) or an infinite universe (open).

1.1.2 Friedmann equations

With the metric defined by Eq. 1.4, it is then possible to express the geometrical part of the Einstein equations, the $G^{\mu\nu}$ tensor considering an isotropic, homogeneous and expanding Universe. A detailed discussion on the derivation of the Einstein equations and their geometrical part can be found in Weinberg (2015). The $G^{\mu\nu}$ tensor can be expressed in terms of its temporal G^{00} and spatial part G^{ii}, which are the only non-zero contributions in this case:

$$G^{00} = 3\left(\frac{\dot{a}}{a}\right)^2 + \frac{3k}{a^2}, \tag{1.9}$$
$$G^{ii} = (2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2})g^{ii}. \tag{1.10}$$

Eq. 1.9 and 1.10 describe the geometrical part of the Einstein equations, namely how the metric is affected through the expansion of the Universe. Recalling the Einstein equations (1.1), this geometrical part both affects and is affected by the energetic content of the Universe whose information is encoded in the tensor $T^{\mu\nu}$.

Considering all that is in the Universe as a perfect fluid, with isotropic pressure, the energy-momentum tensor of the energetic content of the Universe can be written as a diagonal tensor:

$$T^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & -P & 0 & 0 \\ 0 & 0 & -P & 0 \\ 0 & 0 & 0 & -P \end{pmatrix}, \tag{1.11}$$

where ρ corresponds to the density and P is the pressure of the fluid. Thus, we have:

$$T^{00} = \rho g^{00} = \rho \tag{1.12}$$
$$T^{ii} = -P g^{ii}, \tag{1.13}$$

which, through Einstein's relation yield:

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G\rho}{3} + \frac{\Lambda}{3}, \tag{1.14}$$
$$(2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2})g^{ii} = -8\pi G P g^{ii} + \Lambda g^{ii}. \tag{1.15}$$

Combining the above equations, we recover the Friedmann equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{k}{a^2}, \tag{1.16}$$
$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) + \frac{\Lambda}{3}. \tag{1.17}$$

We define $\frac{\dot{a}}{a} = H$, the Hubble parameter H which gives us the expansion rate of the Universe at time t.

These equations represent the most general solutions to the Einstein equations obeying homogeneity, isotropy and expansion as conditions. However, in this general framework, they cannot be solved while the relation between the energy density ρ and the pressure P of the content is not formulated.

1.1.3 Dynamic contribution of the content

The Friedman equations are valid for any type of expanding Universe respecting the cosmological principle. The solutions to these equations necessitate an inventory of the content of the Universe and the relation between their density and pressure in order to expand Eq. 1.17.

From Eq. 1.16, we can denote three types of contributions driving the expansion of the Universe:

– **Matter** encoded in the existence of the $\rho(t)$ term which is affected by gravity. Two species of matter exist: *ultra relativistic*, near massless matter such as the radiation (photons, neutrinos, etc) and *non-relativistic* baryonic matter.

– **Curvature** carried out in the k dependent term describes the curvature of space-time, in the event of a non-flat universe.

– **Cosmological constant**, $\boldsymbol{\Lambda}$, in the event of a non-zero constant.

While all of these species are not encoded in the energy-momentum tensor $T^{\mu\nu}$, the existence of curvature in the metric or of a non-zero cosmological constant have repercussions on the general fate of the Universe. Considering the contributions listed above as different perfect fluids, the relation between pressure and density is of the form:

$$P = w\rho \tag{1.18}$$

where w is the equation of state of the considered content.

The energy-momentum conservation law in an expanding Universe goes as follows:

$$a^{-3}\frac{d(\rho a^3)}{dt} = -3\frac{\dot{a}}{a}P. \tag{1.19}$$

So, combined with Eq. 1.18, the evolution of density in terms of the expansion can be expressed as:

$$\rho(t) = \rho_0 a^{-3(w+1)}, \tag{1.20}$$

where ρ_0 is the density $\rho(t_0)$ at present time. Isolating each specie to an associated density, pressure and therefore equation of state allows to identify the evolution of their energy density in terms of the expansion.

The matter content is non-relativistic and considered as collisionless at cosmological scales, thus $w = 0$. Radiation, on the other hand, is ultra-relativistic and thus suffers collisions, therefore $w = 1/3$. The curvature can be considered as a fluid with density $\rho_{k,0}$ defined as :

$$\rho_{k,0} = \frac{-3k}{a^2 8\pi G}, \tag{1.21}$$

which yields a parametric equation of state of $w = -1/3$. Finally, a supposed Λ will have an equation of state $w = -1$ as it is constant throughout the expansion of the Universe. As a result, we have:

$$\rho_r \propto a^{-4}, \tag{1.22}$$
$$\rho_m \propto a^{-3}, \tag{1.23}$$
$$\rho_k \propto a^{-2}, \tag{1.24}$$
$$\rho_\Lambda \propto 1. \tag{1.25}$$

and the Friedmann equation (1.16) can thus be expressed in terms of the different content of the Universe:

$$H^2 = \frac{8\pi G}{3}(\rho_{r,0}a^{-4} + \rho_{m,0}a^{-3} + \rho_{k,0}a^{-2} + \rho_{\Lambda,0}). \tag{1.26}$$

Let us consider the critical density ρ_c, corresponding to a threshold density at which the Universe halts its expansion and is considered flat ($k = 0$.) This density, evaluated at present-day then takes the value:

$$\rho_c = \frac{3H_0^2}{8\pi G}. \tag{1.27}$$

The sum of all contributions $\rho_{i,0}$ should thus equate the critical density. This allows us to parametrize the densities ρ_i in terms of their rescaled quantities $\Omega_i = \rho_i/\rho_c$. The total energetic content becomes:

$$\Omega_{tot} = \begin{cases} \sum_i \Omega_{i,0} > 1 & \Omega_k < 0, \text{ Spherical/Closed} \\ \sum_i \Omega_{i,0} = 1 & \Omega_k = 0, \text{ Flat/Open} \\ \sum_i \Omega_{i,0} < 1 & \Omega_k > 0, \text{ Hyperbolic/Open} \end{cases} \tag{1.28}$$

The Hubble parameter can be rewritten in terms of the dimensionless densities and H_0:

$$H^2 = H_0(\Omega_{m,0}a^{-3} + \Omega_{r,0}a^{-4} + \Omega_{k,0}a^{-2} + \Omega_{\Lambda,0}) \tag{1.29}$$

The above equation expresses the expansion rate in terms of dimensionless quantities. Expressing the scale factor in terms of the redshift, we obtain the following expression of the expansion rate:

$$H(z) = H_0\sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{r,0}(1+z)^4 + \Omega_{k,0}(1+z)^2 + \Omega_{\Lambda,0}} \tag{1.30}$$

In practice, one express H_0 as follows:

$$H_0 = h100 \text{ km}\,\text{s}^{-1}\text{Mpc}^{-1} \tag{1.31}$$

where h is the reduced Hubble constant.

The behaviour of our Universe can thus be described by a singular association of the cosmological parameters $H_0, \Omega_{0,m}, \Omega_{0,r}, \Omega_k$ and Ω_Λ which we will call through the misnomer: cosmology. The cosmology can only be inferred through observational means.

1.1.4 Distances

Throughout the section, it was mentioned that the metric of the Universe is embedded in the framework of an expanding Universe. This time-evolving spatial coordinate system raises the question of the distance definition and measure in our Universe: how should we quantify the distance when comparing the distance of an object A to an object B?

1.1.4.1 In regard to the metric

The spatial part of the metric depends on both a, the scale factor and χ the radial coordinate. The latter is set to be an expansion independent distance, along the line of sight which we define as **comoving distance**:

$$\chi = c\int_t^{t_0} \frac{dt}{a(t)} \tag{1.32}$$

We use χ from now on to mention the comoving distance. In practice, it is more convenient to express χ in term of the redshift. This yield:

$$\chi = c\int_0^{z} \frac{dz}{H(z)} \tag{1.33}$$

This distance, however, does not amount to the true euclidean distance in the presence of a non-null curvature of the metric. .

Indeed, let us consider object A and B, both situated at a distance χ from us but separated by a wide angle $\Delta\theta$. In regard to us, both object will present the same radial comoving distance as it is directly the distance along the line-of-sight. Were we to consider the comoving distance between A and B, then, their distance would change according to the curvature of our universe. This **transverse comoving distance** will thus be defined as follows:

$$D_M = \begin{cases} D_H\sqrt{|\Omega_k|}\sin\left[\sqrt{|\Omega_k|}\chi/D_H\right] & \Omega_k < 0, \\ \chi & \Omega_k = 0, \\ D_H\sqrt{|\Omega_k|}\sinh c\left[\sqrt{|\Omega_k|}\chi/D_H\right] & \Omega_k > 0, \end{cases} \tag{1.34}$$

where $D_H = \frac{c}{H_0}$, the Hubble distance. It is to be noted that we recover the Euclidean distance in the case of a flat Universe. To refer to the comoving distance, the notation $D_c = \chi$ will be used henceforth.

The two distances presented above, however, do not relay the physical distance between the two A and B objects. If an observer out of the Universe were to take snapshots along the expansion, the true distance between the object should have changed. This physical distance between the object, or **proper distance** goes as follows:

$$d_p(t) = a(t)D_c, \tag{1.35}$$

Considering the velocity of the object measured with Eq. 1.35, it is possible to recover an expression of the recession velocity of a considered object:

$$v = \frac{d\,d_p(t)}{dt} = \frac{da(t)D_c}{dt} = \dot{a}D_c = H d_p(t), \tag{1.36}$$

which is similar to the Hubble law. For sufficiently small redshifts, it can be considered that objects are nearly simultaneous to us. It follows that $d_p \sim D_c$ and $H \sim H_0$, allowing us to recover the Hubble law:

$$v = H_0 d_p. \tag{1.37}$$

1.1.4.2 In regard to observation

The main information of a distant object is the light emitted by it and the subsequent photon flux that travels up to us. It allows us to see the object and decipher its general shape provided that it is close enough. The information thus provided by astronomical observations amount to the luminosity L and an angular diameter $\delta\theta$.

The angular diameter distance is thus defined as:

$$D_A = \frac{D_M}{1+z} \tag{1.38}$$

where D_M is the transverse comoving distance, and the (1+z) term comes from the fact that the apparent distance subtended $\delta\theta$ between the extremities of the object will be also be affected by the expansion.

Let us now consider the luminosity of the object. The flux of photons which arrive to the instruments is dependent on the distance of the source, thus:

$$F = \frac{L}{4\pi d_L^2}, \tag{1.39}$$

where d_L is the luminosity distance. The luminosity is defined as:

$$d_L = (1+z)D_M \tag{1.40}$$

1.2 The Standard model of cosmology

The expansion history of the Universe is defined as a function of several parameters whose association describe a single Universe. The information relayed by different observations seems to converge toward a single favoured model which is called Λ-CDM. The parameter Λ stands for a non-zero cosmological constant and 'CDM' for Cold Dark Matter, an unknown kind of invisible matter but nevertheless highly massive with non-negligible impact on the observable Universe. Recent constraints on cosmological parameters are given by the 2018 Planck collaboration results (Aghanim et al., 2020c):

$$\Omega_m = 0.3111 \pm 0.0056, \tag{1.41}$$
$$\Omega_\Lambda = 0.6889 \pm 0.0056, \tag{1.42}$$
$$\Omega_k = 0.001 \pm 0.002. \tag{1.43}$$

As a consequence, our Universe can be considered flat because the curvature is compatible with $k = 0$ and would be mainly dominated by the so-called *dark sector*, which qualifies the unseen quantities represented by the dark matter and the constant Λ. The latter reflects the fact that our Universe is currently experiencing a late-time acceleration of its expansion. The existence of the dark sector represents a great challenge for cosmology: the measured Ω quantities tell us nothing about their nature and their underlying physical process.

1.2.1 Dark matter

The first observations of the velocities of galaxies made by F. Zwicky (Zwicky, 1933) within the Coma cluster gave a curious result: the underlying mass of the cluster estimated from the velocities of the galaxies implied a quantity of galaxies much larger than that estimated by the quantity of visible light in the cluster. This led to the hypothesis of the presence of a kind of dark – non-visible – matter in the clusters. Later, the observation of the rotation curve of spiral galaxies by Rubin & Ford (1970) reinforced this idea at the scale of galaxies. Indeed, the velocity of stars is related to both the mass of the galaxy and their distance from the centre of the galaxy. However stars at the periphery of these galaxies rotated as fast as those located at the centre, as shown in Fig. 1.1, while the velocity was expected to decrease in $\propto r^{-1/2}$.

This implied the presence of an invisible massive contribution: dark matter. This dark matter would take the form of a non-luminous halo surrounding the galaxies and clusters (Ostriker et al., 1974)

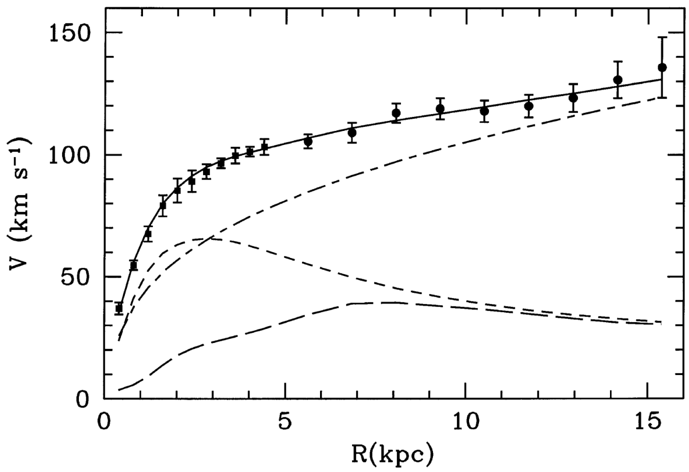


Figure 1.1: M33 rotation curve observed with 21cm meter estimation (full circles) as well as the theoretical prediction in dark matter paradigm (full line). The contributions from the stellar disc (normal dashed lined), the gas (long dashed line) and dark matter halo (dot-dashed line) are also displayed (Corbelli & Salucci, 2000).

A final consecration to the existence of these dark matter haloes in the neighbourhood of the galaxies is provided by the observation of lensing phenomena. Photons, as per their nature of massless particles, travel along the geodesics of the space-time metric taking the shortest path. General Relativity tells us that gravity and thus mass bends the fabric of space-time causing photons to deviate from their original paths. The amount of deflection of the photons' trajectory is proportional to the amount of mass contained by the cluster. Such lensing effect, when detected in its strongest form (and rarer, unfortunately), can thus relay an indication of the total amount of matter contained within a structure. From this total mass can be removed that of the visible contributions, yielding constraints on the fraction dark matter present in the object (Massey et al., 2018; Tyson et al., 1990)

In a generalized view, if the dark matter field is akin to that of the galaxies, then it affects all the photons emitted by the observed galaxies on their way to our cameras. As a result, it is possible to recover a statistical depiction of the dark matter field through the use of this form of lensing, called *weak lensing* (Bartelmann & Schneider, 2001; Kaiser & Squires, 1993).

(a) **(b)**

Figure 1.2: (a) Strong lensing effect in the A3827 cluster with inferred dark matter halo distribution in blue isocontours, (b) Lensing imprint in the Bullet cluster in the form of convergence map (red) and sigma map (blue), Credits: (a) NASA/ESA HST, composite image by Massey et al. (2018) , (b) NASA/ESA HST composite image by Clowe et al. (2006)

The presence of non-luminous, albeit massive, quantity suggested that dark matter (DM) had to be collisionless interacting primarily through gravitation. As per its nature of matter, the particles that constitute DM are defined on the point of view of particle physics. Three leading hypothesis were advanced to describe such matter:

– **Cold dark matter (CDM)**: Non-relativistic particles that interact weakly with ordinary matter are the preferred candidates for dark matter. The most popular candidates for CDM (Frenk & White, 2012) are WIMPs (Weakly interacting Massive Particles) with masses $\sim 0.1 - 10$ TeV. They are predicted in the context of beyond standard model particle physics, for example, in supersymmetry as lightest supersymmetric particles (LSP) or in extra-dimensional theories as lightest Kaluza-Klein particles (LKP). Finally, axions are theoretical particles introduced to solve the lack of observed CP violation in the strong interaction, with very light masses < 0.1 eV (Blumenthal et al., 1984).

– **Warm dark matter (WDM)**: Well-motivated elementary particle candidates that have appreciable thermal velocities at early times, could behave as warm dark matter. The best-known example is a sterile neutrino with a particle mass of order a keV (Asaka & Shaposhnikov, 2005; Kusenko, 2009).

– **Hot dark matter (HDM)**: Ultra Relativistic particles of very low mass that interact very weakly. The most likely candidates in this framework are the three flavour of neutrinos (ν_e, ν_μ, ν_τ) which have evidenced oscillations suggesting massive particles (Fukuda et al., 1998).

A final explanation to the observation of this dark matter in the overall distribution of matter has been advanced in the form of a modification of Newton's dynamics like the MOND theory (Milgrom, 1983). The latter has been ruled out by observational evidence of lensing in the Bullet cluster whose effects were not predicted by MOND (Clowe et al., 2006).

Dark matter represents the major contribution of matter in the Ω_m energy density term, amounting to $\sim 80\%$. The remaining 20% are attributed to ordinary matter. Observations such as the large-scale distribution of matter and fluctuations in initial density, which will be discussed later, have shown a preference for the cold dark matter paradigm (CDM).

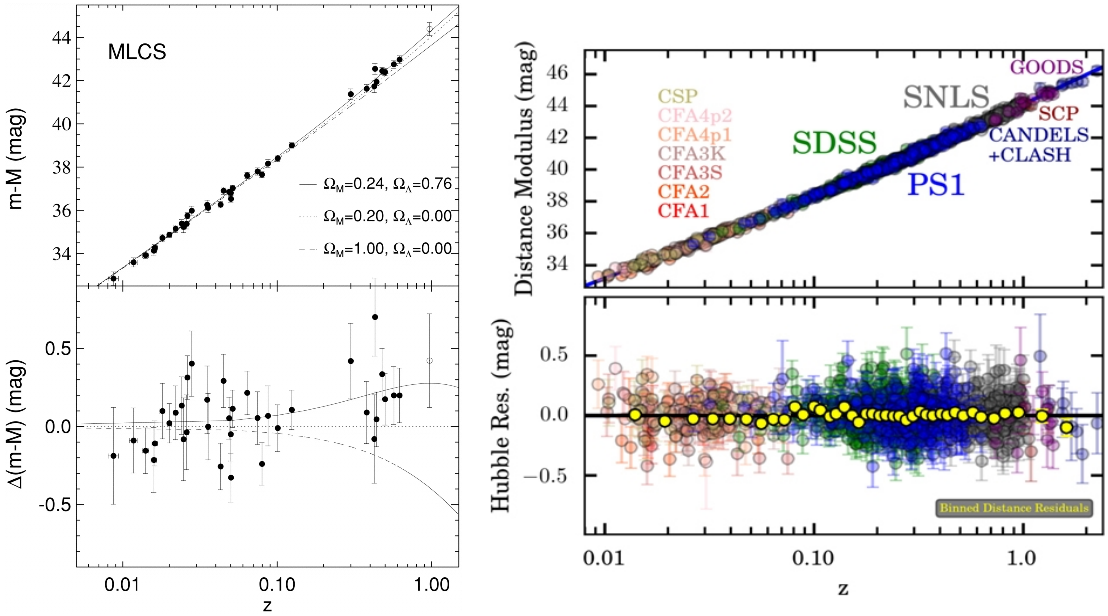
1.2.2 Dark energy

While the dark matter contribution of the dark sector has been a long-standing assumption regarding the content of our Universe, the discovery of a dark energy component is quite new. In the 1990s, the view of a matter-dominated Universe is no longer in favour of most of the observation, such as the age of the oldest clusters (Ostriker & Steinhardt, 1995) or preliminary results of the temperature fluctuations of the CMB (see next section).

In this context, the study of the type Ia supernovae (SNIa), thermonuclear SN, uncovered what seemed to be a late-time cosmic acceleration of the expansion of the Universe. SNIa stems from a binary system composed of a white dwarf and a giant star: the white dwarf through gravitational interaction accretes the gas released by the giant star up until its mass exceeds $1.4\mathrm{M}_\odot$, the Chandrasekhar mass. At this point ensues a thermonuclear explosion: the Supernovae of Type Ia. The light emitted as a result of this process generally follows the same profile and display the same intrinsic luminosity, regardless of the epoch. In this sense, they are called standard candles.

A consequence of the observation of the SNIa is that, when considering the flux of the luminous object as seen in Eq. 1.40, an acceleration of the expansion rate of the Universe will cause the luminosity of the supernovae to appear fainter in redshift or brighter in the event of a deceleration of the expansion rate.

To this end, both the Supernovae Search Team and the Supernoave Cosmology Project observed the apparent luminosity of SNIa at several redshifts in order to reconstruct the luminosity to redshift relation, dependent on cosmology. In their analysis, it became apparent that there was a non-negligible contribution in the form of a non-zero dominant Λ evidencing a late-time acceleration of the expansion of the Universe (Perlmutter et al., 1999; Riess et al., 1998). As can be seen in Fig.1.3a, the observed Hubble diagram of the supernovae favours a non-zero Ω_Λ configuration.



(a) Discovery　　**(b)** Latest result

Figure 1.3: 1.3a Hubble diagram of the early Supernovae Search Team (Riess et al., 1998), 1.3b displays the most up-to-date Hubble diagrams of SNIa (Scolnic et al., 2018). Lower panels both show the residuals between the best-fit model and the data.

More recent results on both SNIa and other probes (see next section) have established the Λ constant as a contribution in the overall expansion of the Universe. The latest constraints concerning the SNIa light-curve, displayed in Panel 1.3b, are brought by the Pantheon sample defined by 1048 SNIa (Scolnic et al., 2018). The matter-energy density, within a flat ΛCDM framework ($\Omega_\Lambda = 1 - \Omega_m$) is estimated at $\Omega_m = 0.307 \pm 0.012$ and the Dark Energy equation of state amounts to $w = -1.026 \pm 0.041$, in accordance to a cosmological constant $w = -1$. However, the explanation of a cosmological constant as the source of an acceleration of the Universe poses the well known cosmological constant problem. Considering Λ as vacuum energy density, the measured contribution in cosmology is 120 order of magnitudes too low compared to the predictions on the point of view of quantum mechanics. As a result, the physical process underlined by this non-zero Λ constant is known as an unexplained *dark energy*.

Attempts have been made to model such dark energy which can be separated in two kinds: Dynamical dark energy and their likes which attempt to conserve the idea of an energetic contribution leading to a late-time cosmic acceleration and modifications of the law of gravity at large scale which involve a modification the $G^{\mu\nu}$ tensor.

In the following paragraphs are given a broad depiction of both kind of theories, detailed descriptions can be found in Sami (2007) and Yoo & Watanabe (2012):

Dynamical dark energy This modelling in its general form involves the use of a time-evolving scalar field ϕ and its potential $V(\phi)$ to describe dark energy as a time-evolving field acting as an opposite to the gravitational field. In their most simple forms, these models assume a spatially homogeneous field which would propagate, in the framework of GR and FLRW metric, as follows:

$$\ddot{\phi} + 3H\dot{\phi} + \frac{dV}{d\Phi} = 0 \tag{1.44}$$

Such modelling generally yield an observable that can be quantified in terms of a time-evoluting equation of state:

$$w_\phi = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\ddot{\phi}^2 - V(\phi)} \tag{1.45}$$

In order to account for the evolution of the dark energy equation of state throughout the expansion, it is convenient to use the Chevallier-Polarski-Linder parametrization (Chevallier & Polarski, 2001; Linder, 2003):

$$w(z) = w_0 + (1-a)w_a = w_0 + \frac{z}{(1+z)}w_a \tag{1.46}$$

As such, it is possible to define bounds on the value of the equation of state today, denoted w_0, to be situated in the range $-1 < w_0 < -1/3$ where the high boundary is directly inferred from Eq. 1.20 as the exclusive maximal value which guarantees a late-time acceleration of the expansion. The w_a quantity, however, varies in time.

Such classes of dynamical dark energy models are called quintessence. The resulting new definition of the dark energy equation of state leads to the following definition of the dark energy density: $\rho_{DE} = \rho_{DE} a^{-3(1+w_0+w_a)} e^{-3w_a(1-a)}$.

These classes of models, as they allow for several parametrization of ϕ and $V(\phi)$ and subsequent equation of state, remain in line with general relativity theories.

Alternative or Modified Gravity model In this case, instead of considering a modification of the energetic content, one considers a direct modification of the definition of the geometrical part of the Einstein field equations, changing the laws of gravity themselves on large scales. Some involve the addition of extra-dimensions that would cause a change of the general laws of gravity on large scales, such as the DGP brane-world model (Deffayet et al., 2002; Dvali et al., 2000). Others consider the introduction of an arbitrary function of the scalar curvature $f(R)$ (De Felice & Tsujikawa, 2010; Sotiriou & Faraoni, 2008) that involves the apparition of higher-order terms in the formulation of the Einstein equations. Such models are bound by the obligation to recover the Newton gravity at small scales.

Alternative-theories to the Λ constant should lead to fine observable effects in the form of the equation of state of dark energy or deviations from direct prediction of Einstein's gravity as presented above. A major problem remains that, if there were to be modifications of the properties of dark energy or of the gravitational interaction itself, current measurements do not allow to discriminate from a Λ constant framework.

1.2.3 Test of the standard model

1.2.3.1 Early Universe: Cosmic Microwave Background

An expanding Universe implies that it went through, in its early stages, a very dense and hot period during which matter and radiation alike were part of a plasma. This density field was considered to be generally smooth, composed of small perturbations. The photons trapped in the plasma pushed matter into gravitational potential wells while matter, in turn, resisted pushing the radiation outward. As a result, the baryon-photons interactions in the primordial universe led to contraction-dilation processes of the over-densities, causing perturbations in the plasma to propagate like sound waves: the baryonic acoustic oscillations (BAO) (Hu & Sugiyama, 1995).

In time, the Universe cooled down causing it to become less dense. Photons were suddenly released from the primordial plasma carrying out an imprint of the initial density fluctuations and baryons were free to assemble to form the first atoms. This period is called recombination (or decoupling). The resulting radiation was first predicted by Gamow, Alpher and Bethe (Alpher et al., 1948; Gamow, 1948) in the late 1940s, but unfortunately discarded. At the time, cosmology was considered as a nearly metaphysical science, due to the lack of experimental validation of the theories. Years later, however, the prediction of isotropic radiation originating from an early dense Universe was recovered by Dicke et al. (1965) in an independent manner. The latter having been emitted so early in time should be redshifted to reach the microwave wavelength. The same year, a corresponding signal was detected as white noise through an antenna, in the microwave regime (Penzias & Wilson, 1965).

The existence of relic radiation which had travelled throughout the expansion was a powerful incentive to understand our Universe from its early stages. Due to the influence of the atmosphere smearing the signal, space missions were sent in order to properly characterise the black-body signal. The COBE (Mather et al., 1994) experiment measured a near-perfect black-body spectrum of a temperature of 2.73 K today at a wavelength of the microwave, however, such spectrum presented anisotropies in its spectrum of the order $\frac{\delta K}{K} \sim 10^{-5}$ (Smoot et al., 1992). The measured fluctuations of temperatures in the spectrum were actually the remnants of the initial density fluctuations. The analysis of the statistical properties of these fluctuations thus enabled to gain knowledge on both the early Universe but also, the value of the cosmological parameters today (Hu et al., 1997).

The first inkling toward a flat Λ-CDM model was brought by the BOOMERANG (Hanany, 1997) experiment which consisted in balloons sent high in the atmosphere to map the fluctuations of the density. The study of the resulting spectrum favored strongly an $\Omega_k = 0$ universe (de Bernardis et al., 2000). Two space missions dedicated to more precise measurements of the cosmic microwave background temperature fluctuations followed. First, WMAP (Bennett et al., 2013; Hinshaw et al., 2013; Komatsu et al., 2011) provided the first evidence the Λ-CDM model to be obviously favoured by multiple probes and later, the Planck mission, which released its last cosmological results in 2018 (Aghanim et al., 2020a,c).

The study of the temperature fluctuations of the CMB, in the form of the angular power spectrum, yielded very precise constraints on the values of the cosmological parameters. As seen in Fig. 1.4, the power spectrum features oscillations in its spectrum which directly relate with the initial baryonic acoustic oscillations happening in the early plasma. The first peak is directly dependent on the energy density Ω_m and the overall spectrum is also dependent on Ω_b, the total amount of baryonic matter.

The existence of such a cosmic microwave background provided several confirmations concerning our Universe. A first validation was the existence of primordial density fluctuations which are considered to be the origin of the large scale distribution of matter observed today. The validity of the cosmological principle at large scale was also confirmed and finally, it enabled to assume that the Universe has a flat geometry.

1.2.3.2 Baryonic Acoustic Oscillations in galaxy clustering

With the expansion and gravity, the initial density fluctuations present in the Universe evolved to form a peculiar large scale matter distribution. These present a similar feature in their power spectrum as to the Cosmic Microwave Background radiation. As the Universe cooled down, so did the plasma in which the perturbations propagated up until the baryons stop interacting with photons: the drag epoch. At this time, gravitational interaction led to dark matter following the behaviour of the baryon plasma, yielding a distinctive feature in the overall density field: BAO. The latter being embedded directly in the density field and subject to no observational effects that affect the photons of the CMB (Eisenstein & Hu,

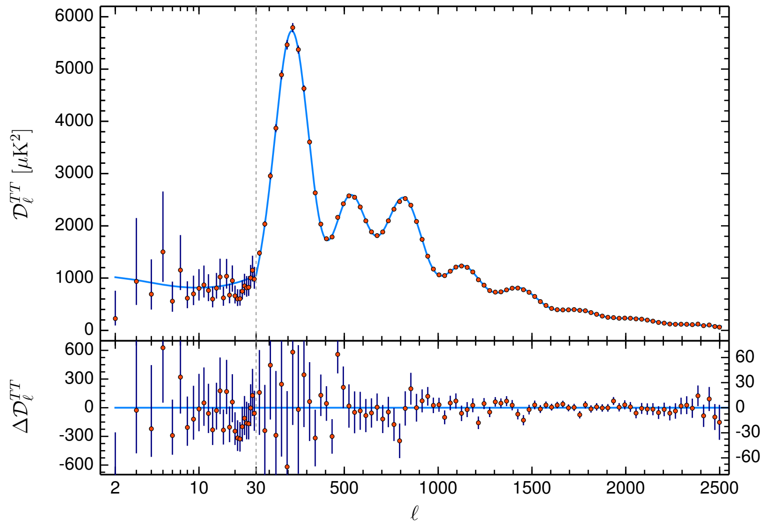


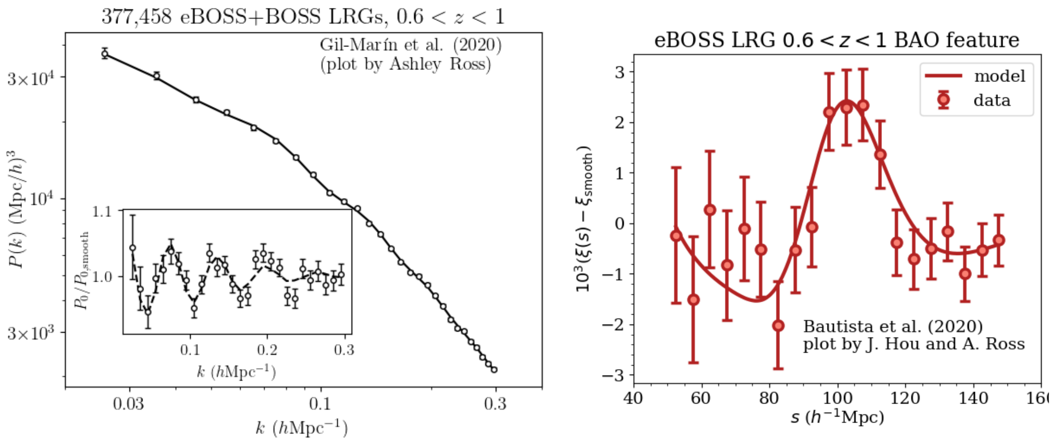
Figure 1.4: Upper panel: In red, estimated angular power spectrum of the temperature fluctuations of the cosmic microwave background of the Planck 2018 data release. In blue is shown the best fit theoretical prediction from a $\Lambda - CDM$ model. Lower panel: Residuals in regards to the model prediction. Errors displayed are the 1σ errors which include the cosmic variance. Credits: (Aghanim et al., 2020b)

1998).

This freeze of the BAO feature in the density field led to statistical over-density found at a specific scale of $\sim 100h^{-1}$Mpc , thus providing a standard ruler embedded in the general density field. The comoving distance is directly dependent on cosmology, as such, the comparison of this standard scale to the predicted standard ruler enables to place constraints on the cosmological parameters. Furthermore, the information on the galaxy density field is a late-time probe which can give interesting constraints on the expansion rate of the Universe or look for eventual discrepancies between early and late time measurements.

The study of the BAO feature is a recent development in cosmology. This excess density pattern was first noticed by Broadhurst et al. (1990) as a puzzling feature in the distribution of matter at large-scale for which an explanation was given later on by Eisenstein et al. (1998). The first significant detection was made in the Luminous Red Galaxies (LRG) sample of the SDSS-I program (Eisenstein et al., 2005) and has since been a powerful and robust probe of cosmology, contributing to precise measurements at different epoch allowing to extract the cosmological parameters, consistent with a flat ΛCDM cosmology (Alam et al., 2017, 2020b). Fig. 1.5 displays the most recent measurement of the BAO feature in both Fourier and Configuration space for the LRG sample of the Sixteenth Data Release (DR16) of the extended-Baryon Oscillation Spectroscopic Survey (eBOSS)[1]. The BAO feature in Fourier

[1]https://www.sdss.org/science/final-bao-and-rsd-measurements/



(a) Power spectrum　　**(b)** Two-point correlation function

Figure 1.5: Baryonic Acoustic Oscillation feature in the density field of the eBOSS DR16 Luminous red galaxy (LRGs) sample in (a) Fourier space and (b) Configuration space. Circles represent the measured data and the full line correspond to the Λ-CDM model. Credits: (Bautista et al., 2020; Gil-Marín et al., 2020)/SDSS/eBOSS Collaboration.

space is shown in the embedded subpanel of Panel 1.5a, marked by the wiggles at the lower scales of the measured power spectrum between $k = 0.1\ h\mathrm{Mpc}^{-1}$ and $k = 0.3\ h\mathrm{Mpc}^{-1}$. In configuration space, the BAO is seen as a near gaussian density excess around $\sim 100\ h^{-1}\mathrm{Mpc}$.

1.2.3.3 Cosmic Ladders

The probes presented above are targeted toward the determination of the energetic content of the Universe. However, the expansion rate of the Universe depends on the value of the Hubble parameter today H_0, which is primordial to trace its expansion history. To this end are generally used cosmic ladders, that is, probes of the distance that are generally independent of the cosmological parameters. The estimation of the H_0 constant can be done in the early Universe through the study of the CMB which yield $H_0 = 67.4 \pm 0.5\mathrm{km.s}^{-1}$ in the Planck analysis (Aghanim et al., 2020a) and $H_0 = 67.6 \pm 1.2\mathrm{km.s}^{-1}$ for the ACT-WMAP (Aiola et al., 2020) analysis (both in combination with BAO measurements as well as lensing). Most of the constraints on the Hubble constant are done in the local Universe with a calibration of the distance through the use of the SNIa luminosity distance or Cepheids. The latter are stars whose luminosity varies periodically, allowing the recovery of their intrinsic luminosity and therefore, their distance. The latest results from SH0ES (Reid et al., 2019) yield a constraint of $H_0 = 73.5 \pm 1.4\mathrm{km.s}^{-1}$. The use of time-delays from strong quasar lenses can also be used to estimate the present time Hubble parameter, the most recent analysis from the H0liC0W collaboration (Wong et al., 2020) gives a value of $H_0 = 73.3^{+1.7}_{-1.8}$.

The estimation of Hubble constant with cosmic ladders evidences a strong tension between the Early and Local Universe.

1.2.3.4 Limitations of the ΛCDM cosmological model

While cosmological constraints have reached unequalled precision in the last decade, several issues have yet to be overcome in regard to the ΛCDM scenario:

- *Flatness problem*: The idea of a flat universe, separated in Ω_m and Ω_Λ as major contributions, suggests that in its early times, both quantities had to be extremely finely tuned in order to obtain the Universe observed today. Any deviation would have caused a different scenario to occur. The flatness problem emphasises the fact that were are part of a unique non-reproducible experiment which cannot be thoroughly tested.

- *The nature of the late time cosmic acceleration*: While still embodied by the existence of a non-zero cosmological constant, no satisfying scenario has been settled on. A precise estimation of the dark energy equation of state has taken very high stakes as it could enable to disentangle between several dark energy or modified gravity models.

- *The mass of neutrinos*: Massive neutrinos can have a non-negligible impact on the clustering of galaxies, it is thus primordial to properly measure their mass to understand where is their place in the overall energetic content of the Universe. Can they be accounted for as a part of dark matter?

- *Tension between local and large-scale H_0 estimation*: A consequence of the wealth of observables is the variety of estimations of the present-day expansion rate H_0. However, there seems to be a tension between the measures performed in the local Universe on SN and Cepheids and those carried out on the distant Universe with the large-scale structure analysis and cosmic microwave background.

The ΛCDM Universe, while seemingly favoured by the various constraints, is not quite satisfactory. No definite explanation have been given on the general unknowns listed above. To properly discriminate between available scenarios, both a large statistic and a variety of cosmological probes are needed. In this context are developed the next generation of large scale surveys such as LSST, Euclid (Amendola et al., 2018; Laureijs et al., 2011), and DESI (DESI Collaboration et al., 2016a,b), destined to provide a large amount of data such as SNIa, galaxy and quasar positions both in photometry and spectroscopy.

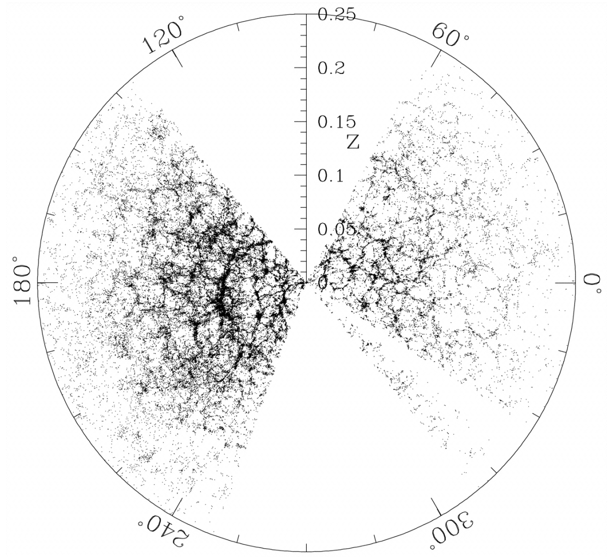


Figure 1.6: The Large-Scale structure of the Universe traced by galaxies in the early data of the Sloan Digital Sky Survey in which can be seen the different cosmic structures voids, filaments and clusters. Credits: Blanton et al. (2003)

1.3 Cosmic structures

As part of the modern cosmology development, the distribution of matter at large scale was found to form an intricate pattern of galaxies. As shown in Figure 1.6, matter on large scales clusters along and in structures delineating filaments, walls and forming clusters and superclusters. As a result, regions seemingly empty of matter can also be detected: cosmic voids.

The cosmic structures, well known as the Large Scale Structure (LSS) of the Universe, results from the concurrent effects of the expansion of the Universe and the gravitational interaction. In the context of precision cosmology, the LSS abounds of cosmological information that can be found, for example, in the Baryon Acoustic Oscillations but also in the dynamics at work in the cosmic structures. In this section, I give an overview of the linear theory of perturbations which allows understanding how these observed structures came to be and how their statistical properties are dependent on the cosmology. For references on the subject, I refer the reader to the following sources: Bernardeau (2012) and Peebles (1980), as well as Bothun (2000) that heavily influenced this chapter.

1.3.1 A model of the structure formation

The cosmic microwave background provides observational evidence of the early state of the Universe, especially in regard to the existence of initial density fluctuations which can be understood as the seed of the Large-scale distribution of matter. In this framework, two models of the structure formation were advanced relying on the same assumption that the observed large scale structure resulted from the growth of initial density perturbations:

Top - bottom formation: The top-bottom formation picture consists in the growth of over-density to ultra-dense large-scale structure, 'pancake-like', which later on fragmented to form the smaller structure such as superclusters, then clusters and finally galaxies. This formation process advocated by Zeldovich (Zel'dovich, 1970; Zeldovich, 1976) favoured the existence of a hot dark matter specie, subjected to pressure, yielding diffuse clustering.

Bottom-up formation: This second model proposed a hierarchical formation of the structure, in which the initial density fluctuation grew to form the first galaxies. This formation process relies on the existence of cold dark matter, that is a very weakly interacting matter which interacts dominantly through gravitational interacting. CDM would then start the structure formation process forming the skeleton of the LSS, forming gravitational potential wells in which ordinary matter would be attracted and then trapped, resulting in the formation of the first galaxies. This model implies a continuous formation of the LSS, where the galaxies through gravitation would assemble in clusters and superclusters.

In the current Λ-CDM paradigm, the bottom-up formation process is favoured by the amplitude of density fluctuations of the CMB.

1.3.2 Gravitational instabilities

The existence of initial density perturbation implies a variation of the mass distribution and subsequent perturbation of the gravitational potential. An over-dense region will attract matter with more intensity than an under-dense one, this uneven repartition of the gravitational potential is called *gravitational instabilities*

In this framework, the Large-Scale structure (LSS) formation is driven by the evolution of gravitational instabilities in a background expanding Universe. Considering self-gravitating non-relativistic fluid such as matter, we can describe its motions with the following set of equations:

$$\frac{\partial \rho}{dt} + \nabla \cdot (\rho v) = 0, \tag{1.47}$$

$$\frac{\partial v}{\partial t} + (v \cdot \nabla) v = -\frac{1}{\rho} \nabla P - \nabla \Phi, \tag{1.48}$$

$$\nabla^2 \Phi = 4\pi G \rho, \tag{1.49}$$

where ρ denotes the density of the fluid, v its velocity, P the pressure and Φ the gravitational potential. Eq. 1.47 is the continuity equation of a non-relativistic fluid which governs the transport of density in time, Eq. 1.48 is the Euler equation which depicts the variation of the fluid velocity through the interaction of opposite forces at play: in our case the pressure and gravity. Finally, Eq. 1.49 is the Poisson equation which relates the gravitational potential to its surrounding density field.

To understand the growth of the structure, we are interested in initial perturbations in the medium. Therefore, the quantities of interest are described as the sum of two contributions: their initial background value and an infinitesimal perturbation, assuming that the latter is adiabatic:

$$\rho = \rho_0 + \tilde{\rho} \tag{1.50}$$
$$P = P_0 + \tilde{P} \tag{1.51}$$
$$v = v_0 + \tilde{v} \tag{1.52}$$
$$\Phi = \Phi_0 + \tilde{\Phi}. \tag{1.53}$$

Considering the fluid to be homogeneous and isotropic, the equations of motion of the fluid rewrite as:

$$\frac{\partial\tilde{\rho}}{\partial t} + \nabla \cdot (\rho_0 \tilde{v}) + \tilde{\rho}(\nabla \cdot v_0) + v_0 \cdot \nabla\tilde{\rho} = 0 \tag{1.54}$$
$$\frac{\partial\tilde{v}}{\partial t} + (v_0 \cdot \nabla)\tilde{v} + (\tilde{v} \cdot \nabla)v_0 = -\frac{1}{\rho_0}\nabla\tilde{P} - \nabla\tilde{\Phi} \tag{1.55}$$
$$\nabla^2\tilde{\Phi} = 4\pi G\tilde{\rho} \tag{1.56}$$

where the product of the perturbed components are considered to be negligible.

The set of equations represents a Eulerian description of the fluid, that is a description of the evolution at a specific coordinate. As we want to express the displacement of the perturbation in regard to both gravitational interaction and the expansion of the background, we use the Lagrangian description:

$$\frac{\partial\tilde{\rho}}{\partial t} = \frac{d\tilde{\rho}}{dt} - (v_0 \cdot \nabla)\tilde{\rho} \tag{1.57}$$
$$\frac{\partial\tilde{v}}{\partial t} = \frac{d\tilde{v}}{dt} - (v_0 \cdot \nabla)\tilde{v}, \tag{1.58}$$

and substitute them in the continuity equation 1.54 and Euler equation 1.55 respectively, which reduces to:

$$\frac{d\tilde{\rho}}{dt} - \rho_0(\nabla \cdot \tilde{v}) = 0 \tag{1.59}$$
$$\frac{d\tilde{v}}{dt} + (\tilde{v} \cdot \nabla)v_0 = -\frac{1}{\rho_0}\nabla\tilde{P} - \nabla\tilde{\Phi} \tag{1.60}$$

Let us consider a density fluctuation δ, also known as the density contrast. This represent a dimensionless quantity describing the density field in regard to a background ρ_0:

$$\delta(\mathbf{x}, t) = \frac{\rho(\mathbf{x}, t) - \bar{\rho}_0}{\bar{\rho}_0}. \tag{1.61}$$

The continuity equation then rewrites in term of the dimensionless quantity:

$$\frac{d\delta}{dt} + (\nabla \cdot \tilde{v}) = 0 \tag{1.62}$$

The whole system of equations can then be considered in terms of the dimensionless density δ.

Finally, to properly account for the expanding background influence on the spatial reference frame suggested by ∇, we use the comoving distance:

$$\mathbf{r} = \frac{\mathbf{x}}{a(t)}, \tag{1.63}$$

where $\mathbf{x}$ is the proper coordinate system and $a(t)$ the scale factor of the Universe. The spatial derivative thus redefines as:

$$\nabla_r = \frac{1}{a}\nabla. \tag{1.64}$$

The time derivative has already been modified to take into account the expanding background through the equations 1.57. The velocity rewrites as:

$$\mathbf{v} = \frac{d\mathbf{x}}{dt} = \mathbf{r}\frac{da}{dt} + \frac{d\mathbf{r}}{dt}a, \tag{1.65}$$

where the first term corresponds to the expansion of the universe and the second term to our perturbation. We recover here the condition that the velocities of the objects have to be negligible compared to the Hubble expansion at large scales. The perturbed velocity vector rewrites as follows:

$$\tilde{\mathbf{v}} = a\mathbf{u}. \tag{1.66}$$

Applying our new definitions, the continuity equations becomes:

$$\frac{d\delta}{dt} + (\nabla_r \cdot u) = 0, \tag{1.67}$$

and the Euler equation reduces to:

$$\frac{du}{dt} + 2Hu = -\frac{\nabla\tilde{P}}{a^2\rho_0} - \frac{\nabla\tilde{\Phi}}{a^2}, \tag{1.68}$$

Taking the divergence of the Euler equation, we can relate the evolution of the perturbation δ with the pressure and the gravitational potential thanks to the relation in eq. 1.66 :

$$\frac{d^2\delta}{dt^2} + 2H\frac{d\delta}{dt} = \frac{\nabla^2 \tilde{P}}{a^2\rho_0} = \frac{\nabla^2\tilde{\Phi}}{a^2}. \tag{1.69}$$

The gravitational potential can be expressed through the Poisson equations 1.56. Considering the density fluctuations to be adiabatic, the pressure and density perturbations relate through $c_s = \tilde{P}/\tilde{\rho}$, with c_s the velocity of sound in the medium. Eq. 1.69 rewrites:

$$\frac{d^2\delta}{dt^2} + 2H\frac{d\delta}{dt} = \frac{\nabla^2 \delta c_s}{a^2} + 4\pi G\delta\rho_0. \tag{1.70}$$

This equation describes the rate at which the gravitational instability is affected in time, as the Universe expands. Considering the perturbation δ in a linear manner, the spatial and time component can be considered independent:

$$\delta \propto \exp(ik \cdot r - \omega t)$$

, we obtain the final equation for the time-evolution of perturbations in a self-gravitating fluid embedded in an expanding background:

$$\frac{d^2\delta}{dt^2} + 2H\frac{d\delta}{dt} = \delta(4\pi G\rho_0 - \frac{k^2 c_s}{a^2}). \tag{1.71}$$

Considering δ as defined by two independent component: $\mathbf{x}$ and t, we can write:

$$\delta(\mathbf{x}, t) = \delta(\mathbf{x}) \exp(i\omega t)$$

which gives the following equation:

$$\omega^2 + 2Hi\omega = 4\pi G\rho_0 - \frac{k^2 c_s}{a^2}. \tag{1.72}$$

The above equation relates to an oscillating system. If we consider the system at equilibrium, we can define the Jeans length:

$$\lambda_j = \frac{2\pi}{k_j} = \frac{2\pi c_s}{a\sqrt{4\pi G\rho_0}} \tag{1.73}$$

The Jeans length corresponds to the equilibrium between the gravitational force and the pressure of the gravitational instability. It allows us to consider two limiting cases:

- $\lambda \ll \lambda_j$, the perturbation oscillates due to the pressure-density interaction in the medium
- $\lambda \gg \lambda_j$, the perturbation time evolution is dominated by the expansion of the medium and grows or decays exponentially.

The latter case dominates in the context of structure formation and growth.

1.3.3 Linear theory of perturbation

Considering the very large scale, the oscillating perturbation can be discarded. The growth of an over(under) density follows the following equation:

$$\frac{d^2\delta}{dt^2} + 2H\frac{d\delta}{dt} - \delta 4\pi G\rho_0 = 0, \tag{1.74}$$

As the LSS formation process is driven by the gravitational interaction, it is considered that the matter content contributes dominantly. In this scenario, the background density $\rho_0(t)$ thus relates to that of the matter $\rho_m(t)$ which we express in terms of the critical density:

$$\rho_{dm}(t) = \Omega_m(t)\rho_c(t) \tag{1.75}$$

where $\rho_c(t) = \frac{3H(t)^2}{8\pi G}$ is the time evolutive background density. The equation of describing the time evolution of the perturbation thus becomes explicitly only on the expansion rate of the Universe $H(t)$.

$$\frac{d^2\delta}{dt^2} + 2H\frac{d\delta}{dt} - \frac{3}{2}H(t)^2\Omega_m(t)\delta = 0, \tag{1.76}$$

The evolution of the perturbation $\delta(\mathbf{x}, t)$ presented above is thus only dependent in time while the spatial evolution remains unchanged. It is thus possible to look for a solution to the equation 1.76 of a general form:

$$\delta(\mathbf{x}, t) = D_+(t)\delta(\mathbf{x}) + D_-(t)\delta(\mathbf{x}) \tag{1.77}$$

which separates spatial component and time component. $D_+(t)$ corresponding to the growing mode overtime and $D_-(t)$ the decaying mode.

In an Eistein-de Sitter $\Omega_m = 1$ universe, the solutions to this equation yield:

$$D_+(t) = t^{2/3} \tag{1.78}$$
$$D_-(t) = t^{-1}, \tag{1.79}$$

where D_+ is responsible for the growth of perturbations and subsequent formation of the LSS while D_- decreases and becomes negligible in time.

In ΛCDM Universe, however, the solutions yields:

$$D_+(t) = H(t)\int \frac{dt}{aH}, \tag{1.80}$$
$$D_-(t) = H(t). \tag{1.81}$$

Again, the growing mode is dominant before the decaying mode. For a time, the growth will be near that of the Einstein-de Sitter solutions up-until dark energy starts to dominate before the matter. From that point, the growth of the LSS will slow in time. D_+, the growing mode is known as the growth factor of structure and gives us the time evolution of the δ field.

1.3.3.1 Peculiar velocities and growth rate of structure

The derived equations above evidence that the density perturbation will be subjected to two types of velocities, one is that falls under the expansion of the Universe (Hubble flow) and one, that we call peculiar velocity.

Considering the Euler's equation at large scale, meaning we neglect the pressure:

$$\frac{du}{dt} + 2Hu = -\frac{\nabla\tilde{\Phi}}{a^2}, \tag{1.82}$$

Once again, the equation are considered in a linear manner. The velocity can thus be expressed as the product of two separate contributions, spatial and time.

$$u(\mathbf{x}, t) = u(t)u(\mathbf{x}) \tag{1.83}$$

In this case, we can write $u(\mathbf{x}) = -\nabla\Phi$ as the $\nabla\Phi$ is the only spatial component. Taking the divergence of the velocity:

$$\nabla \cdot u = -u(t)\nabla^2\tilde{\Phi} \tag{1.84}$$

The divergence of the velocity is also defined from the continuity equations 1.67:

$$\nabla \cdot u = -\frac{d\delta}{dt} = -\dot{a}\frac{d\delta}{da} \tag{1.85}$$

Combining both equations thus allows us to express the time component of the velocity:

$$u(t) = -\frac{2f}{3Ha^2\Omega_m}. \tag{1.86}$$

were the $\delta(\mathbf{x})$ contributions have disappeared in favour of the growth function D_+ and f, the growth rate of structure is defined as:

$$f := \frac{d\ln D_+}{d\ln a} \tag{1.87}$$

Considering a matter-dominated Universe, the growth rate of structure describes the rate at which structure forms and can be parametrized as follows (Peebles, 1980):

$$f \approx \Omega_m(a)^\gamma, \tag{1.88}$$

where γ is of premier importance as it explicitly depends on the dark energy equation of state (Linder, 2003):

$$\gamma = \frac{3(w_{\mathsf{DE}} - 1)}{6w_{\mathsf{DE}} - 5} \tag{1.89}$$

In the framework of GR, the prediction gives $\gamma \approx 0.55$. As a consequence, any deviation from this expected value would imply a non-standard dark energy such as modified gravity (Linder & Cahn, 2007). The dynamics involved in the creation of the LSS represent a powerful probe of the nature of the late-time cosmic acceleration.

1.3.4 The density field

The growth of the initial density fluctuations is considered to be the mechanism at cause for the observed large scale structure of matter. As the initial density field is assumed to be a random variables $\delta(\mathbf{x})$, a favoured paradigm to describe its distribution is the multivariate gaussian probability density function:

$$\mathcal{P}(\delta_1\delta_2...\delta_N) = \frac{1}{(2\pi)^{N/2}|C|^{1/2}} \exp\left[\frac{1}{2}\sum_N \delta_i C_{ij}^{-1}\delta_j\right] \tag{1.90}$$

This description has been corroborated by the results of the Planck collaboration (Aghanim et al., 2020a). A benefit from this description is given by the Wick theorem, which states that a multi-variate gaussian can simply be described by the knowledge of its covariance, here C and its average value. The latter is considered to be $\langle\delta\rangle = 0$. The former, the covariance of the density field is also known as the two-point correlation function or two-point covariance function:

$$\xi(\mathbf{x_1}, \mathbf{x_2}) = \langle\delta(\mathbf{x_1})(\delta\mathbf{x_2})\rangle. \tag{1.91}$$

By virtue of the cosmological principle, the correlation function is actually dependent on the separation between the objects instead of their specific location, as well as independence on the direction of the pair separation $\xi(\mathbf{x_1} - \mathbf{x_2}) = \xi(\mathbf{x_2} - \mathbf{x_1})$.

The density field can also be describe in terms of the power spectrum, corresponding to the Fourier transform of the density field in configuration space:

$$P(\mathbf{k}) = \int \delta(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} d^3\mathbf{r} \tag{1.92}$$

The power spectrum between two modes is formally described as:

$$P(\mathbf{k_1}, \mathbf{k_2}) = \frac{1}{(2\pi)^3}\langle\delta(\mathbf{k_1})\delta(\mathbf{k_2})\rangle, \tag{1.93}$$

which, considering statistical homogeneity and isotropy reduces to:

$$P(\mathbf{k_1}, \mathbf{k_2}) = \frac{1}{(2\pi)^3}\delta^D(\mathbf{k_1} - \mathbf{k_2})P(\mathbf{k_1}). \tag{1.94}$$

The δ^D denotes the Dirac function. Inflation, a model describing the expansion of the early universe predicts a power law power spectrum of the initial density fluctuation:

$$P(k) = A_s k^{n_s} \tag{1.95}$$

where A_s is the amplitude of the initial density perturbation and n_s is the spectral index. As of today, both quantities have been estimated in the analysis of the CMB anisotropies (Aghanim et al., 2020c), yielding $A_s = (2.105 \pm 0.030)10^{-9}$ and $n_s = 0.9656 \pm 0.0042$.

In linear theory, the power spectrum of the density field can thus be fully described from initial perturbations with the following form:

$$P(k, z) = A_s k^{n_s} D_+(z)^2 T(k)^2 \tag{1.96}$$

where $D_+(z)$ is the growth function defined in the previous section that accounts for the time evolution of the density field. $T(k)$ is the transfer function, it encodes the scale-dependent evolution of the density field, such as the impact of pressure that suppresses the growth of perturbations or the contribution of each specie and their interactions.

The power spectrum and subsequent correlation function that describe the underlying density field can be predicted in the context of linear theory. In practice, the power spectrum of the density field is estimated with tracers of the under-lying matter field in the form of galaxies or clusters.

However, the density field estimated from the observed objects δ_g does not exactly render the underlying matter field δ_m. Due to their different properties such as their luminosity, mass or simply environment, the sampling of the density field with discrete astrophysical objects will yield a biased estimate of the density field.

To take that into account, a simple assumption has been applied in the relation between the measured density field and the real density field, that is that both quantities are related in a linear manner:

$$\delta_g = b\delta_m \tag{1.97}$$

where b is the linear bias. At large scale, this bias is assumed to be scale-independent, however, it is expected to evolve with the redshift. This linear bias encodes the non-linear information of the environment, such as the galaxy formation.

When considering the power spectrum or the two-point correlation function estimated from biased tracer of matter we thus have:

$$\begin{aligned} P_{gg}(k,z) &= b(z)^2 P(k)_{\delta\delta} \\ \xi_{gg}(r,z) &= b(z)^2 \xi(r)_{\delta\delta} \end{aligned} \tag{1.98}$$

where the subscript gg denotes the galaxy auto-power spectrum or correlation function and $\delta\delta$ denotes the matter auto-power spectrum.

The Large-Scale structure of the Universe can therefore be probed and studied in a statistical manner in order to constrain cosmology.

2 Voids in the Large-Scale Structure of the Universe

2.1 Exploration of the large scale structure

The understanding of the current picture of the Universe has been lead through the discovery and exploration of the large-scale structure. The Large Scale Structure (LSS) of the Universe refers to the distribution of matter on scales larger than that of galaxies. Previous framework (Rood, 1988) relied on the thought that the matter was uniformly distributed in the distant sky, in which clumps of matter appeared as clusters or superclusters of galaxies. The advent of redshift surveys has allowed astrophysicists and cosmologists to probe the LSS and to uncover some key aspects of the large-scale distribution of matter. The increasing number of objects probed as well as the technical advances have made LSS studies the focus of most cosmological constraints in recent years.

2.1.1 Early redshift surveys : Discovery of voids

With the confirmation that nebulae in the night sky are extra-galactic objects (Slipher, 1913, 1917), and as such are galaxies in their own right, along with the steady rise in the number of observed galaxies, the distribution of galaxies on a large scale became a key question for the cosmologist. While some authors, including Zwicky (Bahcall & Joss, 1976; Zwicky, 1967), argued that the distribution of galaxies was entirely random in order to respect the cosmological principle of homogeneity and isotropy, others were convinced of the non-random distribution of galaxies on length scales from 0.1 Mpc to at least 100 Mpc (de Vaucouleurs, 1975, 1976; Vaucouleurs, 1971). Indeed, evidence for a specific clustering of galaxies has been shown by the observation of high-density regions of galaxies suggesting the existence of clusters. A very general picture of the distribution of matter was given in Abell (1958) as follows: "*There is a general field of galaxies, the surface numerical density of which varies from point to point in the sky. Whether this field is composed of isolated individual galaxies, of clusters of galaxies overlapping in projection, or both, is considered immaterial. In any case, superposed upon the general field there are occasional very rich clusters of galaxies which stand out conspicuously and which we shall assume to be physical associations.*" In a second paper, Abell argued that in addition to the existence of clusters, there would be traces of second-order clustering, meaning clusters of clusters, also called superclusters, found in the large-scale structure of the Universe (Abell, 1961). With the slow accumulation of galaxy positions and velocity maps, a statistical work applied on catalogues by Peebles (1974), which consisted in evaluating the angular correlation function of two different cluster/galaxy catalogues, disfavored the picture of a general field from which galaxies were drawn to form clusters as there were no anti-correlations, or ”no holes” as he says. In conclusion, from the discovery of galaxies to the end of the 1970s, no real scientific consensus had emerged on the distribution of galaxies. In 1976, Tifft & Gregory endeavoured to observe galaxies in a specific portion of the sky : the Coma/A1367 supercluster and found evidence of structuring of matter in what they called clusters, groups (smaller clusters) and isolated galaxies, although the latter were less numerous (Tifft & Gregory, 1976). Thompson and Gregory continued this work by accumulating more data in a magnitude-limited sample (a catalogue of all observable galaxies in a specific magnitude range) in order to obtain a more complete picture of the Coma cluster and its surroundings (Gregory & Thompson, 1978). At the same time, another group compiled the available data from different galaxies and cluster catalogues to obtain a clearer picture of the distribution of the galaxies (Jôeveer et al., 1978).

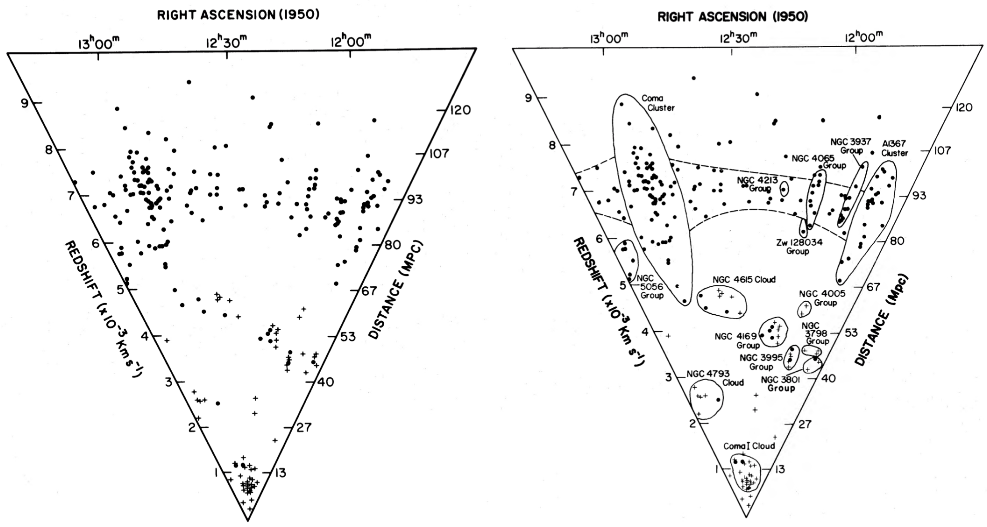


Figure 2.1: Redshift map of the magnitude-limited survey of Gregory & Thomson (Gregory & Thompson, 1978) showing the Coma/A1367 supercluster region. On the right panel, the redshift map and on the left panel, an interpretative depiction of the different clusters and groups. The first voids can be recognized in the foreground between the bridges formed by the galaxies.

As it is, Gregory & Thompson (1978) presented a first result for which there is evidence of structuring of matter, as shown in Fig. 2.1. Even more so, it seemed that the surroundings of the cluster contained a region *devoid* of galaxies. The same year, Jôeveer et al. (1978) presented similar findings from their compiled catalogues which highlighted the presence of chains of clusters and galaxies linking different structures. They also noted the presence of large empty regions in these distributions. The Large scale structure and the voids therein were found.

Due to the lack of consensus on the distribution of galaxies at large-scale, various criticisms shed a shadow on the discovery regarding the methodology used to create these galaxy maps. One of them, concerning Gregory and Thompson's work, was that their selection of galaxies was magnitude-limited, so that there could be fainter galaxies, unseen by the instrument, that could populate those supposedly empty regions (Tifft & Gregory, 1978). Another criticism, concerning the work of the Estonian team and acknowledged by the authors, was that their data were a compilation of very different catalogues and could be inhomogeneous, thus not depicting a real vision of the large-scale structure (Jôeveer et al., 1978; Thompson & Gregory, 2011).

Despite these misgivings of the scientific community, the large-scale structures of the Universe were the subject of an in-depth study in the following years. This '*lace curtain*' (Jôeveer et al., 1978) was the first inkling of a cell-like (or foam-like) pattern of the Large-Scale Structure of the Universe and has deeply affected our understanding of the distribution of matter in the Universe. The presence of prominent holes in this distribution is the first step towards a non-matter dominated Universe.

2.1.2 Confirmation of the existence of voids

Following the first observations, the telescopes continued to focus on the distribution of galaxies in the vicinity of clusters or super-clusters. The insights of a complex large-scale structure given by the first observation were confirmed on many occasions by larger statistics. The newfound structuration of matter found provided an impetus in the general study of cosmology.

Voids became a real feature of the large-scale structure landscape. Kirshner et al. (1981) detected a very large void of size $\sim 50\ h^{-1}$Mpc : the Boötes void, although further analysis (Kirshner et al., 1987) downgraded the size to $\sim 34h^{-1}$Mpc . In his review, Zeldovich et al. (1982) presents the true consequence of the observed structuring of matter in the sky: clusters and superclusters are not isolated objects at all but seem to form a complex and connected pattern in which voids play an important role. His first estimate was that voids represented more than 90% of the Universe.

Further work from de Lapparent et al. (1986) confirmed this structuration. They collected the redshift of galaxies in the vicinity of Coma/A1367, the same region of the sky initially covered by Gregory & Thompson (1978) wherein was revealed the existence of voids. With a much larger field, they were able to capture a more sharply defined image of the large-scale structure, as shown in Fig. 2.2. This picture perfectly illustrates the LSS: the filamentary structure develops from the Coma cluster, highlighting the idea of a connected network of galaxies, clusters and superclusters. This cone diagram cemented the existence of voids which became central in the study of LSS.

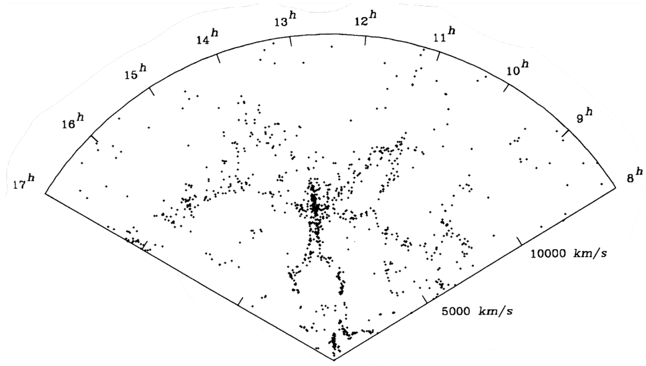


Figure 2.2: Redshift map of the CfA redshift survey in units of velocity (cz) and right ascension in h (hours) showing the distribution of galaxies around the coma cluster from de Lapparent et al. (1986).

Discovery of such connectedness of matter, however, had very large implications from a cosmological point of view. First of all, it very obviously broke the cosmological principle in that this distribution was obviously non-homogeneous! This pitfall was successfully avoided with the assumption that on a larger scale, the cosmological principle remained true.

The size of the structure also questioned the modelling of our universe, because the Universe was supposed to be dominated by matter, hardly explaining the existence of such large empty regions as voids. Several explanations were evaluated to explain the extent of these empty regions, among which the fact that these regions were formed by some kind of explosive process (de Lapparent et al., 1986; Ikeuchi et al., 1983), pushing matter out of them and thus creating these large empty regions. On the other hand, some considered, according to the observation of N-body simulations (Melott, 1983), that the possible existence of massive neutrino particles of mass in the MeV range led to void comparable in size to those observed in the sky. Nowadays, it is considered that the formation of the LSS – over and under dense – is quite well understood by a hierarchical formation process stemming from gravitational instabilities in the early Universe and in which Cold Dark Matter has a prominent role (Fry et al., 1989).

2.2 Identification of voids

The major challenge when considering voids in LSS is the ability to define them using tracers of matter and ensure that they represent the underlying true matter void regions as faithfully as possible. Indeed, voids are not objects that can be observed directly through a telescope but are defined from galaxies spread over the cosmic web. The definition of the voids will thus depend on the void extraction procedure, or more precisely on the algorithms for finding voids in tracer samples.

In this section, I will focus on void finder algorithms that allow recovering the position of under-densities and some of their properties in a tracer sample, with a special emphasis on the algorithms used during my thesis: the VIDE and Revolver algorithms, which are both wrappers of the ZOBOV algorithm.

2.2.1 Overview of void finding algorithms

The void finding step is paramount to any void studies undertook in the realm of cosmology. With the large number of galaxy positions collected over the years, visual inspection of redshift maps is no longer sufficient or even feasible to probe the under-dense regions of the Universe. The extraction and characterisation of voids in galaxy samples has to be an automated process in order to process the positions of the galaxies in an efficient and fast way. Over the course of the years, a variety of void finding algorithms have been developed in order to best extract the information from various LSS tracers, primarily galaxy or other point-like objects (QSOs).

The choice of the void finding algorithm is primordial as the void finding choices seep into the properties of the found voids. It is thus necessary to consider the different categories of void-finders in order to understand the possible bias that percolates in the void definition.

Overall, two general features of void finding can be discerned. The first is the dimensionality of the void finder, that is, whether the void finder is applied to two or three-dimensional information. Three-dimensional void finders are generally suited to find voids in tracer distributions or 3D estimation of the density field. Two-dimensional void finders can be most commonly used in photometric surveys, allowing to compensate for the error in redshift estimates by stacking angular positions and thus recovering a depth in h^{-1}Mpc (Sánchez et al., 2017). The second feature concerns the shape of the voids. One of the objectives of void finders is to locate voids in the density field and to evaluate their approximate extent. One approach used in void algorithms is to assume that voids are spherical. Although voids are not necessarily spherical, as can be seen in N-body simulations, such an assumption allows to speed up the void search process – which can take time – and still be able to locate under-densities with accuracy.

The working of a void finding algorithm is fairly universal and, in principle, follows three main steps:

1. Estimation of the density field which is methodology dependent,

2. Identification of the under-dense regions in the inferred density field,

3. Characterisation of the under-density to define voids.

The large existing variety of void finding will differ in the first aspects, that is the estimation of the density field and the definition of the voids which enters after the identification of the under-dense parts. The methodologies themselves can be classified into three general kinds (Lavaux & Wandelt, 2010), although those may not be mutually exclusive.

– **Density criterion void finders :** These are among the most developed and used algorithms. Such void finders estimate the density contrast of given data in order to identify over and under-dense regions. The subsequent regions are qualified as voids if their density is below a given density threshold. The latter is generally set as $\rho < 0.2$, in accordance with linear theory prediction (Sheth & Van De Weygaert, 2004). These voids are usually found using a spherical assumption in order to define the voids as the largest sphere containing the threshold density. Void finders defined by the density criterion have the benefit of being applicable to a wide range of data, as discrete distributions (Gottlöber et al., 2003; Hoyle & Vogeley, 2004; Kauffmann & Fairall, 1991; Micheletti et al., 2014; Padilla et al., 2005), 2D projected field (Sánchez et al., 2017), or 3D reconstructed density field such as tomographic maps, (Clampitt & Jain, 2015; Stark et al., 2015). The density criterion is usually applied as a step which corresponds to both the identification of the under-densities and the void definition. Indeed, the void size is defined as the general volume in which the density criterion is satisfied.

– **Geometrical void finders :** As the name suggests, such void finders define voids as constructs of geometrical structure, the latter being generally defined at the step of the density field estimation. As such, voids can be defined as a cluster of Voronoi polygons (Nadathur et al., 2019b; Neyrinck, 2008; Sutter et al., 2015) or their dual, Delaunay tetrahedra DIVE (Platen et al., 2011; Zhao et al., 2016), spheres or ellipsoids (Colberg et al., 2008) or grid cells (Müller et al., 2000; Shandarin et al., 2006). Such methods, except for DIVE, have the advantage of not assuming the general shape of voids, which can be quite aspherical.

– **Dynamical void finders :** In these kinds of void finders, the tracers are considered as test particles of the velocity field induced by the gravitational evolution of the LSS. They rely on the identification in Lagrangian coordinates rather than the Eulerian scheme (Elyiv et al., 2015; Hahn et al., 2007; Hoffman et al., 2012; Lavaux & Wandelt, 2010). In the local Universe, the velocities of the galaxies can also be used to map the underlying density field, thus enabling the identification of under-dense regions through velocity outflows, like the Local Void (Courtois et al., 2011; Tully et al., 2019).

Choosing a void finder from the list of possibilities is a non-trivial matter. The choice depends first of all on the choice of the definition of what a void should be: an under-dense volume or an empty volume. If the void search is based on a density criterion, then an under-density will be considered as a void only below a density threshold, but by construction, may not be totally empty, since these voids are composed of cells defined by the matter tracers themselves. This density threshold may be stringent, imposing empty voids or on the contrary allows for only under-dense voids in which some of the matter tracers can be found, thereby affecting the very definition and properties of voids. Geometrical void finders or dynamical void finders in this regard are less sensitive to such definitions, generally yielding under-dense regions with varying depth.

The variety of void finding choices corresponds to an unending search for a proper definition and parameter-free void detection. Another aspect to be taken into account in the void finding choice is the geometry of the survey. For instance, a pencil beam like survey such as VIPERS (Guzzo et al., 2014) is very deep but covers a limited area in the sky. The density reconstruction of a geometric void finder would thus not properly estimate the volume of the survey and is very dependent on the shot noise. A spherical density criterion void finder such as the one developed specifically for the survey (Micheletti et al., 2014), at the contrary, is less restrained by such specificities.

The major consequence of these choices is to impart some *a priori* of what a void should be. It follows then, that the recovered properties are affected by the choice of the definition, disabling the possibility to identify and compare the entirety of the void characteristics from one algorithm to another. A prime example of that is the shape. While finding spherical empty structures can ease the void extraction from available data, a part of the information is lost in the process. The sizes are also affected by the choice of both shape and density criterion: considering spherical voids may cause the estimated volume of the void to be smaller than in reality. This incomplete characterization of the voids can therefore be problematic when the very properties distribution can be of interest to constrain cosmology (as we will see later on), such as the shape of the voids in the case of the Alcock-Paczynski test, see Chapter 4.

Some authors have attempted to draw a comparison of various void findings. Figure 2.3 shows a selected few void finders compared against the same dark matter N-Body simulation in the Aspen-Amsterdam void finder comparison project (Colberg et al., 2008). It evidences the disparity in the void finding algorithms along with their extracted properties, such as the void centre displayed in red. Although voids are identified in the same region, they obviously differ in size, shape and void centre definition. The spherical assumption based void finders are easily noticed. The geometrical void finders, however, tend to have different behaviour, although, some provide more refined contours of voids. The latter, however, find voids in the DM particle distribution or DM density field which can explain this level of definition.

It is to be noted that the definition of the void centre still differs, even when considering the same category. A similar study, with 6 different void finders, has also been lead more recently by Cautun et al. (2018), where the different void positions and properties are compared in regard to their sensitivity to different cosmologies like GR or f(R) cosmologies. Void properties are studied in both 2D and 3D and display similar discrepancies. However, such form of comparison is primordial in void science. It is necessary to make sure that, regardless of the void finding algorithm, the properties inferred are self-consistent and display the same overall discriminating behaviour.

2.2.2 The ZOBOV algorithm

In this part, I present the algorithm which is the core of the two algorithms used in my thesis to recover voids and associated properties in samples of data, galaxies or mock galaxies. In order to have a comprehensive understanding of the void science, the algorithm must be properly studied in order to understand the objects that we recover.

ZOBOV, as in ZOnes BOrdering on Voidness (Neyrinck, 2008) is a *parameter-free* void finding algorithm which relies on the combination of the Voronoi Tesselation Field Estimation (VTFE) and a Watershed technique. It aims to find voids in a discrete sample of points as defined by galaxies or quasars. These tracers of matter will be referred to as particles hereafter. Prior to any of these methods, the sample is translated to comoving coordinates in h^{-1}Mpc through the use of flat Λ-CDM fiducial cosmology of choice. After this transformation, the particles are placed in a cubic volume, as ZOBOV was developed only run on cubic boxes.

2.2.2.1 Voronoi Tessellation

The Voronoi tessellation paves the 3D space in small volumes. It carves the whole volume in small volume entities. These small volume entities, called cells, have three properties:

- Each cell contains only one particle, and its centre is defined as the position of the particle.
- The volume of the cell is defined by the set of points in 3D space which are closer to the centre of the cell than to any other centre of another cell.
- The boundaries of the cell correspond to the mid-segment between the centre of the cell and the adjacent cell centres.

The Voronoi tessellation enables to partition the space in small volumes centred on the input particles in the form of complex polygons. The volume of each cell, by definition of the Voronoi tessellation field estimation (VTFE) (Schaap, 2007), represents a local estimate of the volume in the vicinity of the cell. The resulting local density estimation corresponds $\rho_{cell} = 1/V_{cell}$, providing a local density associated to each tracer. It is a simple matter of comparing the different

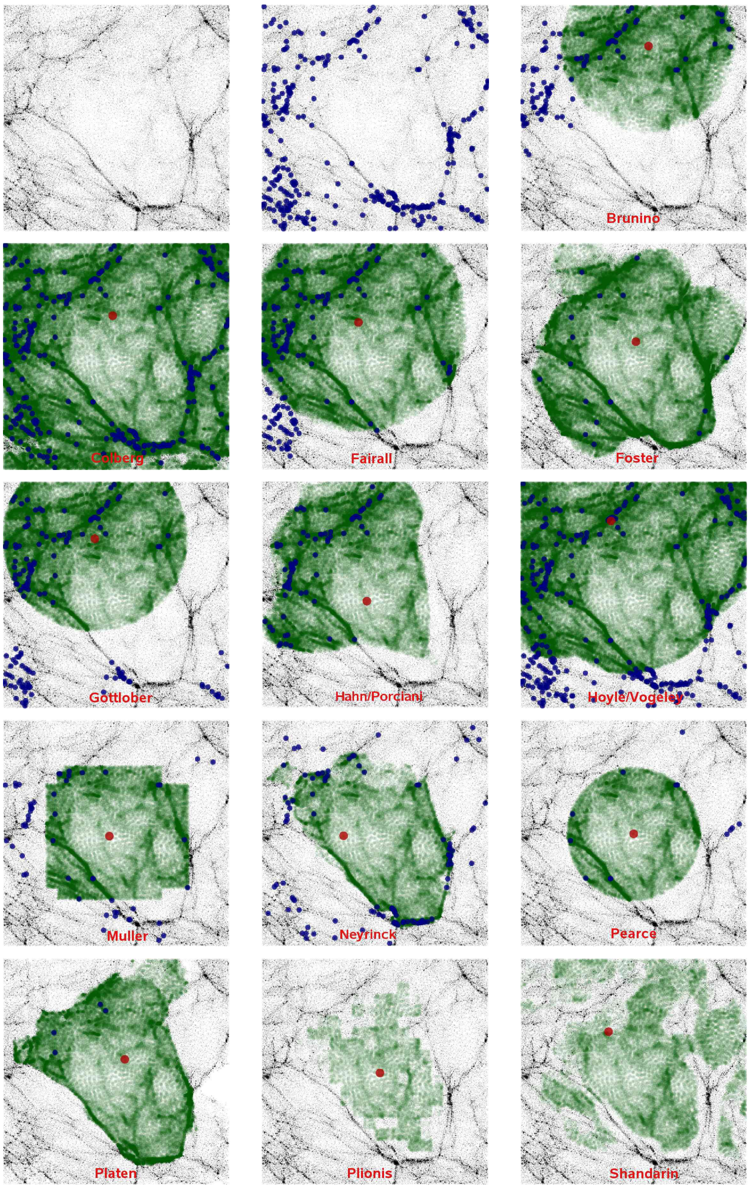


Figure 2.3: The Aspen-Amsterdam Void comparison project : A variety of void finding algorithm compared in the same slice of thickness $5h^{-1}$Mpc of the central 40 h^{-1}Mpc region of the Millenium simulation. The top-left corner displays the original simulation and top-middle shows tracer galaxies. The rest of the panels display the void identified by each algorithm in the same region. Red circles show the void centre positions position, green dots show the member dark matter particles and the blue dots display the member tracer galaxies. A review of these algorithms can be found in the source of the plot: Colberg et al. (2008).

volumes/density to find the local minima of the distribution, and thus, local under-densities in the sample.

Preliminary zones are built from the different adjacent cells up until a local maximum is reached. At which point, the merging process ceases. Fig. 2.4 illustrates the process from the point distribution (Panel a) to the resulting zones (Panel d) in a 2D plane. The tessellation allows to directly identify the most isolated galaxies by the largest cells. At the contrary, the smallest cells allow to separate and trace the high-density regions.

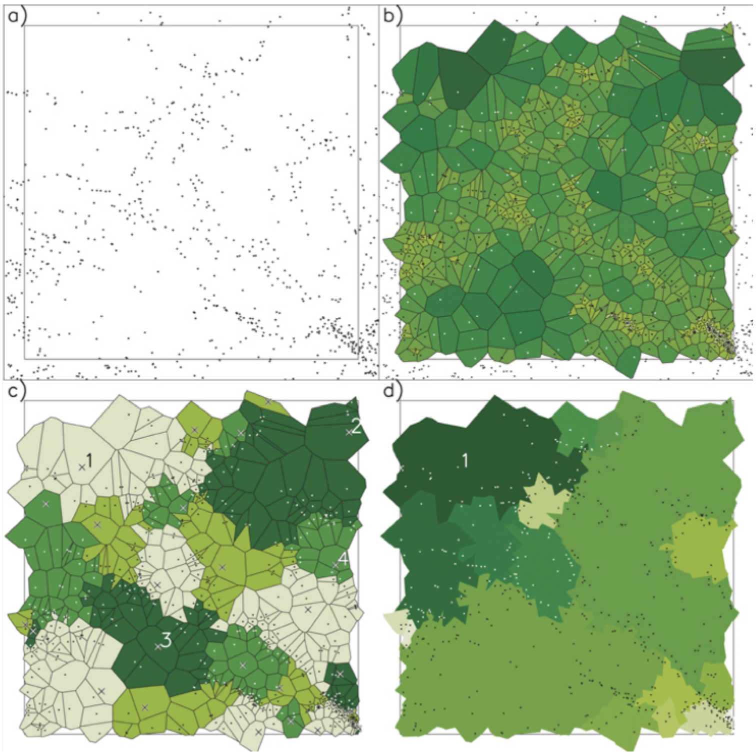


Figure 2.4: 2D illustration of the Voronoi tessellation procedure from Neyrinck (2008), (a) shows the initial disposition of the galaxies, (b) the issueing Voronoi tessellation, (c) Preliminary zone building and (d) The resulting void candidates.

2.2.2.2 Watershed Transform

In order to properly explore the density field, and to, eventually, identify a possible structure hierarchy, the resulting preliminary zones are merged together according to the Watershed Transform. This method consists in merging adjacent zones so that the dense wall separating both doesn't cross a given density threshold. It takes its name from the fact that, if one were to have adjacent basins and were to pour water in those, then, the water may connect if the separation between the two basins is low, otherwise, the basins remain distinct. It is the same for the voids, if the density threshold is obeyed to, then the distinct proto-voids are merged in a larger void. At the opposite, if the wall is too dense, the voids remain unconnected. Such a threshold is generally taken as $0.2\bar{\rho}$, where $\bar{\rho}$ is the mean density of the sample, as mentioned earlier. Others have found this threshold to be too arbitrary to really provide information on the void hierarchy in a survey (Nadathur & Hotchkiss, 2015a).

The process is illustrated in three steps in Fig. 2.5, a slice of a 2D density field is shown on the left panel for which under-dense regions can clearly be identified. The Watershed process then 'floods' the density field linking and identifying the ridges separation the basins (middle). Finally, once the water level has risen, one can identify the different separation between the catchment basins. The higher the ridge, the less likely the adjacent zones are to be merged.

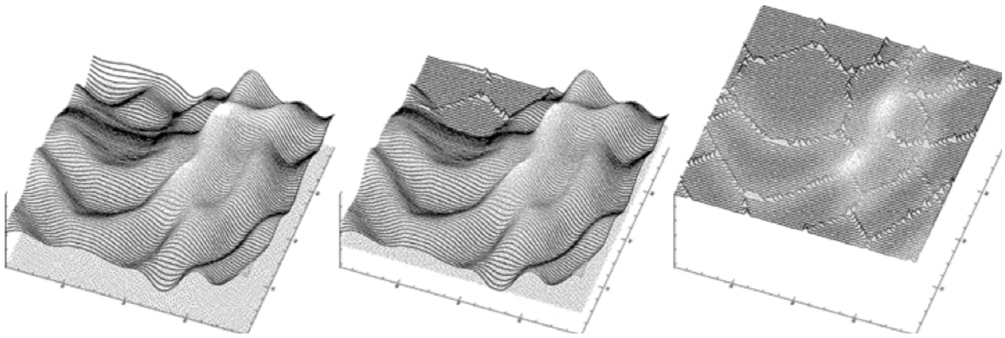


Figure 2.5: Illustration of the watershed transform from Platen et al. (2007), from left to right, the flooding level increase up until only the ridges separating the depressions are seen.

The resulting zones obtained through the combination of the VTFE and the Watershed transform are qualified as under-dense zones in a given sample of points. They correspond to an aggregate of complex polygons defining the entirety of the under-dense volume as the sum over all individual cell volume. The zones are further characterized by the particles at the core of the Voronoi cells. Some use them as *full* voids, in the sense that, they do not apply further post-processing.

From this, in order to obtain the position of the void, some post-processing is needed and as such, a final definition on the void centre is needed as well. The possibilities are numerous: emptiest centre of the cell, geometrical average of the cell contributions etc...

2.2.3 VIDE and Revolver algorithms

VIDE (Void Identification anD Examination toolkit) (Sutter et al., 2015) and REVOLVER (Nadathur et al., 2019b) are ZOBOV-based algorithms. As mentionned above ZOBOV was initially created to identify voids in cubic dark matter simulations boxes, while both VIDE and REVOLVER are wrapper of this algorithm developed to be able to be applied on survey data. Survey data are generally characterised by their geometry (or footprint) which depends on the targeted area and redshift depth. These are observational data, thus making their footprint sometimes discontinuous due to the presence of bright objects or other observational systematics. This results in vetoed/masked regions within the footprint. The existence of such boundaries has to be taken into account in order to properly extract voids, without them leaking out of the footprint.

2.2.3.1 Footprint boundaries

Voids and associated properties are extracted from galaxy samples or their likes (mocks or light-cone N-Body simulations). This process is not trivial when it comes to observational data or mocks. The existence of bright stars or other observational systematics in the measurement of the galaxy spectra results in an uneven footprint of the galaxy distribution, necessitating stringent cuts in the footprint in areas too deeply affected by those. Another aspect of these observational data is that the footprint is restricted to the specific observed area of the sky where the survey probed the positions of the galaxies.

The final galaxy catalogue is necessarily affected by these systematics and observational conditions in the form of masked out areas in the galaxy distribution. The consequence in terms of void finding is highly important because these parts of the sky must not be identified as voids in the void search process.

Both VIDE and REVOLVER tackle this issue through the use of a binary pixelized angular mask in (RA, Dec) coordinates, in which each pixel indicates an angular region in the sky. The presence or absence of galaxies can then be mapped, tracing the footprint of the survey. As a result, the empty pixels at the boundary of the survey are identified and used to generate buffer particles in the vicinity of the boundaries, the redshift boundaries of the surveys are carefully controlled also. The buffer particles, or mock particles are random positions generated within the footprint boundaries. These are then tracked in the tessellation process in order to prevent the identification of the masked regions. Fig. 2.6 illustrates the pixelized angular map corresponding to the footprint of the sample on the upper panel. The lower panel displays the identified boundary pixels within which will be generated the buffer particles.

This step is capital in the void finding process, the VTFE on which both algorithms rely, needs a properly defined volume in order to perform consistent paving of the density field. The consequence of an ill-defined survey boundary, other than finding voids in 'forbidden' regions can lead to a general misidentification or truncation of local density minima at the root of the void finding process. As such, a large number of mock particles is generated in the empty pixels corresponding to the survey angular boundary to avoid them being taken into consideration as under-dense zones. The empty pixel positions providing the angular (RA, DEC) positions. Although this step is common to both algorithms, they differ in the

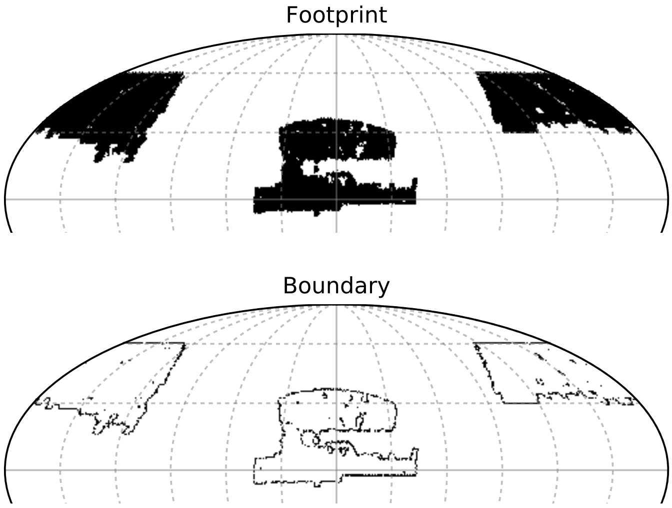


Figure 2.6: Identification of the boundary angular pixels of an Healpix map with NSIDE= 128. The upper panel displays the full pixel given by the particles' angular positions and the lower panel displays the boundary pixels

generation of such boundary particles.

Assuming the survey volume to be embedded in a spherical volume, the VIDE algorithm generates the radial positions of the particles in the whole range 0 to $D_c(z_{max})$, the maximal radial distance of the tracer, where z_{max} is the higher edge of the redshift distribution. On the other hand, REVOLVER draws the radial positions between $D_c(z_{min})$ and $D_c(z_{max})$, the radial boundaries of the survey.

A second step in the generation of the buffer mock concerns the redshift boundaries, VIDE only places a thin layer of buffer particles on the complete angular area of the survey at positions $D_c(z_{min})$ and $D_c(z_{max})$, while REVOLVER generates a wider layer which corresponds to 1.5 times the mean tracer separation of the galaxy ($\sim 1.5(3V/4\pi\bar{n})^{1/3}$). Additionally, the density of the buffer particles can be controlled in REVOLVER with a multiple of the mean tracer density.

Although these two processes seem inconsequential to the void finding process, the handling of the survey boundaries is primordial and has a great impact on the recovered void positions, as we will see in 3.1.3.1.

After tessellation, VIDE rejects any galaxy and information that is adjacent to the generated buffer particles. On the contrary,REVOLVER retains the volume of this galaxy to perform the identification of local density minima.

2.2.3.2 Under-density definition

Post-processing of the ZOBOV output covers another aspect of the vacuum search choices. VIDE allows a watershed transformation of the tessellated field with a density threshold estimated at $0.2\bar{\rho}$, $\bar{\rho}$ being the mean tracer density of the sample. This step is performed in order to extract an eventual void hierarchy which is favoured in the void excursion set theory. REVOLVER allows the merging without density threshold and defuses the resulting zones in its final post-processing steps in order to obtain the smallest entity corresponding to a under-density.

From the final information provided by the zones and the galaxies that define them, the void centre and subsequent properties are computed. The volume of the void is defined as:

$$V_v = \sum_{i}^{N_p} V_c^i, \tag{2.1}$$

where N_p is the number of galaxies which are part of the void and V_c^i is the Voronoi volume of the i^{th} member galaxy cell built in the tessellation stage. In the case of REVOLVER, the Voronoi volume pertaining to each galaxy can be corrected to take into account an eventual weighting of the considered tracers.

An effective void radius is then defined as the radius of a sphere of equal volume V_v:

$$r_v = (3/4\pi V_v)^{1/3}. \tag{2.2}$$

Finally, the centre of the void is defined as the Voronoi volume-weighted barycentre of the tracer :

$$\mathbf{x_v} = \frac{1}{V_v} \sum_{i}^{N_p} V_c^i \mathbf{x_i}, \tag{2.3}$$

where x_i is the position vector of the i^{th} member galaxy. REVOLVER also provides an alternative definition of the void centre: the circumcenter, which is computed as the centre of the circumcircle of the four largest volumes defining the void. The knowledge of the galaxy positions and input fiducial cosmology used to find the void is used to recover the voids position in observational $(\mathrm{RA}, \mathrm{Dec}, z)$ space.

2.2.4 Limitation of void search processes

Indirect void finding is the only method to pinpoint under-dense zones in an estimated density field. While it was possible to decipher these zones visually in the early void work, where the sampling statistics and subsequent number of voids were limited, it is nowadays impossible not to process the data in a big data approach. These "black box" type of methodologies are based on theoretical or geometrical assumptions and sometimes require fine-tuning in order to extract voids properly. The lack of a unified context on what to expect from the voids, their distribution and characteristics leads to several drawbacks.

2.2.4.1 Void definition and properties

Void finding techniques are numerous and can relay very different results and void properties. While the general under-density 'recognition' is independent of the void finding algorithm, extraction processes of the true underlying properties of the void are not so universal.

The consequence is that, from one algorithm to another, the behaviour of the observed properties may not be exactly the same, or even comparable. As void finder comparison projects (Cautun et al., 2018; Colberg et al., 2008) illustrate, the void centre and the sizes can be quite different depending on the type of tracer probed (Dark Matter particle, haloes or galaxies) and the adopted void definition. Theoretical expectations on the void sizes may provide a framework of comparison. Nevertheless, sizes of voids extracted with void finding algorithms have not been able to be directly compared to the theoretical distributions from the excursion set formalism of Sheth & Van De Weygaert (2004) or more recently, the 'Vdn' model of Jennings et al. (2013).

A new method has been developed in order to re-evaluate the properties of voids in regard to theoretical expectations. It enables to compare the distribution of voids to that of the expected abundance (Ronconi & Marulli, 2017), as the raw catalogue from void algorithms (even when treated through cuts) are highly different from the predicted distributions. This procedure relies on the rescaling of the under-densities in regard to theoretical under-density levels which leads to a rescaling of the void radius.

When probing for the spatial distribution of matter – clustering – around voids, this disparity in terms of void finding measurements does not seem to present a massive problem. The clustering estimation is quite consistent despite the different algorithms used, and the theoretical models explored still yield similar results, see 2.4.3 and reference therein. That being said, probing cosmology through the use of the void properties is another matter entirely, the need to modify the void size distribution, for example, raises an underlying issue of the void science: the lack of true and thorough sanity check of the void finding procedure.

2.2.4.2 The matter of significance

Most of the conclusions drawn on voids and their properties have been done through the extracted information of N-Body simulations (van de Weygaert & van Kampen, 1993) or observational data (Hoyle & Vogeley, 2004; Mao et al., 2017; Müller et al., 2000; Nadathur & Hotchkiss, 2014; Sutter et al., 2014a, 2012b). It follows that the void finding procedure has never been tested against a hypothetical simulation in which the underlying void positions are known. The fidelity of the void reconstruction, as such, as never been attested for. The lack of significance index on the validity of the under-density found thus raises some questions on the methodology itself. Do we find a majority of 'actual' voids? Or can they be found by our algorithm solely due to the fact that they probe a discontinuous field? Such questions are in part answered by the measurement of lensing signature in voids (Fang et al., 2019; Sánchez et al., 2017) which confirms that voids they found are truly under-dense statistically.

However, the question of the significance of individual voids found remains. ZOBOV provides a significance level in the form of a p-value calibrated with a Poisson sampling in a cubic box, as such, it may not be properly applied to the case of survey data. In the same vein, a multivariate analysis of the different information extracted by the void finders was developed in Cousinou et al. (2019) to compare them to under-density found in truly random distributions.

2.2.4.3 Bias in voids

A last and primordial aspect of the void finding is the role of the bias. The matter field, when probed through the prism of tracer distributions, evidences a bias in relation to the underlying true matter (DM) field. In the same way, the identification of voids rely on the estimation of such density field and should thereby be affected by the bias of the tracers. It is thus not a stretch to assume the same linear relation between true underlying dark matter voids and those inferred from luminous tracers.

$$\delta_v^g = b\delta_v^m \tag{2.4}$$

where the superscript g or m relays the density field in which voids were extracted. A primary study on the effect of the bias on voids was led by Little & Weinberg (1994) with N-Body simulations, wherein it was shown that the bias affected the void size estimation along with the surrounding density. As a result, a higher bias made the void appear emptier, seemingly consistent with the hypothesis of a linear bias affecting both over and under density. The linear bias identified in Eq 2.4 is generally considered to be the tracer linear bias b_t. This was investigated by Pollina et al. (2017) which lend credence to this assumption, at large scale. Identification of the influence of the galaxy bias in the general void properties such as their sizes and depth is made difficult by the influence of the number density of the tracer – in the event of a point-like tracer – (Nadathur & Hotchkiss, 2015b; Pollina et al., 2016; Sutter et al., 2014a). The effect of the bias furthermore affects the sensitivity to eventual deviations from the Λ-CDM model.

The linear biasing relation presented above, however, does not relate the prediction of the underlying density field δ_m and the underdense density field. As such, a second formulation of the biasing of voids can be inferred:

$$\delta_v^m = b_v\delta_m \tag{2.5}$$

where b_v is the void bias. The void bias in this form becomes a scale-dependent quantity that has yet to be quantified. The void bias has been investigated in simulations (Chan et al., 2014, 2020; Jamieson & Loverde, 2019; Schuster et al., 2019), and data samples (Clampitt et al., 2016). In the latter, a dual behaviour was noticed: large scale contributions identify a linear bias relation. However, this bias also displayed a scale-dependent behaviour, changing sign as larger voids were considered.

It follows the total void bias contribution of voids found in a galaxy sample is considered to be the product of the void bias and tracer linear bias: $b_v b_g$. However, for the moment a scale-independent bias on large scale is assumed $b_v b_g \sim b_g$. This bias was extended to an affine relation with a constant term calibrated in simulations (Contarini et al., 2019; Ronconi et al., 2019).

The relation of the observed void in regard to the true under-dense field, the void bias, is still under investigation. The void bias and its nature represent an unresolved issue in the study of voids and their properties for an application to cosmological constraints. It is an essential part of the void science to be which prevents to have a predictive power regarding the under-dense density field. In addition, while it can be considered distinct from the linear bias of the galaxies, it probably encodes an additional bias from the non-linear transformation that is the void finding process.

2.3 Cosmic voids

Linear evolution of the density field in time is well modelled in the context of the standard Λ-CDM model. It helps to understand the process at work in the formation and evolution of the Large-Scale structure and its associated content: filaments, clusters and voids.

As such, the evolution of voids embedded in the Large-Scale structure relies on the cosmological model as well. However, due to their nature as non-linear extended objects, it is not possible to predict the clustering of matter around voids and deduce the underlying properties. Voids have been studied in detail since their discovery using numerical models and N-Body simulations and observations. In this section, I give a general overview of what is known today in the field of cosmic voids.

2.3.1 Growth of voids

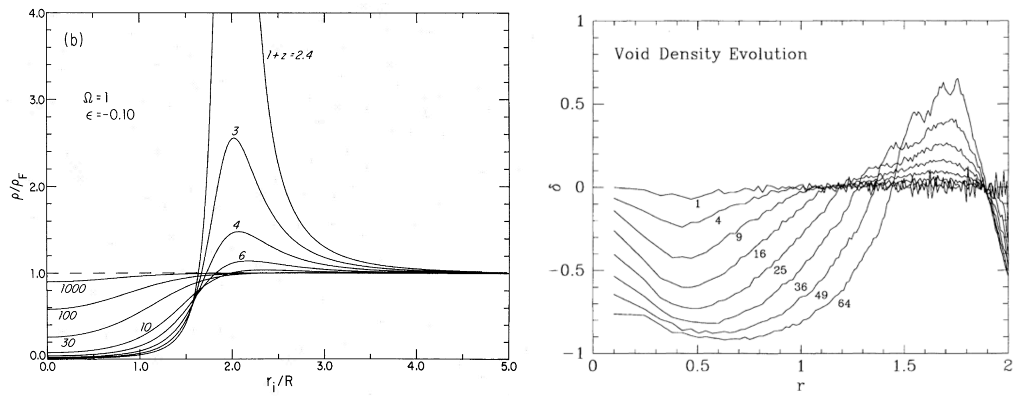
Voids are fully part of the LSS, their definition and observation are dependent on the surrounding matter field. Their evolution and growth also impact the structure formation and affect the distribution of the surrounding over-dense field.

2.3.1.1 Void Density Profile

Discovery of voids in the LSS forced cosmologists to consider how they became these salient features of the cosmic web. On the basis that the structure is a marker of past density fluctuations (Bardeen et al., 1986), the early void science was driven to explain the growth of the voids in time and space through the study of the void density profile and its evolution in an expanding Universe.

First and foremost, the void density profile traces the density field from the inside of the voids to their outskirts. It is generally evaluated by way of the void two-point correlation function (or void-galaxy cross-correlation) or modelled through a parametric function.

The first explorations concerning the growth of voids consisted in an attempt to understand their development within the Large-Scale Structure. Through the use of numerical solutions and adopted density profile, the growth of initial under-density was explored in the general growth of structure picture. The density profile was assumed to grow toward sphericity and spur from initial conditions. In time, the pressure caused by the expansion of the voids, happening more rapidly in the centre of the voids than at its outskirts should push the matter outward along the line of sight (Peebles, 1982), creating an accretion of matter around the voids that we define as its wall. This behaviour has been predicted several times over (Bertschinger, 1985; Fillmore & Goldreich, 1984; Fujimoto, 1983; Hausman et al., 1983; Hoffman et al., 1983; Icke, 1984; Lake & Pim, 1985). It is to be noted that the growth of structure is then dependent of both the growth of over-densities and-under densities. Such studies were derived using different assumptions for the void density profile and, *a fortiori*, different parametrizations.



(a) Numerical density profile **(b)** N-Body void density profile

Figure 2.7: Void density profile evolution in an Einstein de-Sitter $\Omega_m = 1$ at different epochs. (a) Parametric density profile from $1 + z = 1000$ to $1 + z = 2.4$ in comoving radius coordinate from Hausman et al. (1983) (b) N-Body simulation from $a = 1$ to $a = 64$ (past to now) Dubinski et al. (1993)

Figure 2.7 displays the density profile of voids at different epoch. Panel 2.7a displays the evolution of the density profile parametrized by an arbitrary function and evolved in time from $1 + z = 1000$ to $1 + z = 2.4$. Panel 2.7b shows the estimated density profile from N-Body simulations at different epoch. Both evidence a similar behaviour : the voids grow less and less dense while at their limits forms an over dense ridge. Meanwhile, the void also expands radially, although this behaviour is not emphasized in the Panel 2.7a as the radial coordinates adopted are independent of the expansion. This spatial expansion is responsible for the growth of the surrounding over-density.

The accretion of matter surrounding the void is indicative of the surrounding dense structure of the void. Its amplitude and definition are affected throughout the evolution of the Universe. The depth of the voids represents in some way a deficit in mass ΔM in the overall density field. As such, three types of void density profile scenarios can be inferred (Bertschinger, 1985; Ryden, 1994). The existence of the void in a general matter field, that can be considered over-dense, is evidence for the existence of a mass deficit in the general perturbation. Depending on the voids, the net mass deficit can be $\Delta M \geq 0$, marked by the presence of an over-dense wall. The void can thus be considered as compensated: $\Delta M = 0$, or over-compensated: $\Delta M > 0$. If the net mass deficit is negative, $\Delta M < 0$, the hole is considered as uncompensated and should not feature an over dense shell. This type of compensation affects the rate at which the void expands.

Bertschinger (1985) showed that, in an Einstein-de Sitter (EdS) Universe, over-compensated voids expand at a lower rate than the under-compensated voids. This can be understood in the fact that expanding over-compensated voids will end up 'competing' with the physics in charge of the growth of over-densities and thus will be challenged by those. However, uncompensated voids will not be affected by the presence of the over-dense surroundings and will expand at a different rate.

The impact of the growth of void along the Hubble flow is non-negligible in the formation process of the structure and subsequent pattern observed. Voids are not isolated objects, they do affect the surrounding density field and, as a consequence contribute in shaping up the LSS. The expansion of voids throughout the LSS tends to push matter outward of the under-densities thus compressing matter along the filaments, sheets and walls of the LSS (Dubinski et al., 1993; Icke, 1984; Martel & Wasserman, 1990; Regos & Geller, 1991; van de Weygaert & van Kampen, 1993).

The numerous studies in relation to the density profile evolution have thus agreed on a general behaviour of the void density profile: the void density profile is characterized by an under dense centre with rising density along the separation up until it crosses the mean density of matter. The density then features a wall, depending on its compensation level, before lowering toward the mean background density.

Voids found in both surveys and N-Body simulations have confirmed that this behaviour can be found for all type of voids, evoking a universal behaviour (Hamaus et al., 2014a), which was similarly confirmed in other studies (Ricciardelli et al., 2014). Some even go further arguing that observed voids can be considered as self-similar objects (Nadathur et al., 2015). All the work cited above put forward a fitting function that seemingly reproduces the void density profile:

$$\delta(r) = \delta_c \frac{1 - (r/r_v)^\alpha}{1 + (r/r_v)^\beta} \tag{2.6}$$

where δ_c is the density contrast, an indicator of the under-density level of the void, r_s is the scale at which the density profile crosses the average $\delta(r_s) = 0$ density threshold, and α and β governs the shape of the surrounding wall and the slope of the profile. Finally, r_v is the radius of the void.

Although there is a good agreement between the numerical explorations performed in the early void works and more recent works on observed voids (Padilla et al., 2005; Paz et al., 2013; Sutter et al., 2012a, 2014b), the true density profile of voids has yet to be predicted in terms of the initial density power spectrum and initial conditions as for the standard galaxy clustering.

2.3.1.2 Hierarchy

Along with the study of the evolution of voids in terms of their density profile, the development of numerical simulations of dark matter contributed significantly to the understanding of the behaviour of voids. Voids interact with their surrounding, be it over-dense or under-dense. Looking at the evolution in N-Body Dark matter simulation through random walk initial conditions, Sheth & Van De Weygaert (2004) were able to discriminate voids formation in three scenarios:

- *void-in-cloud* : a void is included in a high-density region and thus surrounded with clusters
- *void-in-void* : an under-denser region is encapsulated in a deeper void, creating one large void with local density minima.
- *cloud-in-void* : a void surrounds an over-density

Those three scenarios, combined with the void evolution, evidence the role of the void expansion in shaping the LSS: in the case of the cloud-in-void, the matter inside the void is pushed in the expansion leading to the formation of walls, while in the void-in-void case, the expansion of two voids leads to filamentary structure. Finally, in the case of the void-in-cloud process, voids tend to disappear as the surrounding over-densities merge. Those several cases, when considering the evolution of a void in time, evidence that voids follow, as for over-densities, a hierarchical formation process. All three scenarios are illustrated in Fig. 2.8.

2.3.2 Void properties

Through the void finding procedure or visual inspection, properties of note have been investigated in regard to voids. The evolution of under-densities along the expansion has consequences on their general properties: their size and their shape.

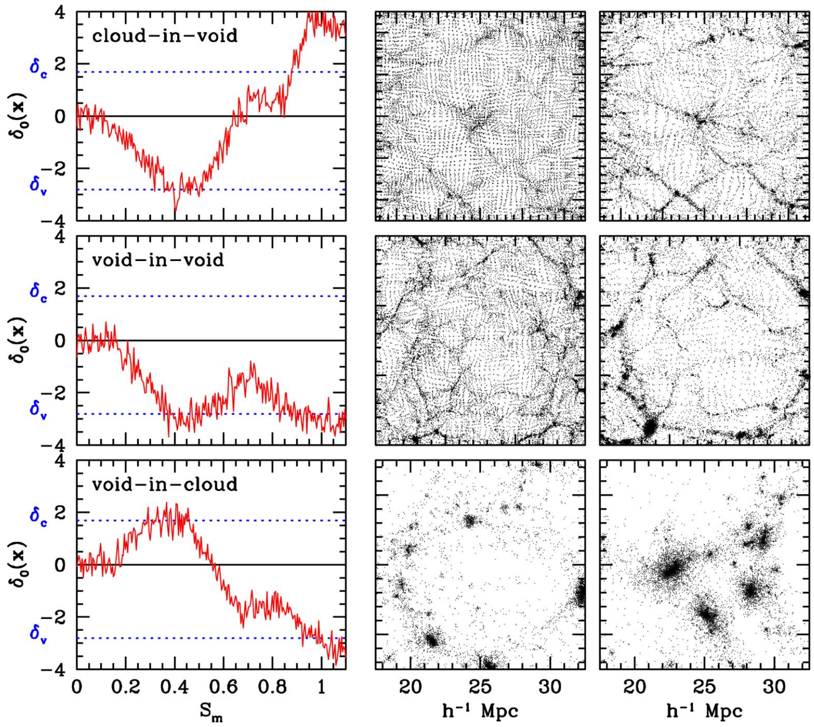


Figure 2.8: Void hierarchy clustering from Sheth & Van De Weygaert (2004), the left column shows the different initial density fluctuations at early times in the N-body simulation, while the two other columns illustrates the void scenario with a dark matter distribution. The void-in-void scenario presents substructure within its rank, while the void-in-cloud scenario presented on the right-most panel shows a void that is about to disappear through the gravitational collapse.

2.3.2.1 Sizes

Ever since their discovery in the cosmic web, the sizes of the void have been a stringent feature with cosmological implications. Indeed, the existence of such large objects, ranging in terms of tens to hundreds of h^{-1}Mpc provided an insight into the initial conditions necessary to produce such large under-dense zones. The size of voids has been a premium interest since day one of their discovery: the extent of the patches of the sky devoid of galaxies (Kirshner et al., 1981; Kirshner et al., 1987) was indubitably in tension with the cosmological mindset of the epoch of a matter-dominated universe. The size of such structure was actually part of the progressive rejection of an EdS model (Blumenthal et al., 1992).

The size of these under-densities is generally quantified in terms of length, through an estimated radius r_v or diameter, but, it is important to note that these empty regions represent an under-dense *volume* in the LSS of which the measured radius or diameter represents an effective 1-dimensional observable.

Voids being scaled objects, their volume or effective radius grow along the Hubble flow and thus, evolve as the universe expands and are affected by the cosmology and their initial conditions, like the energy density of matter or non-zero cosmological constant, (Bertschinger, 1985) or the existence of a very massive neutrino (Melott, 1987). With the numerical exploration of the void properties, as the statistic increase, follows the estimation of those experimentally. Kauffmann & Melott (1992) advanced that the size of voids was proportional to the non-linearity scale $\frac{2\pi}{k_{nl}}$, while Dubinski et al. (1993) and Goldwirth et al. (1995) estimated that at each redshift, and thus epoch, voids should have characteristic sizes.

Following the Excursion set formalism used for the clustering of haloes of dark matter (Press & Schechter, 1974), a two-barrier excursion set model was derived by Sheth & Van De Weygaert (2004) in order to take into account spherical collapse and void-in-cloud phenomenon. This model can be used to predict the void size function (or void abundance): the number of voids as a function of their size for a given cosmology and epoch. The modelling was improved by Jennings et al. (2013) and used by Pisani et al. (2015) to provide forecasts on the estimation of the dark energy equation of states, depending on which, the void abundance varies. The void abundance has since been thoroughly studied, especially in regard to its sensitivity to various aspects of the current cosmological conundrum: the acceleration of the expansion of the Universe. It was shown that the void abundance is sensitive to different dark energy scenarios as well as modified gravity (Cautun et al., 2018; Clampitt et al., 2013; Falck et al., 2018; Perico et al., 2019; Spolyar et al., 2013; Verza et al., 2019; Voivodic et al., 2017). However, the void size function estimation relies on the distribution of the tracers considered and the underlying dark matter field. Although it was shown that the abundance prediction reproduces very well voids found in underlying DM field (Ronconi & Marulli, 2017), it is not quite so in the case of voids found in galaxies or other tracers of the LSS. The modelling of the abundance has to account for this discrepancy which is a consequence of the biasing of the density field (Contarini et al., 2019). Therefore, despite the promising potential of the void sizes as a discriminating observable of dark energy, it has yet to provide any stringent constraints on the cosmology.

2.3.2.2 Shape

The last void property to be noted is their shape. The early studies of voids considered voids to be spherical objects that grew in time. Initial density fluctuations as origins for the LSS picture started to be favoured as an explanation for the peculiar distribution of matter at large scale. Works on the shape of initial density revealed those were not necessarily spherical but presented tri-axiality (Bardeen et al., 1986), implying that structures and, by extension voids, did not grow as spherical objects but rather as ellipsoidal objects. However, throughout their evolution along the Hubble flow, they would tend toward sphericity (Icke, 1984; Ryden, 1994).

Observation of the void shapes in simulations (Shandarin et al., 2006), through measurement of ellipticities, have shown that voids are not quite spherical (and sometimes far from it) despite it having been a long-standing assumption in regard to voids. This departure from sphericity has been attributed to the tidal gravitational field created by the surrounding matter. This tidal field, in turn, affects the shape of voids, by virtue of them being under-dense (Lee & Park, 2006; Park & Lee, 2007). A subsequent study of void ellipticities showed that these were indeed dependent on initial conditions (Park & Lee, 2007). A model of the shape distribution of the voids indicated a dependence in the dark energy equation of state (Lee & Park, 2009). The application of this model on several simulations, with various dark energy models, indeed showed a dependency in the equation of state of dark energy (Bos et al., 2012). It was noted, however, that even though this dependency was observed with the DM particles field, it was not so discriminating when considering biased tracers.

2.4 Cosmology with voids

The various properties presented above are tightly linked with the initial conditions of the Universe. As voids are structures which trace the under-dense part of the Universe, it is expected that their evolution and subsequent properties enable to discriminate different aspects of the cosmology. Research on and with voids has tremendously evolved in the last decades. Where voids once presented a mystery whose sole presence in the LSS led to near cosmogonic implications, they have now been widely accepted as a distinct part of the LSS. This change of paradigm of singularities in the LSS to real objects, in the same vein of galaxies or clusters, has allowed starting to shape voids as a very interesting object in order to constrain cosmology.

With the advent of N-Body simulations which enable probing the behaviour of the Large-Scale Structure of the Universe depending on the initial conditions, voids have shown considerable sensitivity to the properties of dark energy and dark matter. The impact, however, is not necessarily applicable as of yet in order to place direct constraints on cosmology, nevertheless, it highlights the potential of voids in the study of cosmology.

Indeed, voids are under-dense and so, are expected to be dominated by the only component which doesn't correspond to matter whatsoever, which means that their expansion and related physics should be governed by physical processes based on the presence of dark energy. It is thus expected that any hint on the nature of dark energy is to be found in those large under-dense regions.

In addition to the energetic contribution Ω_Λ, voids are expected to be able to place severe constraints on the dark energy equation of state $w(z)$ which could help disentangle between dark energy models. To this end, several probes of dark energy have been investigated in the context of under-densities. In the previous section were highlighted the properties of voids and their sensitivity to the current cosmological concerns: is the late-time acceleration a consequence of dark energy or of a modification of the laws of General Relativity?

The distributions of the shape and abundance have shown to be sensitive to a variation of both of them and as such, voids are expected to be able to contribute to the overall understanding of cosmology. As of today, the study of void properties has not yet provided any observational constraints on the nature of dark energy but their sensitivity to it highlight its potential. In this part, I will give a brief overview of the additional probes of cosmology that can be considered in regard to voids.

2.4.1 Voids and gravitational potential

Presence of under-density and their differing growth in regard to over-density have several effects on the overall gravitational potential. Such effects can be probed in order to detect traces of new physics in the dark energy framework.

2.4.1.1 Weak lensing imprints

The presence of very massive objects has a consequence of curving the space-time in their vicinity. A consequence is that the light rays emitted by a luminous source behind a massive object, like a cluster, will see its trajectory deviated depending on how massive the object is. This cluster will thus act like a lens on other light-emitting source situated behind it.

Voids, on the other hand, are generally surrounded by massive structures such as filaments, clusters, walls, whose presence will curve the space-time, on different levels. However, the centre of voids should not affect the fabric of space-time, leading to a similar effect as with massive objects: weak lensing. The light rays emitted from a bright source and sufficiently distance from the void centre should be repulsed from the void-centre as in the massive lensing case if the mass was negative, (Amendola et al., 1999). In the case of voids, it was considered that measuring the individual weak lensing signal in the void yielded too noisy results and were too dependent on whether the void was properly identified, as such, weak lensing studies started to consider stacks of voids to infer an average weak lensing signal, (Higuchi et al., 2013; Krause et al., 2013). A first detection and measurement of the weak lensing was performed in Melchior et al. (2014) and further refined with Clampitt & Jain (2015). The most recent results on the weak lensing imprints in voids have successfully confirmed their under-dense nature (Fang et al., 2019; Sánchez et al., 2017). The latter arguing that strong deviation from ΛCDM, provided proper theoretical modelling, could very well be traces of new physics. A primary study of the weak lensing signal in the presence of modified gravity was led by Baker et al. (2018), in which was highlighted the potential of the weak lensing in voids to investigate deviations from Λ-CDM with voids extracted from the seventh data release of the SDSS program. However, a limiting factor proved to be the statistics and subsequent errors on the weak lensing signal that prevented decisive discrimination. An interesting take on weak lensing has been investigated by Davies et al. (2018), in which is considered finding voids directly in weak lensing map, allowing to increase the significance of the lensing signal.

Following the work on the existence of potential lensing imprints on distant sources, it is possible to consider the CMB. Indeed, the primordial photons carrying the CMB signal are well enough distant in order to be affected by the existence of voids. As such, the CMB measured by Planck (Aghanim et al., 2020a) has probably been impacted by the gravitational lensing of voids. Chantavat et al. (2016) provided a forecast of the study of CMB lensing with voids in and argued that this could break degeneracies to allow for a precise measurement of cosmological parameters. Such analysis was performed with voids found in the last BOSS data release (DR12) (Cai et al., 2017; Raghunathan et al., 2020) and in the first year release of the DES survey (Vielzeuf et al., 2019), all recovering a signal.

2.4.1.2 Integrated Sachs-Wolfe effect

The effect of CMB lensing can, however, be degenerate with the integrated-Sachs Wolfe (iSW) (Sachs & Wolfe, 1967) effect inside voids. The latter is a result of the decay of gravitational potential over time due to the effect of the late-time cosmic acceleration, as it is, probing the amplitude of the iSW effect provides a direct probe of dark energy (Sołtan, 2019). This effect is expected to be stronger in the vicinity of superstructures of which voids are part. Such effect is generally probed in combination with the CMB temperature measurements, which are affected by the superstructures encountered by the CMB photons. The first detection of such a signal was performed by Granett et al. (2008). Since then, the technique has been thoroughly investigated (Kovács et al., 2017; Nadathur & Crittenden, 2016; Nadathur et al., 2017) and recent results in the DES survey have reported an excess iSW amplitude in tension with the standard Λ-CDM prediction (Kovács et al., 2019).

2.4.2 Alcock Paczynski test

The Alcock-Paczynski test consists in measuring the shape of voids in regard to a cosmological template. The shape of the void should vary if the cosmological template is wrong, incurring a distortion of the shape. The measure of such a distortion allows for recovering the true underlying cosmology. This test was first proposed by Alcock & Paczyński (1979), and by Ryden (1995) in regard to its application to voids. The work of Lavaux & Wandelt (2012) regarding the use of stacks of voids to apply the Alcock-Paczynski test provided an impetus, arguing that the Alcock-Paczynski test, when applied to voids, should compete in time with the measurement applied to the BAO. First works did not provide a strong constraint (Sutter et al., 2012a), but later works showed an improvement (Mao et al., 2017; Sutter et al., 2014c). The Alcock-Paczynski test has also shown promise in voids found in 21cm intensity maps (Endo et al., 2020). The Alcock-Paczynski test applied to void will be investigated in Chapter 4.

2.4.3 Clustering of voids

The study of the clustering can be used in order to constrain f, the growth rate of structure, which is a quantity predicted by the General Relativity in the ΛCDM framework. Deviation from the expected measurements would thus imply a deviation from Einstein's gravity, laying the ground for modification of gravity on large-scale in order to explain the accelerated expansion of the Universe.

The clustering of voids in regard to the surrounding matter field has been shown to be sensitive to the dynamics of the galaxies which carry the growth rate information (Padilla et al., 2005). As the statistics of galaxies increased with the advent of large redshift surveys, so did that of the voids, allowing to measure more and more precisely the correlation function between voids and tracers of matter (e.g galaxies). Since Pan et al. (2012) provided the first measurement of the growth rate of structure of the galaxies in the vicinity of voids in the DR7 data release of the BOSS (or SDSS) survey, the potential of voids as a probe of f, and thus gravity, has been recognized over and over with various analysis applied to data (Achitouv, 2019; Achitouv et al., 2017; Aubert et al., 2020; Hamaus et al., 2017, 2015; Hawken et al., 2017, 2016; Hawken et al., 2020; Nadathur et al., 2019a, 2020b; Paz et al., 2013), the precision of the estimation tending toward competitivity with standard galaxy clustering technique. The study of the growth rate will be addressed in details in Chapter 3.

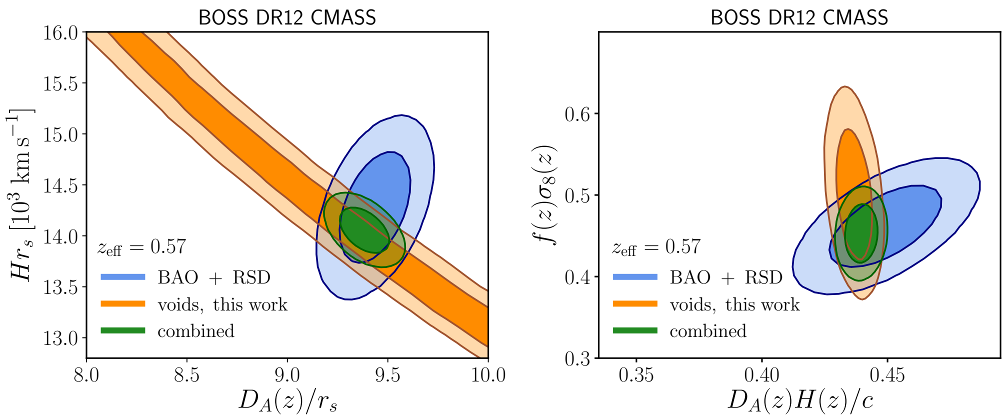
In conjunction with the growth rate can also be probed the cosmology with the two-point correlation function (Hamaus et al., 2015; Nadathur et al., 2019a) through the inclusion of a parameter taking into account the Alcock-Paczynski effect (summarily described above), rendering the void-galaxy correlation function very interesting in terms of cosmological constraints.

A final part of the interest of voids in terms of clustering is that they also include a BAO feature(Chuang et al., 2017; Kitaura et al., 2016) which can contribute to increasing the signal to noise and better the constraining power of this technique.

2.4.4 Degeneracy breaking

Voids are thus very sensitive to the cosmology and can be probe through numerous techniques and physical processes. Combined altogether or with other standard cosmological probes, they can relay significant improvement to cosmological analysis as they consider different measurements which, in turn, display different degeneracies between cosmological parameters. It is of great importance to be able to break these degeneracies in order to obtain the clearer picture possible to make significant constraints on the cosmology, and especially, dark energy models or traces of modified gravity.

Forecast focusing on the contribution of the void to the constraint of the equation of state was done on the void abundance (Pisani et al., 2015) or the shape distribution (Biswas et al., 2010). The former evidenced that the combination of the measurements of abundance with joint estimation from the CMB, Hubble Space Telescope with the void expectations from Euclid-like or WFIRST-like surveys provided a significant improvement on the constraint of the equation of state measurement. The latter presented the forecast on the combination of the shape distribution with SN forecast from LSST and DES and Euclid-like voids and showed that the voids provided independent measurements allowing to break degeneracies and thus obtain more constrained results on the estimation of the dark energy equation of state.



(a) (b)

Figure 2.9: (a) Combination of the galaxy clustering standard constraints with the two-point correlation function in the $Hr_s - D_A(z)/r_s$ plane. (b) Same but on the $f\sigma_8 - D_A(z)H(z)/c$ plane. Credits : taken from Nadathur et al. (2019a)

The combination of voids and galaxy clustering standard techniques was studied in Nadathur et al. (2019a), which evidenced very much tighter constraints as the Alcock-Paczynski parameter and growth rate that were probed in this work do not show the same degeneracies as the standard galaxy-clustering estimation, amounting to an improvement on the parameter's measurement of 30% to 60%. The improvement thanks to the combination of all the LSS information is illustrated in Fig. 2.9, showing that voids provide very complementary information, yielding much tighter constraints than standard galaxy-clustering alone, as they break the degeneracies in the parameters.

The same analysis was performed on later eBOSS data, yielding an improvement ranging from 13% to 28% on the parameter estimation (Nadathur et al., 2020b). Combination of void-galaxy clustering measurements along with these of standard galaxy clustering and their consequence emphasize the benefit of considering all of the volume traced by the redshift survey: over-dense and under-dense.

Finally, the combination of the added information provided by the voids, along with other unrelated probes such as the CMB or the SN1a allows for a very precise measurement of all the parameter space regarding the cosmology, among those: the dark energy and H_0 which remains to this day in tension with that found by the CMB (Nadathur et al., 2020a). The void potential is again very much noticeable in combination with the available probes, as can be seen on the estimation of Ω_Λ in Fig. 2.10.

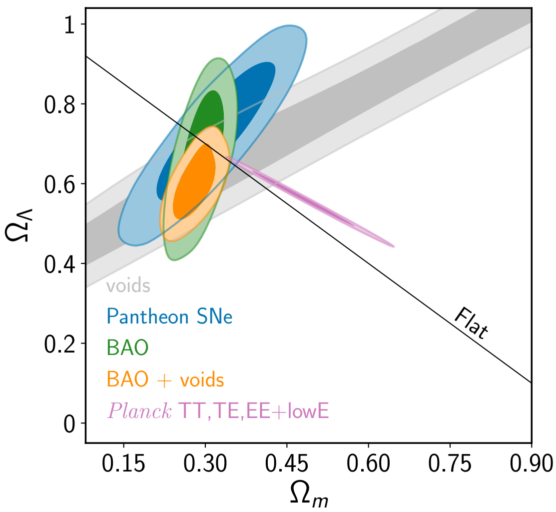


Figure 2.10: Concordance measurements with the combination of the BOSS voids + galaxy , SNIa and Planck measurements. Significant improvement on the estimation of Ω_Λ can be seen. Credit : from Nadathur et al. (2020a)

Voids are very relevant to the present cosmological constant which tries to access the nature of dark energy. They are sensitive through their properties, sizes, shapes or density profiles, to the dark energy equation of state or eventual modifications of gravity. Their evolution is driven by the current energy content of the Universe and the behaviour of dark energy, which means that their relation with the matter distribution is also dependent on the initial conditions. The abundance of voids has not yet been used to derive the initial conditions from available data, the analysis is still undergoing some fine-tuning in order to be able to correctly constrain the cosmology regardless of the void finding algorithm, while the shape distribution of voids has been put aside in favour of the attractive Alcock-Paczynski test (Shoji & Lee, 2012). Several probes have been developed in order to search for hints or clues on the late-time cosmic acceleration, voids lensing has been found to occur in voids and to affect the CMB signal, but it has not yet provided tight constraints on the initial conditions as it lacks the prediction available of the matter distribution. The latter is currently unknown for voids. On the other hand, the study of the void-galaxy clustering has been a long-standing process and shows significant improvement in signal recovery as time goes on, yielding complementary observations to the standard galaxy clustering technique, allowing for much tighter cosmology constraints.

The quest of precise measurement in cosmology, or precision cosmology, is reaching a new era in which voids may have their say, allowing to disentangle between measurements and discriminate between the favoured scenario of the cosmic evolution.

3 Probing cosmology with dynamical distortions around voids

In the search for the nature of dark energy and specifically, breaking the degeneracy between dark energy models and modified gravity, a key test is provided by the estimation of the linear growth rate of structure. This growth rate, as presented in section 1.3.3.1, encodes the rate at which cosmic structures grow with the expansion and can be predicted in the context of General Relativity. Constraints on the growth rate can be provided by galaxy redshift surveys.

In this chapter, we will show how cosmic voids can be used to constrain the growth rate of structure at three different effective redshifts. First, we will present the final Data Release 16 of the Sloan Digital Sky Survey from which will be extracted the growth rate estimates. Then we will introduce the linear redshift-space distortion model used to estimate the growth rate of structure, describe its application on mocks and evaluate systematic errors from different sources. Finally, we will present the final constraints on the growth rate of structure using voids in the final galaxy catalogues of the extended Baryon Oscillations Spectroscopic Survey.

3.1 The Sloan Digital Sky Survey: Mapping the Universe

3.1.1 Overview of The Sloan Digital Sky Survey

The Sloan Digital Sky Survey (SDSS) (York et al., 2000) has been working for more than 20 years to map the Universe and has, as of today, gone through four phases.

In the first two phases, SDSS-I and SDSS-II ranging from 2000 and 2008, the SDSS Legacy Survey was conducted by imaging the sky in five bandpasses (u, g, r, i and z) (Fukugita et al., 1996) using the SDSS imaging camera (Gunn et al., 1998). The spectra of more than 1.3 million targets were observed with the SDSS Legacy spectrograph (Smee et al., 2013) over 8000 deg^2 of the sky, covering a large contiguous region in the Northern Galactic Cap (NGC) and three thin stripes in the Southern Galactic Cap (SGC). The subsequent release of the SDSS main galaxy sample at low redshift in the seventh data release, as well as Luminous Red Galaxy sample, (Abazajian et al., 2009, DR7) provided the most precise picture of the Large-Scale structure at that time. In these first stages was made the first detection of the Baryonic Acoustic Oscillations, presented in section 1.2.3.2. In concurrence with the analysis of the growth rate of structure, redshift surveys were put forward as the best environment to probe dark energy in its numerous forms.

In Summer 2009 started the Baryon Oscillation Spectroscopic Survey(BOSS) (Dawson et al., 2013), part of SDSS-III (Eisenstein et al., 2011). Targeting principally Luminous Red Galaxies, the BOSS main objective was to extract and constrain the BAO at the percent level. Targets were observed using double-armed fibre fed optical spectrographs (Smee et al., 2013) mounted on the 2.5-meter Sloan Telescope at the Apache Point Observatory (APO) (Gunn et al., 2006).

By the end of SDSS-III, in 2014, BOSS spectroscopically surveyed 10,338 $\deg^2$, gathering 1.2 million galaxy spectra of Luminous Red Galaxy (LRG) to $z = 0.7$ and 180,000 observed Ly-α quasars to map the fluctuations in neutral hydrogen at redshifts $2.1 \leq z \leq 3.5$. The complete SDSS-III data set was released in 2015 January in DR12 (Alam et al., 2015). The last data release (DR12) of the SDSS-III provided the biggest catalogues of galaxies of the same specie to date as well as percent level estimation on the BAO distances (Alam et al., 2017).

3.1.2 The DR16 of the extended-Baryonic Oscillation Spectroscopic Survey

Part of the fourth phase of SDSS (Blanton et al., 2017, SDSS-IV), the extended Baryon Oscillation Spectroscopic surveys (Dawson et al., 2016, eBOSS) endeavoured to probe deeper in redshift, specifically targeting three types of tracers: Luminous Red Galaxies (LRGs), Emission Line Galaxies (ELGs) and Quasi-Stellar Objects (QSOs). The latter were probed both for galaxy clustering and Lyman-α forest cosmology. The eBOSS survey had a duration of five years during which were collected nearly a billion spectra thanks to the BOSS instrument described below. It ended in March 2019 recording the last spectra measured to constrain cosmology. The completed survey provided its last cosmological analysis in July 2020 (Alam et al., 2020a), as part of the sixteenth date release of SDSS (Ahumada et al., 2020). The work presented in this chapter was carried out as part of the eBOSS collaboration.

3.1.2.1 The BOSS instrument

At the heart of the eBOSS survey is the precise determination of the redshift of the targeted tracers. This necessitates an optimal measurement of the spectra of the considered targets to allow for the identification of the signature lines on which are based the galaxies classifications and their subsequent redshift. To this end, the eBOSS survey made use of the two-armed fibre fed optical spectrographs previously used in the BOSS program. This instrument is built on the legacy instrument of the two first stages of the SDSS program (I and II) which consisted of two fibre fed spectrographs of 320 fibres each. It went through a re-design in order to suit the

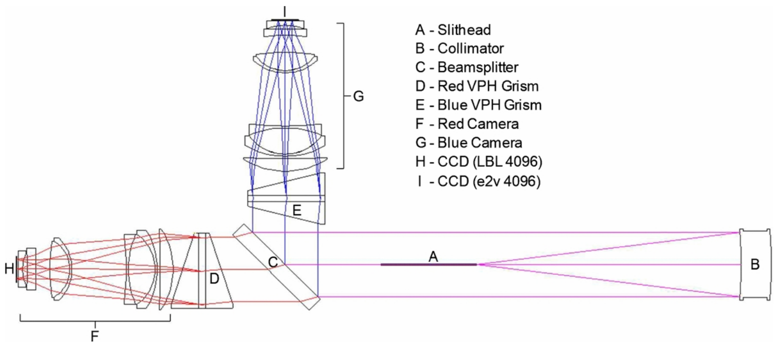


Figure 3.1: Optical scheme of one BOSS spectrograph. From:Smee et al. (2013)

requirements of the BOSS program. An in-depth description of both SDSS and BOSS instrument can be found in Smee et al. (2013).

As such, the BOSS instrument consists of twin two-arms spectrographs, each arm probing blue and red wavelength respectively. Each of the spectrographs is fed by 500 optical fibres of sizes $2''$ ($\sim 120\ \mu$m). The thousand optical fibres are hand-plugged on aluminium plates in which are drilled the positions of the targets, connecting the spectrographs to the focal plane of the telescope. The other end of the fibres connects to two slit-heads, respective to each spectrograph. The structure guaranteeing the assembly of the plates, fibres and slit-heads is called the cartridge. Several cartridges are provided for in order to load in advance the plug-plate and fibres to optimize the acquisition time.

An optical scheme of the one of the BOSS spectrograph is displayed in Figure 3.1. The light collected by the fibres then goes through the slit-heads (A) and is collimated (B) toward a dichroic plate (C). The dichroic plate splits the beam and redirects it to the two arms of the spectrographs. Each beam is diffracted through a grism (E, D), respective to the wavelength range of the cameras. The resulting spectrum is then collected by a camera sensitive to the wavelength considered (F, G).

The resulting spectra are fitted according to the noticeable emission line or absorption lines in the spectra, allowing to recover a precise estimation of the redshift of the observed object.

The fibres cannot be plugged on the plate at a distance inferior to $62''$ in order to exclude fibres colliding with each others (Reid et al., 2016).

3.1.2.2 The DR16 clustering sample

Prior to the collection of the spectra, the target selection of the QSOs and LRGs was performed with the DR13 SDSS imaging photometry (Albareti et al., 2017) with additional information from the WISE satellite (Wright et al., 2010). The ELG target selection was performed using the DECaLS part of the DESI Legacy

Imaging Surveys[1] (Dey et al., 2019).

Luminous Red Galaxies Luminous Red Galaxies sample the overdense part of the Universe, as the brightest galaxies in the vicinity of galaxy clusters. With their $\lambda4000$ Å line, their spectra can be identified and used to estimate their redshift (Eisenstein et al., 2001). In eBOSS, the LRGs are sampled at a redshift $z > 0.6$, where they appear fainter. This allows the eBOSS LRGs to be selected on a colour cut basis and to discriminate them from the high redshift tail of the BOSS LRGs (CMASS). The detailed LRG eBOSS selection is described in Prakash et al., 2016 to provide a sample of galaxies with redshift between $0.6 < z < 1.0$. The eBOSS LRG sample totals 174,816 objects over a footprint of 4,242 deg^2, the full process of the catalogue creation is described in Ross et al. (2020).

As in galaxy clustering analyses on the LRG sample in Fourier space (Gil-Marín et al., 2020) and configuration space (Bautista et al., 2020), the eBOSS LRGs are combined with the high redshift tail of the BOSS CMASS galaxies ($z \geq 0.6$). The combined LRGpCMASS catalogue contains 377,458 galaxies with $0.6 < z < 1.0$ over a total footprint of 9,493 deg^2. All eBOSS LRGs are assumed to be within the CMASS footprint. From this point forward, the LRGpCMASS sample will either be mentioned as LRGpCMASS or LRGs in an exchangeable manner, unless stated otherwise.

Emission Line Galaxies The star-formation density increases with redshift in the range $0 < z < 2$ and this process is expected to happen mostly within galaxies in the late Universe (Madau & Dickinson, 2014). This process leads to marked emission lines in the spectrum of the host galaxy, the most characteristic among them being the $[\mathrm{O_{II}}]$ doublet emitter at ($\lambda3727$, $\lambda3729$ Å) or the Hα ($\lambda6563$ Å). These wavelengths, due to the expansion, can be observed in the optical range in the case of the oxygen doublet lines and in the infrared in the case of the Hα line. The galaxies emitting these spectra are classified as Emission Line Galaxies. Their increasing density at high redshift has made them prime targets in the case of the eBOSS survey (for the $[\mathrm{O_{II}}]$ ELGs) but also in future large spectroscopic surveys such as DESI ($[\mathrm{O_{II}}]$ emitter; DESI Collaboration et al., 2016a,b) and Euclid (Hα emitter; Amendola et al., 2018). The target selection performed for eBOSS is described in Raichoor et al. (2017). The creation of the Large-Scale structure catalogue, specifically build toward cosmological analysis, is detailed in Raichoor et al. (2020). The resulting catalogue counts 173,736 objects in the redshift range $0.6 < z < 1.1$ covering a footprint of 1170 deg^2.

Quasars Quasars, as in Quasi-Stellar Objects (QSOs), are highly luminous objects powered by active galaxy nuclei (AGN) which are presumed to be supermassive black holes. Their high luminosity makes them very interesting targets to probe

[1] http://legacysurvey.org/

the Universe at very high redshift. Within the eBOSS survey, two populations of quasars were targeted: high redshift QSOs ($z > 2.1$) for which the distribution is not quite homogeneous and low redshift QSOs ($0.8 < z < 2.2$). High redshift QSOs emits Lyman-α (Ly-α) line (λ1215.67 Å). As highly luminous objects, the photon flux of the Ly-α is absorbed when neutral hydrogen gas is encountered in the intergalactic medium. This absorption results in a series of absorption lines in the spectrum: the Lyman-α forest. The latter gives an indication of the structures encountered throughout the photon's journey, allowing to map the large-scale structure.

The low redshift QSOs, the CORE sample, represents a homogeneous sample in a uniform volume in order to obtain point-like tracers of the LSS in the same capacity as the LRGs and ELGs. The CORE sample allows bridging the gap between the low $z < 1$ and high $z > 2.2$ redshift probes of cosmology (Dawson et al., 2013). The total CORE and Ly-α QSO target selection is described in Myers et al. (2015) for eBOSS. The resulting DR16 QSO catalogue compiles the entirety of the quasars observed spectroscopically within all the SDSS stages and is presented in Lyke et al. (2020).

Thereafter, we will only consider the homogeneous sample QSOs used for the clustering. The creation of the clustering catalogue is fully described in Ross et al. (2020). The number of eBOSS QSOs is 343,708 covering a sky area of 4,808 deg^2, and spanning the redshift range $0.8 < z < 2.2$.

For each of the samples presented above, random catalogues are generated with at least 40 times the number density of the original tracer sample, in accordance with both their angular and radial distribution. They are used to measure the correlation function respective to each tracer. The description of the random catalogues generation is given in Ross et al. (2020) for the LRGs and QSOs, and Raichoor et al. (2020) for the ELGs.

3.1.2.3 Correction of systematics

The catalogues provided for the analysis of the clustering of those objects report the right ascension (RA), declination (Dec) and measured redshift z of each object, the latter being estimated by fitting the galaxy spectra.

As the galaxy positions depend on both technical and observational conditions, weights are computed to correct for eventual systematic effects and are defined as follows:

- **Fibre collision weights** $w_{\rm cp}$: The minimal separation between two fibres is of 62″ in order to avoid any collisions between fibres. In the event where several galaxies are situated within this range, only one galaxy will have an assigned fibre. As a result, some initially targeted objects were not observed spectroscopically. The likelihood for this case to present itself is higher in over-dense regions. As such, the close-pair weight $w_{\rm cp}$ defined as $N_{\rm targ}/N_{\rm spec}$, the ratio of targeted objects in a given group over the number these objects observed spectroscopically. It is used to up-weight the neighbouring galaxies of the missing targets.

- **Catastrophic redshift weights** w_{noz}: Some observed spectra yield unreliable redshift estimates which have been attributed to the fibre used to collect the signal. As for the close-pair, the resulting object is removed from the catalogue and its presence is accounted for by up-weighting the nearest object of the same type (QSO, ELG, LRG) with a weight w_{noz}.

- **Photometric systematic weights** w_{sys}: Observations are based on photometric data which themselves are subjected to systematics. The density of stars, seeing, airmass or overall sky background can affect both target selection and resulting data acquisition. These effects are corrected for in a single w_{sys} weight.

- **FKP weights** w_{FKP}: At the contrary of the three weighting schemes presented above, the FKP weights represent an optimization of the galaxy clustering estimation. They are set to minimize the variance in the clustering measurement that arise due to a non-uniform redshift distribution (Feldman et al., 1994). In eBOSS, they are parametrized to optimize the estimation of the BAO signal. The FKP weights are defined as follows:

$$w_{\text{FKP}} = \frac{1}{[1 + \bar{n}(z) P_0]} \tag{3.1}$$

where $\bar{n}(z)$ is the mean number density of the sample at redshift z and P_0 is the amplitude of the power spectrum at the scale where the BAO signal is highest. For the different eBOSS tracers:

$$P_{0,\text{LRG}} = 10000\, h^{-3}\,\text{Mpc}^3, \tag{3.2}$$
$$P_{0,\text{ELG}} = 4000\, h^{-3}\,\text{Mpc}^3, \tag{3.3}$$
$$P_{0,\text{QSO}} = 6000\, h^{-3}\,\text{Mpc}^3. \tag{3.4}$$

These resulting weights balance the contribution of every redshift bin considered in the sample.

The final weight for each object can then be written as:

$$w = w_{\text{noz}} \times w_{\text{cp}} \times w_{\text{syst}} \times w_{\text{FKP}} \tag{3.5}$$

Both randoms and galaxies alike are provided weights.

3.1.2.4 Mock catalogues

Mock data are used to compute the covariance matrix of the two-point correlation function as well as test eventual systematical effects. Two types of mocks are produced, approximate mocks and N-body simulation-based mocks.

EZmocks : EZMOCKS are approximate mocks that reproduce the clustering of the original samples. They are sampled according to the radial and angular distributions of their respective tracers and are downsampled according the observational systematic effects. EZMOCKS are based on the Zel'dovitch approximation to generate a dark matter field at a given redshift (Chuang et al., 2015). The process of the creation of the EZMOCKS specific to each eBOSS samples is detail in Zhao et al. (2020). Each tracer sample is attributed a set of 1,000 realizations of light-cone mocks and respective random catalogues. The latter are required to properly normalize the clustering measurement without incurring any bias in the estimation of the mean density estimation. The fiducial cosmological model used for constructing the EZMOCKS is flat ΛCDM with:

$$\begin{aligned} \Omega_m = 0.307,\ \Omega_b = 0.0482,\ h = 0.678, \\ \sigma_8 = 0.8225,\ n_s = 0.96. \end{aligned} \tag{3.6}$$

which are the best-fit values from the Planck 2013 results (Ade et al., 2014).

In order to assess the eventual modelling systematics, N-body based mocks are provided which correspond to each tracer specificities.

OuterRim OUTERRIM mocks were created in the framework of the eBOSS mock challenge whose purpose was to provide N-body based mocks to study eventual systematic effects of the HOD models on standard galaxy clustering measurements. These mocks are based on the N-body OUTERRIM simulation (Heitmann et al., 2019a,b) of $10,240^3$ particles in a $(3\ h^{-1}\mathrm{Gpc}\ )^3$ volume and built from snapshots of the simulation. The underlying cosmology for OUTERRIM simulation is close to the best-fitting model from WMAP-7 (Komatsu et al., 2011):

$$\begin{aligned} \Omega_m = 0.2648,\ \Omega_b = 0.0448,\ h = 0.71, \\ \sigma_8 = 0.8,\ n_s = 0.963. \end{aligned} \tag{3.7}$$

From the sets of OUTERRIM haloes simulations are derived two sets of mocks, tuned to the specificities of the eBOSS ELGs and QSOs such as the galaxy masses $M_\odot$, tracer density and effective redshift.

- OUTERRIM ELG mocks are built from a single snapshot at $z = 0.865$, close to that of the DR16 ELG sample. Six sets of mocks were produced, each with a different Halo Occupation Distribution (HOD) model. The detailed description of the mock construction and HOD models can be found in Alam et al. (2020b). In this work, we only use one blind mock of the ELG mock challenge with a galaxy number density similar to that of the data and populated with the HMQ3 (HighMassQuenched-3) HOD model. This mock contains 30 pseudo-independent realizations with periodic boundary conditions.

– OuterRim QSO mocks are built from a snapshot at $z = 1.433$. From this snapshot, 20 sets of mocks were created and populated with 20 different HOD models. In order to include the effect of quasars redshift uncertainties, an additional redshift smearing was added to mocks, providing 4 variations of the same mock with a redshift smearing of varying intensity. The detailed description of the mock construction, HOD modelling and redshift smearing along with their impact on standard clustering measurements are described in Smith et al. (2020). We use a 'non-blind' mock populated with the HOD10 model in a realistic redshift smearing configuration. It contains 100 pseudo-independent realizations with a tracer density comparable to that of the QSO sample.

The LRG sample does not have a specifically tuned N-body simulation to test the modelling systematics. Instead, the Nseries mocks are used. These mocks are N-body simulation snapshots populated with a HOD model. They were generated in the framework of the DR12 of the BOSS collaboration (Alam et al., 2017) to reproduce the BOSS CMASS sample at the effective redshift $z = 0.56$. The number of available realizations is 84 and the underlying cosmology for Nseries mocks used to generate them is:

$$\begin{aligned} \Omega_m = 0.286,\ \Omega_b = 0.0470,\ h = 0.700, \\ \sigma_8 = 0.82,\ n_s = 0.96. \end{aligned} \tag{3.8}$$

3.1.3 Void Identification in eBOSS

In this thesis, I make use of two void finding algorithms vide and revolver, with a common root that is the zobov algorithm at their core, both are described in section 2.2. In considering the eBOSS tracers, vide was found lacking in the sense that it did not take into account the weighting schemes of the galaxies in survey samples. In addition, some issues were encountered when considering small footprint such as the ELGs. As a consequence, the use of revolver appeared to be the most appropriate alternative due to its similarity with vide.

While both algorithms generate buffer particles to take into account the finiteness of the footprint and overall survey, they do not do so in the same way. As there are no 'real void' catalogues to refer to when qualifying a void finding algorithm, both algorithms were confronted with the same datasets (a set of eBOSS mocks) in order to decipher if both extract voids with similar properties.

The comparison is based on the redshift distribution and related void positions, as well as their radii which are dependent on the cosmology. These quantities enters directly into the estimation of the clustering of voids.

The catalogues investigated are 500 mock catalogues mimicking the eBOSS only LRG galaxy catalogue in the redshift range $[0.55, 1.05]$. The comparison is done between both void finding algorithms at the minimal post-processing steps, that is, preliminary removal of objects which are decidedly not voids in the catalogues.

3.1.3.1 Impact of the buffer particles proximity

As presented in Section 2.2.3.1, both algorithms place buffer particles in order to bound the volume of the survey that will be tessellated. This step is compulsory when using light cone or survey data as, unlike cubic simulation, the geometry of the survey is complex and non-homogeneous – some vetoed part of the mask have to be discarded.

The placement of buffer particles becomes some kind of trade-off the computing time – in the big data era, this is quite an important part – and the precision of the tessellation. The more buffer particles are placed around the survey, the tighter is the constraint on the volume at the tessellation stage, leading to better local density estimation. The placement of the buffer is limited to the periphery of the survey in order to optimize computing time.

While the generation of the buffer within the survey footprint cannot be quite modified, the placement of the buffer at the redshift edges can. This positioning scheme can therefore serve as a basis of comparison to compare both algorithms, in addition to its impact on the derived properties. We investigate the proximity of the buffer particles at the redshift edges of the survey, which can be decided by setting redshift boundaries at which to place the buffers.

Two redshift boundaries were chosen: a WIDE range: $[0.4, 1.2]$, wider than the original redshift distribution and a NARROW range $[0.55, 1.05]$, which corresponds to the redshift range of the mocks. Both void finders were run on each of the 500 mock catalogues in both boundary schemes. It is expected that the narrow range should provide the 'best' void sample, in the sense that it bounds more tightly the surveyed volume.

Table 3.1 summarizes the total number of voids found depending on the algorithm and buffer proximity considered, averaged over the 500 mocks. The proximity of the buffer particles seems to affect more strongly the void recovery in the case of the REVOLVER algorithm, since the average number of voids drops from 2360 to 1480, thus 880 voids, while VIDE has lost only 69 voids depending on the redshift boundary. In both schemes, VIDE finds a larger number of voids. This can be explained by the fact that VIDE actually disposes a very small amount of buffer particles at the survey edges compared to REVOLVER. The majority of the buffer particles is concentrated in the cones bounding the survey footprint between 0 and $D_c^{\max}$. This buffer placement may also explain the stability of the void count from one redshift scheme to another. On the contrary, REVOLVER sticks to a buffer placement close to the survey volume, which makes it more sensitive to the variation of the redshift limits.

Table 3.1: Number of voids found on average in VIDE and REVOLVER algorithms, depending on the proximity of the buffer particles to the redshift boundaries.

Algorithm	WIDE $[0.4, 1.2]$	NARROW $[0.55, 1.05]$	ΔN_v
VIDE	2509	2440	69
REVOLVER	1480	2360	880
Difference	1029	80	-

Figure 3.2 displays the impact of the proximity of the buffer particles on the redshift distribution in the WIDE and NARROW cases. For the VIDE algorithm, the redshift distributions of the voids are slightly affected by the redshift domain used. On the contrary, for the REVOLVER algorithm, the choice of the WIDE redshift boundaries leads to a very sensitive effect with a drop in the statistic between $z = 0.6$ and $z = 0.8$. This confirms the impact of the redshift boundary placement particles deep within the survey.

There is a good agreement between the two algorithms in the WIDE configuration from $z = 0.82$ onwards, and over the redshift range in the NARROW case. In the NARROW case, we note however that an excessive number of voids is present at low and high redshifts, for both algorithms.

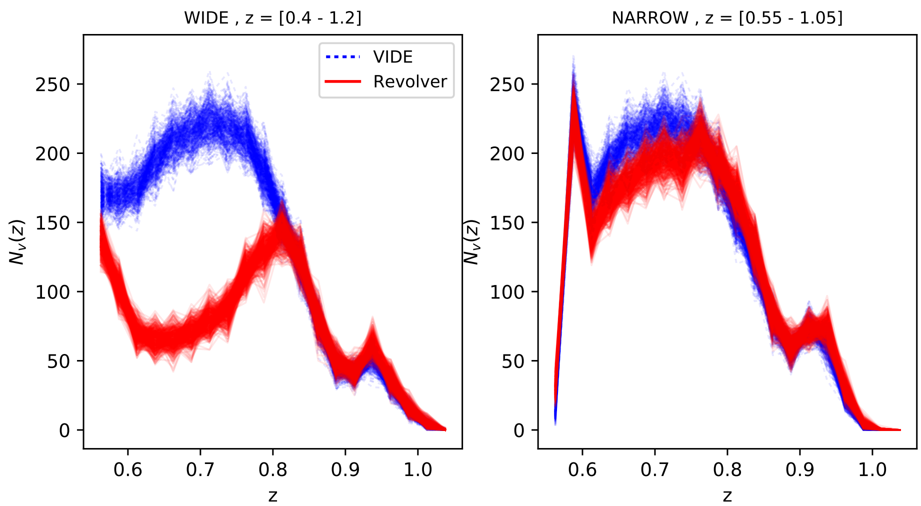


Figure 3.2: Redshift distributions of the voids extracted from REVOLVER (red) and VIDE (blue) in the WIDE (left) and NARROW (right) cases.

The distributions of the sizes of the voids (in terms of their effective radius) are shown in Figure 3.3. While the difference of statistics is the primary difference between both algorithms in the WIDE case, the radius distributions seem to reach a common ground in the larger voids found and evidence a similar behaviour

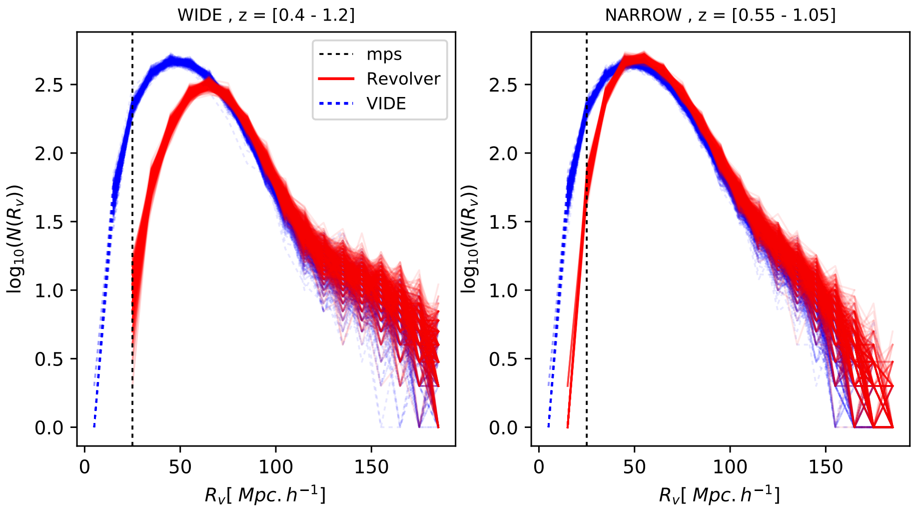


Figure 3.3: Radius distributions of the voids extracted from REVOLVER (red) and VIDE (blue) in the WIDE cas (left) and NARROW (right)

at high radii. In the NARROW case, the distributions are in agreement starting at lower radii. Both void finders display a similar excess of large voids. This feature is especially seen in the WIDE case in the form of an "ankle" like feature starting at $\sim$ 110 h^{-1}Mpc while it is dampened in the NARROW case. In both cases, REVOLVER find larger voids than VIDE. This change in the distribution of the radius evidences the impact of the boundary in the volume definitions of voids. If the volume is ill-defined, larger and more shallow voids will probably be found: Poisson voids. The addition of tighter boundaries dampens the identification of large under-densities but increases the number of small voids found. When comparing the number of voids found and the radius distribution, we can assure that the voids found in excess by VIDE compared to REVOLVER are small voids.

Finally, the edge contamination of the void sample is clearly shown by the excess of voids found near the redshift boundaries. However, the classification of those voids, cosmological or spurious is not easily identified. It could be that the proximity of the buffer particles leads to a prevalence of smaller void-like entities that may not be merged with other local under-densities due to the percolating effect of volume estimation. That is, a random particle too close to the survey will cause a tessellated particle to be part of a too large or too small volume depending on its minimal redshift and angular coordinates, this volume, when compared, will be used or discarded when building the under-dense zones.

The issues raised by this edge contamination show that we cannot completely trust these excessive under-densities not to compromise the cosmological signal. However, the behaviour displayed by both algorithms VIDE and REVOLVER indicates that the NARROW range is the best configuration to use for analysis.

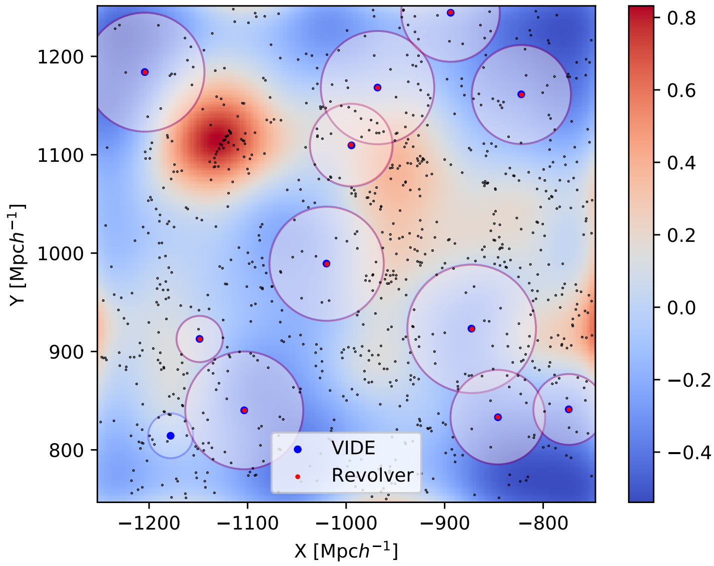


Figure 3.4: A slice of $500h^{-1}$Mpc × $500h^{-1}$Mpc and 50 h^{-1}Mpc deep within the central region of one mock. The black dot corresponds to galaxies as matter tracers, red dots to REVOLVER void centres and blue dots to VIDE void centres. The size of the surrounding circles corresponds to the radius of voids.

3.1.3.2 Comparison of the void positions

The similarity between the radius and redshift distributions between VIDE and REVOLVER hints at similar void centre positions, provided that the same definition is used. Here, the barycentre as the void centre is considered. Fig. 3.4 shows a 500×500 h^{-1}Mpc $X - Y$ slice of depth 50 h^{-1}Mpc in one of the most populated parts of a given mock chosen among the 500 on which the comparison was applied. The redshift boundary scheme is NARROW, as this configuration is the one where the two void finders have the most similarities.

The voids found by both VIDE and REVOLVER clearly trace the same under-densities, although some void centres are slightly shifted from one void finder to another. The sizes of the void in this part of the survey are similar. This is a strong confirmation that both algorithms find the same under-density field and that they can be used interchangeably.

We can note however the presence of a smaller void found only by VIDE, which confirms that the main statistical difference between VIDE and REVOLVER lies in small voids.

3.1.3.3 Mitigation of the edge contamination

The behaviour of the void finding algorithm in regard to the proximity of the buffer particles uncovered non-negligible edge contamination in the recovered redshift distributions, and, to a higher extent, the radius distribution. As there is no discriminating factor between the excess voids and underlying true voids present in the 'horns' of the survey, it was decided to discard these voids, while minimizing the number of objects cut out.

The adopted approach is to consider the volume of the under-densities, as their extent. The voids volume are confronted to their distance to the redshift edges, those being the most subject to the edge contamination. The voids are considered to be voluminous objects whose proximity to the survey boundaries would cause this volume to be truncated. On the basis of symmetry, we consider that a void at the edge of the survey is likely to be continued by the same amount out of the survey, so we base our void selection on a distance comparison to the border of the survey.

The information of interest is the radius of the voids r_v, the distance of the voids to the nearest redshift edge d_{edge} as well as the distance between the void centre and the farthest galaxy defining the void r_{max}. From this information are defined four primary cuts which associate the aforementioned quantities:

- r_v cut: The void is discarded if its distance d_{edge} from the nearest redshift edge is less than its radius r_v ,
- $2r_v$ cut: The void is discarded if its distance d_{edge} from the nearest redshift edge is less than twice its radius r_v ,
- r_{max} cut: The void is discarded if its distance d_{edge} from the nearest redshift edge is less than both its radius r_v and its maximal extent r_{max},
- Directional r_{max} cut: The void is discarded if the r_{max} cut condition is fulfilled in the case where the r_{max} galaxy is situated between the nearest redshift edge and the void centre.

These primary cuts were all tested in both WIDE and NARROW configurations. While it is considered that the NARROW boundary allows for a better estimation of the volume of the Voronoi cells and the subsequent void properties, the use of these cuts on the WIDE sample is quite telling on the reliability of the voids at the boundaries in this configuration, when compared to the NARROW distributions.

Fig. 3.5 and Fig. 3.6 show the impact of the primary cuts on the redshift distribution in the WIDE and NARROW configuration, respectively, for both VIDE and REVOLVER algorithms. Both figures display the redshift distribution of the voids extracted from a single mock realization. We can see that the voids at the edges are easily removed in the WIDE frame with a simple cut in terms of the void radius. This means that voids found in the WIDE scheme tend to have probably

Table 3.2: Relative difference between average number of voids before and after primary cuts over 500 mock realizations.

	VIDE		REVOLVER	
	WIDE	NARROW	WIDE	NARROW
r_v cut	−6.02%	−0.2%	−10.67%	−0.38%
$2r_v$ cut	−17.19%	−10.72%	−26.95%	−15.2%
r_{max} cut	−14.14%	−6.4%	−23.36%	−10.69%
D-r_{max} cut	−8.03%	−2.9%	−13.64%	−4.17%

overestimated volumes due to the wide placement of the boundary particles at the redshift edges, regardless of the algorithm used. In the NARROW range, however, the removal of the edge voids is not so evident with a simple r_v cut, meaning that the volumes and subsequent radius of the void are better bound by the proximity of the buffer and do not allow for an overestimation of the volume.

Visually, the most useful cuts to get rid of excess voids at the redshift limits seem to be the $2r_v$ cut and the r_{max} cut, without taking directionality into account. The $2r_v$ cut is more efficient, however, than the r_{max} cut, which seems to be confirmed in the table 3.2 that shows the loss percentages due to primary cuts.

The WIDE frame is the most sensitive to cuts, especially in the case of REVOLVER as it can cut out up to 27% of voids. In the NARROW case, the r_v cut has a negligible effect, while the $2r_v$ cut is the most stringent cut. The r_{max} cut seems to provide the best trade-off between cutting too many voids and discarding the edge contamination.

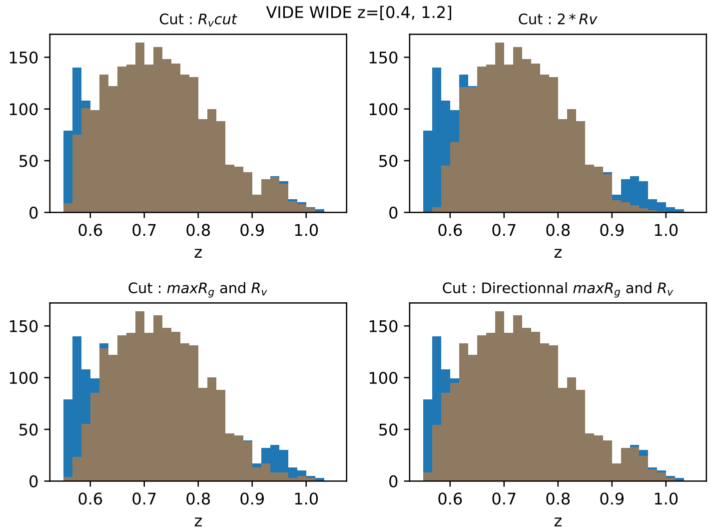
3.1.3.4 Final Selection cuts

Additional selection cuts are applied to voids found by both void finders. A void extracted by VIDE or REVOLVER is defined by a specific number of particles which volumes and positions are used to define the void radius and the void centre. To characterise our final void sample after performing the void finding, the r_{max} cut is applied to remove the visible edge contamination.

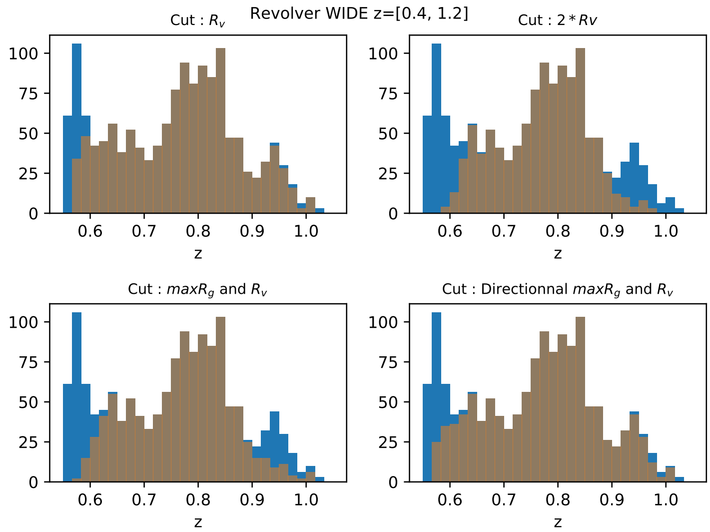
In the case of REVOLVER, the proximity to the footprint buffer mocks and their subsequent contamination are probed during the void finding, yielding a flag telling whether a void is considered as edge or not: the `Edge flag`. VIDE does not provide such information, relying instead on additional cuts linked to the proximity to the footprint buffer particles. For the purpose of our analysis, the void finder of choice is decided to be REVOLVER.

In addition, a selection on the number of galaxies/tracer particles is performed. The minimal number of tracers to define a void is usually two in void algorithms. However, in order to use properly defined voids, we reject all voids with less than five tracers.

Finally, the final selection cuts applied to voids are the $N_{part} > 5$ to remove

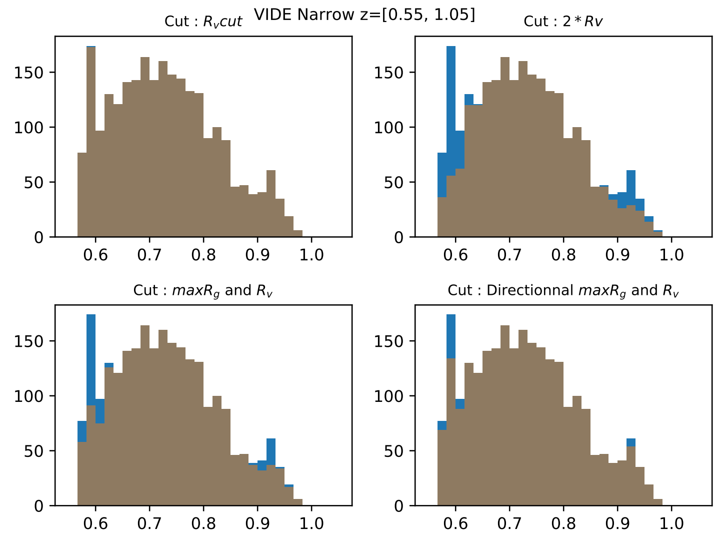


(a) VIDE

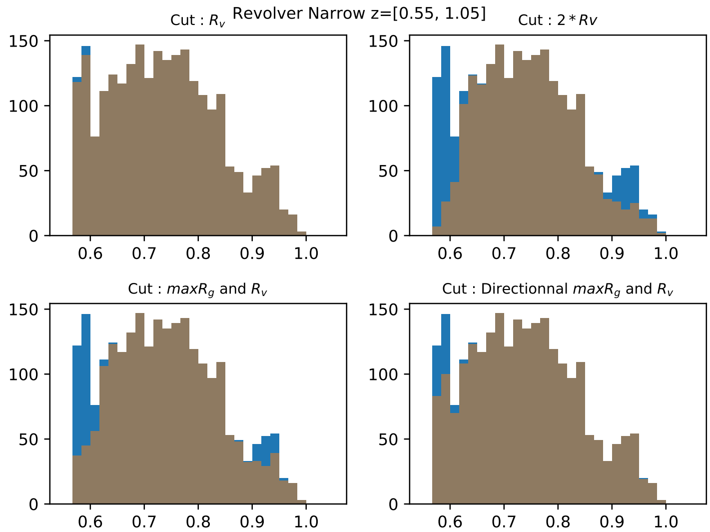


(b) REVOLVER

Figure 3.5: Impact of primary cuts on the WIDE configuration, for VIDE and REVOLVER algorithms. The distribution of voids is shown as a function of the redshift before (blue) and after (brown) primary cuts.



(a) VIDE



(b) REVOLVER

Figure 3.6: Impact of primary cuts on the NARROW configuration, for VIDE and REVOLVER algorithms. The distribution of voids is shown as a function of the redshift before (blue) and after (brown) primary cuts.

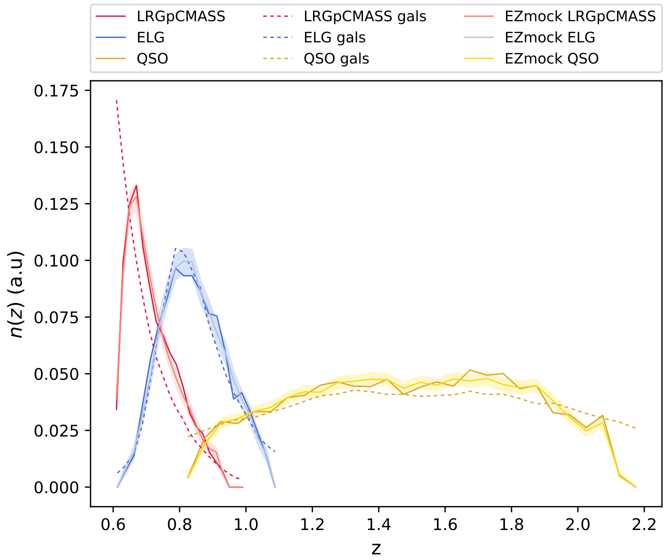


Figure 3.7: Redshift distribution of voids in normalized units after selection cuts for LRGpCMASS samples (red lines), ELG samples (blue lines) and QSO samples (yellow lines). The solid lines correspond to both data and mean of the 1000 realizations of the EZmocks. The shaded areas indicate the 1σ regions evaluated from the 1000 EZmocks realizations. The normalized number density of the galaxies/quasars is drawn for information in dashed line.

poorly defined voids, along with the `Edge Flag` and r_{max} cut for the light-cone samples, those allow us to keep only voids considered to be reliable.

3.1.4 Final DR16 voids catalogues

The REVOLVER void finding algorithm was run on the DR16 tracer samples and their associated thousand realizations of EZmocks and N_{NB} N-body mocks.

The redshift distribution respective to each sample is shown in Fig. 3.7 along with the galaxies n(z), in arbitrary units. First, there is a very good agreement between the redshift distributions of the approximate mocks and their respective dataset. Secondly, the redshift distribution of voids tends to follow that of their respective galaxy sample. However, it can be noticed that the LRGs sample suffers from the selection cuts at low redshifts. The abundance of galaxies in the low z range is degenerate with the effect of the void finder. It results that a large number of voids are cut out of the sample.

The size distribution of voids is displayed in Fig. 3.8. It is quite clear that each void sample has a specific scale of reference: the ELG voids represent the smallest population with voids smaller than $125h^{-1}$Mpc , while void radii from QSOs reach nearly up to $200h^{-1}$Mpc . The sizes of the voids are highly correlated to tracer

Table 3.3: Statistics of void catalogues identified in EZMOCKS catalogues and eBOSS DR16 LSS catalogues. The quantity N_g is the number of galaxies or quasars, N_v and $N_{v,cut}$ are the averaged number of voids and their standard deviation over the 1,000 mocks and data realizations before and after selection cuts as described in 3.1.3.4, respectively. The quantity z_{eff} is the effective redshift of the void catalogues after selection cuts. s_{max} is the maximum separation used in the correlation function.

Sample	N_g	N_v	$N_{v,cut}$	z_{eff}	s_{max}	Area (deg^2)
EZmocks						
ELG	173,736	$2,210 \pm 35$	$1,895 \pm 37$	0.847	3.60	1,170
LRGpCMASS	380,190	$4,305 \pm 54$	$2,850 \pm 47$	0.740	3.52	9,493
QSO	343,700	$5,449 \pm 53$	$4,321 \pm 52$	1.478	3.52	4,808
Data sample						
ELG	173,736	$2,097 \pm 5$	$1,801 \pm 5$	0.847	3.60	1,170
LRGpCMASS	377,458	$4,228 \pm 11$	$2,814 \pm 12$	0.740	3.52	9,493
QSO	343,708	$5,451 \pm 8$	$4,347 \pm 9$	1.478	3.52	4,808

specificities, such as the mass, bias and general sparsity of the survey (or number density) (Jennings et al., 2013; Nadathur & Hotchkiss, 2015b; Sutter et al., 2014a). In our case, the ELG sample is obviously the densest sample while the QSO is the sparsest, so it is not surprising to see that their size reflects this property.

The resulting void statistics are shown in Table 3.3 for each DR16 samples. The total cuts amount to a 34%, 14% and 20% loss in statistics for the LRGpCMASS, ELG and QSO samples, respectively. The cut's stringency is dependent on both the shape of the redshift distribution and the average size of the voids. Indeed ELGs have a quite gaussian redshift distribution and present generally smaller radius compared to the other two samples. As a result, these tracers are less affected by the selection cuts. On the contrary, the LRGpCMASS sample presents a decreasing redshift density distribution where its highest density is situated at the lower edge of the sample, the R_{max} cut thus affect the LRG sample much more strongly. In the case of QSOs, the redshift distribution is quite homogeneous but their scale is greater, leading to more voids crossing the edges of the surveys.

Another crucial information quoted in Table 3.3 is the effective redshift of the void-galaxy sample. Indeed, the voids statistics and clustering are estimated at a specific redshift. It is thus necessary to define the effective redshift for each sample which is attributed to our corresponding measurement. The effective redshift is estimated from the following void-galaxy pair count:

$$z_{\text{eff}} = \frac{\sum_{ij} w_i (z_i + Z_j)/2}{\sum_i w_i} \tag{3.9}$$

where z_i is the redshift of the i^{th} galaxy, Z_j the redshift of the centre of the j^{th} void,

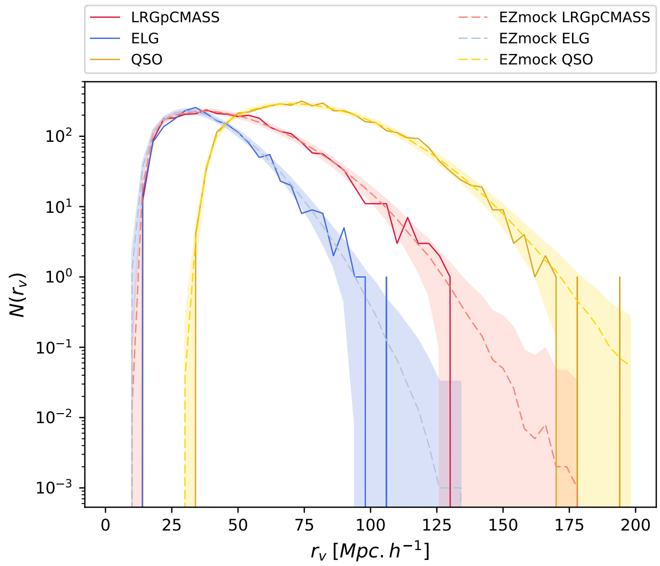


Figure 3.8: Number of voids after selection cuts as a function of their radius r_v for LRGpCMASS samples (red lines), ELG samples (blue lines), and QSO samples (yellow lines). The solid and dashed lines correspond to the data and the mean over the 1000 realizations of the EZmocks, respectively. The shaded areas indicate the 1σ regions evaluated from the 1000 EZmocks realizations.

and w_i the total weight of the i^{th} galaxy, as given by Eq 3.5. The void-galaxy pairs are accounted for only if they contribute to the two-point void-galaxy correlation function in the range $0 <= s < s_{\text{max}}$, where s_{max} is the maximal separation probed in the two-point correlation function, rescaled by the void radius r_v. The s_{max} respective to each sample can also be found in Table 3.3.

Overall, the EZmocks and the data present quite a good agreement in terms of both redshift distribution and size distribution, as well as their void count.

3.2 Estimation of redshift-space distortions around voids in the linear approximation

The interaction and formation process of the Large-Scale Structure is a dynamical process which evolves over time. Tracers of the LSS carry this dynamical information through what is called the Kaiser effect. Indeed, galaxies positions that we observe are affected by different physical interaction taking place in the vicinity of the galaxies. Galaxies move in a coherent flow along filaments and clusters, shaping the LSS at different epoch. These dynamical contributions, called peculiar velocities, provide valuable information on the laws of gravity that govern the movement of galaxies.

3.2.1 Redshift space distortions

Redshift surveys measure the redshift of a luminous object in the sky. As a result of the expansion of the Universe, the farther the object is, the faster it is moving away from us and the more the wavelength of the emitted spectra shifts toward longer wavelengths. The deviation between the emitted wavelength and measured wavelength, due to the Hubble flow, is called the cosmological redshift. However, the observed objects also have a motion of their own within a comoving frame. Hence, the measured redshift is perturbed by the velocities v_p of the galaxies, introducing distortions in real space. It can then be developed in two kinds of contributions:

$$z = z_c + v_p/c, \tag{3.10}$$

where the term v_p/c accounts for the impact of the dynamics of the objects. These peculiar velocities can themselves be decomposed in two types: linear and non-linear. The linear part results from the motion of the object induced by the growth of gravitational instabilities in an expanding Universe. This process is part of the creation and definition of the Large-Scale structure wherein the objects move in coherent flows toward highly over-dense objects, while the non-linear component relates to random velocities at small scales. These peculiar velocities lead to an observational effect, as it interferes with the redshift measurement and induces distortions in the radial component that are called *Redshift-Space distortions* (RSD).

Figure 3.9 illustrates the two kinds of effect induced by the velocities as observed in redshift space:

– *Linear velocities*: Those velocities are most prominent on large scales and relate to the continued growth of the large-scale structure. They materialize as infall velocities in the case of over-dense structure and outflow velocities in the case of voids. As a result, over-dense structures appear squashed along the line-of-site, while under-dense structures tend to be 'stretched'. This apparent distortion is called the *Kaiser effect* (Kaiser, 1987).

– *Non-linear velocities*: They spur mainly from the non-linear collapse of an over-density, resulting in an apparent elongation of the over-dense structure along the line-of-sight, called the *Finger-of-God effect*. Although Fig. 3.9 shows a squashing of an under-dense structure due to the non-linear velocities, it is actually not quite clear how those peculiar velocities affect the observed voids. They can cause voids to disappear, and the general agreement is that the smaller the voids, the more likely to be highly affected by non-linear effects.

These distortions deeply affect the density field estimated from luminous tracers of matter distribution that can be probed through the power spectrum or the two-point correlation function. Thereafter, we will neglect the study of non-linear

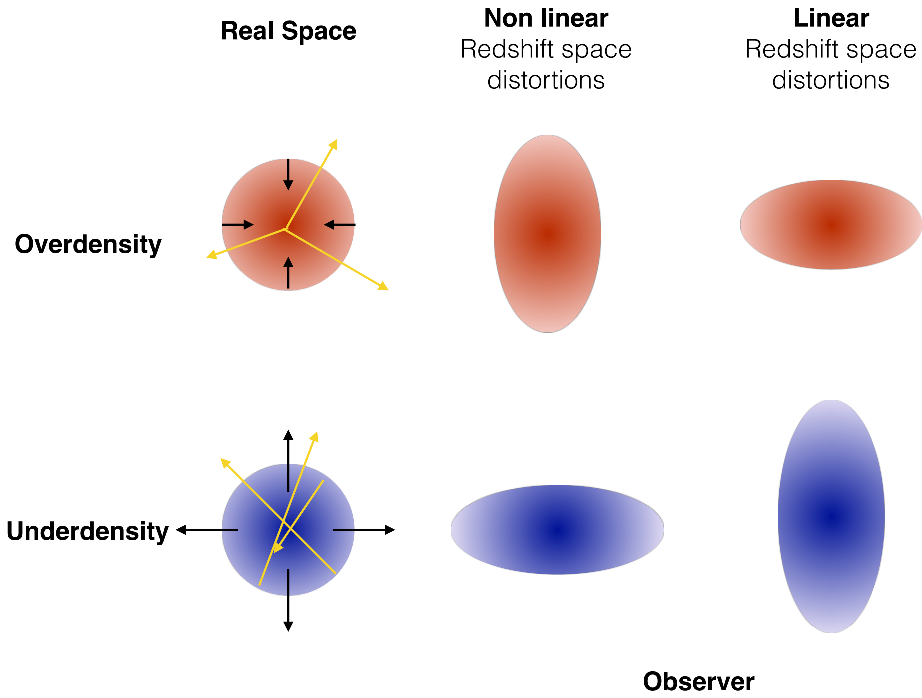


Figure 3.9: Illustration showing how real-space structures (left row) are affected by non-linear (middle row) and linear (right row) velocities, in the case of an over-density (red) and under-density (blue). Adapted from Dodelson (2003, Fig. 9.11 p.277)

velocities when considering the RSD around voids as those structures are considered to have a linear relation to the density field compared to over-dense tracers such as galaxies (Ceccarelli et al., 2013; Hamaus et al., 2014b; Lambas et al., 2016). Thus, the main contribution of RSD in our studies is the Kaiser effect. The easurement of the RSD signal is a powerful probe, as in the linear regime, velocities are directly connected to the growth rate f:

$$v(r) = -\frac{1}{3} f a H \Delta(r) r, \tag{3.11}$$

where a is the scale factor of the universe, H is the Hubble parameter and $\Delta(r)$ the volume-averaged density contrast defined as follows:

$$\Delta(r) = \frac{1}{r^3} \int_0^{r'} \delta(r') r'^2 \, dr \ . \tag{3.12}$$

In consequence, voids are quite interesting objects due to their very linear relation to the density field. The potential of voids as a sensitive probe of the growth rate of structure f have been investigated since the first proof of their sensitivity to dynamical effects in N-body simulations (Padilla et al., 2005). While at first set aside due to the lack of statistics in observational data, the analysis of redshift space distortions with voids regained favour with the first release of the DR7 data sample in which the growth rate was successfully recovered (Paz et al., 2013). Since then, constraints of the growth rate of structure with voids have been tirelessly investigated in numerous surveys: 6dF (Achitouv et al., 2017), VIPERS (Hawken et al., 2017), BOSS (Achitouv, 2019; Hamaus et al., 2017, 2020, 2016; Nadathur et al., 2019a) and eBOSS DR14 (Hawken et al., 2020) with growing accuracy and refinement of the modelling. This work is in the context of the DR16 eBOSS data release in which we probe the growth rate of structure in three different tracer samples at three different epochs. The model used to estimate the growth rate of structure will be presented below.

3.2.2 Linear redshift-space distortions around voids

The main contribution of the redshift space distortions comes from the dynamics of the matter tracers used to reconstruct the voids. The modelling of linear redshift space distortions around voids then relies on the distribution of those objects in redshift space, relative to the void centre.

3.2.2.1 Key assumptions

The distribution of matter around voids is measured through the correlation function $\xi(r)$ which estimates the underlying density field $\delta(r)$. We are interested in the effect of the dynamics of the galaxies surrounding the void on the density field δ. Under the assumption that the void-galaxy pair count is conserved when transforming from real space to redshift space, the mapping between the real-space correlation function and its redshift space counterpart is the following:

$$[1+\xi(r)]dr^3 = [1+\xi(s)]ds^3, \tag{3.13}$$

where $\xi(r)$ denotes the real-space void-galaxy correlation function, r corresponds to real-space separation between the void-galaxy pairs and $\xi(s)$ denotes the redshift-space void-galaxy correlation function for which s corresponds to the redshift-space separation between the void-galaxy pairs.

The mapping between the separations s in redshift-space and r in real-space is fundamental, as it is this shift in the positions which permeates in the correlation function. Under the distant-observer assumption (also known as the plane parallel approximation), wherein the pair void-galaxy are sufficiently distant from the observer to consider that the line-of-sight (LOS) of the pair coincides (Kaiser, 1987), the mapping between r and s is given by:

$$\mathbf{s} = \mathbf{r} + \frac{\mathbf{v} \cdot \hat{\mathbf{X}}}{aH(z)} \hat{\mathbf{X}}, \tag{3.14}$$

where $\hat{X}$ is the unit vector in the LOS direction to the void centre and $\mathbf{v}$ is the velocity of the galaxy considered in the void-galaxy pair. The velocity vector $\mathbf{v}$ corresponds to the coherent velocity of the galaxies directed toward an over-density. When considering under-dense regions, the velocity describes an outflow which drives the galaxy out of the voids. It assumed to be directed isotropically along the radial direction, which gives:

$$\mathbf{v} = v(r)\hat{\mathbf{r}}, \tag{3.15}$$

where $v(r)$ is the linear velocity outflow as defined in eq. 3.11.

In addition to the linear assumption above, the bias in the void-galaxy correlation function is assumed to be linear in regard to the underlying void-matter density field and to have the same value as the effective bias of the tracer considered:

$$\xi_{vg}(r) = b_g \delta_{vm}. \tag{3.16}$$

3.2.2.2 The redshift space void galaxy correlation function

From the key assumptions of linearity defined above, it is possible to recover a formulation of the redshift-space to real-space mapping of the correlation function:

$$[1 + \xi(s)] = [1 + \xi(r)] \frac{dr^3}{ds^3}. \tag{3.17}$$

As the redshift-space correlation function is, originally, a distorted real-space correlation function, the distortion is encoded in the motion induced shift of the separation between voids and galaxies. Such shift can be quantified as the determinant of the Jacobian of the coordinate transformation (Nadathur & Percival, 2019):

$$|J\left(\frac{\mathbf{s}}{\mathbf{r}}\right)| = 1 + (1 - \mu^2)\frac{1}{aH}\frac{v}{r} + \mu^2 \frac{1}{aH}\frac{\partial v}{\partial r} \tag{3.18}$$

where $\mu = \cos\theta$ and θ is the angle subtended between the separation of the void galaxy pair $\mathbf{r}$ and the direction of the LOS. The derivative of the velocity yields:

$$\frac{\partial v}{\partial r} = aH(\delta(r) - \frac{2}{3}\Delta(r)) \tag{3.19}$$

Using the Jacobian of the coordinates transformation and the formulation of the velocity, equation Eq. 3.17 becomes:

$$[1 + \xi(s)] = [1 + \xi(r)][1 - \frac{f}{3}\Delta(r) - \mu^2 f(\delta(r) - \Delta(r))]^{-1}. \tag{3.20}$$

The equation can be expanded at first order yielding:

$$[1 + \xi(s)] = [1 + \xi(r)][1 + \frac{f}{3}\Delta(r) + \mu^2 f(\delta(r) - \Delta(r))]. \tag{3.21}$$

In developing the above equation to obtain a definition of the two-point correlation function between voids and galaxies, it is considered that the contributions of the $\xi(r)\Delta(r)$ and $\xi(r)\delta(r)$ are negligible (Cai et al., 2016). This assumption allows us to obtain the following linear modelling the redshift space void-matter two-point cross correlation function to linear order:

$$\xi(s, \mu) = \xi(r) + \frac{f}{3}\Delta(r) + f\mu^2(\delta(r) - \Delta(r)) \tag{3.22}$$

In the above equations $\xi(r)$ is actually considered a function of $\delta(r)$. The relation between the true underlying matter field is biased by both due to the mass and density of the matter tracer. In the case of this simple modelling, we assume a linear bias relation between the tracer bias and the underlying density field as described previously (Eq. 3.2.2.1). The two-point void-tracer cross correlation function thus becomes:

$$\xi(s, \mu) = \xi(r) + \frac{\beta}{3}\bar{\xi(r)} + \beta\mu^2(\xi(r) - \bar{\xi(r)}), \tag{3.23}$$

with the mapping between the real-space separation r and the redshift-space separation s defined as:

$$s = r(1 - \frac{\beta}{3}\bar{\xi(r)}). \tag{3.24}$$

3.2.2.3 Estimator of the correlation function

The correlation function of given discrete tracers can only be *estimated* from the available information. It directly relates to the joint probability of finding a pair at a distance $r = |\mathbf{x_1} - \mathbf{x_2}|$. An abundance of pair at a distance r should thus denote an excess probability of finding a pair. The estimation of the correlation thus needs to account for this excess in regard to a homogeneous random Poisson distribution of pairs, in which case $\xi(r) = 0$.

First of all, let us define the pair-count variables DD and RR: DD(r) relates to the number of pairs found at a distance r in bins of size Δr for the data information, RR(r) relates to the number of random pairs. We can thus define the estimate of the correlation function through the comparison of the DD(r) pair-count to the RR(r) pair-count:

$$1 + \hat{\xi}(r) = \frac{N_R(N_R - 1)}{N_D(N_D - 1)}\frac{DD}{RR}, \tag{3.25}$$

where the first term is a constant which normalises the estimated pair-counts:

N_D being the total number of data objects and N_R the total number of random objects. The association normalises the pair-counts $DD(r)$ and $RR(r)$ by the total number of pairs available in both data and random samples to take into account a different density of objects. This prescription is known as the Peebles-Hauser prescription (Peebles & Hauser, 1974). It relates to most optimistic dataset possible, in which the galaxies collected in a volume of the universe have all been accounted for and the total footprint of the galaxies present no holes.

In the case of the existence of a survey mask, which masks the regions where the observations are unreliable due to the presence of luminous sources such as bright stars or other systematical effects, the random and data distributions will thus be reliant on the holes in the footprint which may affect the estimation of the correlation function, especially when estimating the pair-count contributions at the edges of the mask. To mitigate these effects, we introduce the cross pair-count $DR(r)$ (or RD(r)) which measures the number of data-random pairs. A modified version of the previous estimator is the following:

$$1 + \hat{\xi}(r) = \frac{N_R N_D}{N_D(N_D - 1)} \frac{DD}{DR}, \tag{3.26}$$

where the normalisation term for the DR pair-count now includes all the considered objects in both samples. This estimator is known as the Davis-Peebles (DP) estimator (Davis & Peebles, 1983).

A final estimator of note to be presented represents a different approach to the estimation of the correlation function. Considering our samples to be an association of the signal data D and R, where the random contribution is removed from available data: $(D - R)$, the total available pair-counts relating to the signal, to compare to a random background should thus be the identity $(D - R)^2$ which considers all the possible association of the samples:

$$\hat{\xi}(r) = \frac{\frac{DD}{N_d(N_d-1)/2} - \frac{2DR}{N_d N_r} + \frac{RR}{N_r(N_r-1)/2}}{\frac{RR}{N_r(N_r-1)/2}}. \tag{3.27}$$

This estimator is called the Landy-Szalay (LS) estimator (Landy & Szalay, 1993) and presents the advantage of having the minimum variance when considering small perturbations ($\xi \ll 1$) which put it in the place of being the optimal estimator of the correlation function (Vargas-Magaña et al., 2013).

While the LS estimator is generally set as the estimator of choice in terms of galaxy clustering, its use is less trivial in the case of the void-galaxy cross-correlation. Indeed, the LS estimator needs to be provided with a random catalogue of void positions R_v. The process of creating such a random void catalogue is not trivial. While some authors use the angular positions and redshift positions (Achitouv, 2019; Nadathur et al., 2019a), it is not straightforward in our case as our correlation function relies on the rescaled separation $r_{res} = r/r_v$ while considering non-overlapping voids.

Two factors have to be taken into account: the random voids have to be non-overlapping extended objects and the random void catalogue has to be statistically superior to that of the data. The generation of a proper void radius distribution in the random sample is thus the problem here. One cannot simply draw a random radius from the existing radius distribution as it would probably lead to an overlapping random void distribution. A work-around solution could be to run the void finding algorithm on a random sample with the same number density as the data, but then the number of available random voids would not obey the superior statistics requirement. In addition, Hamaus et al. (2017) argued that the terms $R_v R_g$ and $D_g R_v$ were of negligible effect in the estimation of the two-point correlation function (2PCF). Therefore the estimator elected is the Davis-Peebles prescription which writes as:

$$1 + \hat{\xi_{vg}}(r) = \frac{N_{R_g} N_{D_v}}{N_{D_v} N_{D_g}} \frac{D_v D_g}{D_v R_g}, \tag{3.28}$$

in the case of the void-galaxy two-point cross-correlation function.

3.2.2.4 Multipole decomposition

Given the existence of galaxy dynamics around voids and their apparent privileged direction along the LOS, it is straightforward to understand that the isotropy of the estimated density field surrounding the void centre is perturbed. The resulting estimated correlation function is thus anisotropic and can therefore be expressed as a sum of even multipoles moment in the basis of Legendre Polynomials $L_l(\mu)$ (Hamilton, 1992, 1998):

$$\xi^s(r, \mu) = \sum_{\ell} (2\ell + 1) \mathcal{L}_\ell(\mu) \xi_\ell(r). \tag{3.29}$$

where μ is the cosine of the angle between the separation vector direction r and the line-of-sight. The odd multipoles cancel out to infinity (Cai et al., 2016; Hamilton, 1992, 1998). On the other hand, the only non-vanishing multipoles are the monopole and the quadrupole. The hexadecapole is expected to be null as opposed to the galaxy-galaxy correlation function. This is attributed to the fact that we assume that the void centres are devoid of dynamical effects.

The two-point correlation function thus reduces to:

$$\xi^s(r, \mu) = \mathcal{L}_0 \xi_0^s(r) + \mathcal{L}_2 \xi_2^s(r), \tag{3.30}$$

with first order Legendre polynomials:

$$\mathcal{L}_0(\mu) = 1, \tag{3.31}$$

$$\mathcal{L}_2(\mu) = \frac{3\mu^2 - 1}{2}, \tag{3.32}$$

and with the multipoles of the correlation being related to Eq. 3.23 by the following expressions:

$$\xi_0^s(r) = (1 + \frac{\beta}{3})\, \xi(r), \tag{3.33}$$

$$\xi_2^s(r) = \frac{2\beta}{3}\, [\xi(r) - \bar{\xi}(r)]. \tag{3.34}$$

Expressing the monopole contribution of the correlation function as a difference between the monopole and its volume averaged counterpart:

$$\xi_0^s(r) - \bar{\xi}_0^s = (1 + \frac{\beta}{3})\, [\xi(r) - \bar{\xi}(r)], \tag{3.35}$$

the multipoles can be reduced to a simple constant through the combination of Eq. 3.35 and Eq. 3.34, removing all dependence to the real-space correlation function. Therefore, this constant represents an estimate of the distortion parameter as presented originally by Cai et al., 2016:

$$G(\beta) = \frac{\xi_2^s(r)}{\xi_0^s(r) - \bar{\xi}_0^s(r)} \tag{3.36}$$

$$= \frac{2\beta}{3 + \beta}. \tag{3.37}$$

In practice, the estimated monopole and quadrupole both tend to the mean density of galaxy/matter $\xi(r) = 0$ at $r \to \infty$, which render the parameter constraint through the formula 4.43 unreliable. As such, we define the residual as follows:

$$\epsilon(\beta) = \xi_2 - (\xi_0 - \bar{\xi}_0)\, \frac{2\beta}{3 + \beta}, \tag{3.38}$$

where ξ_0 and ξ_2 are the measured quantity of the correlation function.

3.2.3 Parameter estimation

The modelling is a straightforward relation that allows a direct measurement of β while only requiring a measurement of the multipoles of the void-galaxy correlation function. To this end, the β parameter is fitted to the data in a frequentist approach using a χ^2 minimization defined as:

$$\chi^2 = \epsilon^T \, \Psi \, \epsilon \tag{3.39}$$

where ϵ is the residual given by Eq. 3.38 and Ψ is the precision matrix. It is to be noted here that the absence of a theoretical prediction regarding the multipoles, the monopole contribution, $\xi_0 - \bar{\xi_0}\frac{2\beta}{3+\beta}$, is used in its stead.

The precision matrix is estimated from the inverse of the covariance matrix. In practice, one estimates a covariance matrix using a set of N_s mock catalogues reproducing the data. It is constructed as follows:

$$C_{ij} = \frac{1}{N_s - 1} \sum_{k=1}^{N_m} (\epsilon_i^k - \langle \epsilon_i \rangle)(\epsilon_j^k - \langle \epsilon_j \rangle), \tag{3.40}$$

where N_s is the number of realizations, ϵ_i^k is the residual of the mock k in the bin i and $\langle \epsilon_i \rangle$ is the mean value of ϵ_i^k in the bin i such as:

$$\langle \epsilon_i \rangle = \frac{1}{N_s} \sum_{k=1}^{N_s} \epsilon_i^k. \tag{3.41}$$

Developing Eq. 3.40 by plugging the proper expression of the residual (3.38), we recover an extended expression of the covariance on a multipole per multipole contribution basis as prescribed in Cai et al., 2016:

$$C_{ij} = G^2 C_{00}^{ij} + C_{22}^{ij} - G C_{02}^{ij} - C_{20}^{ij}, \tag{3.42}$$

where $G = \frac{2\beta}{3+\beta}$.

The covariance is then inverted to recover the precision matrix Ψ. However, when inverting a noisy covariance matrix estimated with a set number of mocks N_s, it is necessary to correct for a possible bias. Hartlap et al. (2007) and Taylor et al. (2013) proposed a correcting factor to the covariance matrix to do just that:

$$\hat{\Psi} = \frac{N_s - N_b - 2}{N_s - 1} \hat{C}^{-1}, \tag{3.43}$$

where N_b is the number of bins and N_s is the number of mocks used to estimate the covariance matrix $\hat{C}$. While $N_b \ll N_m$, the bias on the precision matrix remains pretty low, for example, using a covariance matrix derived from 1000 mocks and 20 bins, the correction factor is about 2%, which is negligible in the case of a survey like eBOSS.

Another implication of using a set number of mocks and bins to build a covariance matrix is that the covariance in itself is an estimated quantity that possesses an uncertainty which causes the precision matrix to be biased. To take that into account, the parameter error obtained through the minimization procedure is rescaled with the following prescription (Percival et al., 2014):

$$\sqrt{m_1} = \sqrt{\frac{1 + B(N_b - N_p)}{1 + A + B(N_p + 1)}}, \tag{3.44}$$

where N_p is the number of parameters fitted, and

$$A = \frac{2}{(N_s - N_b - 1)(N_s - N_b - 4)} \tag{3.45}$$

$$B = \frac{N_s - N_b - 2}{(N_s - N_b - 1)(N_s - N_b - 4)}. \tag{3.46}$$

3.3 Investigation of systematics using mocks

We have shown in Sec. 3.1.4 that voids from mocks and data have similar properties, so we can apply the methodology described in Sec. 3.2 to the mocks in a blind manner before applying them to the final eBOSS data. We will take the opportunity to investigate a number of systematic effects. Moreover mocks are necessary to estimate the covariance matrix of the void-galaxy two-point cross-correlation function and therefore, estimate the error on our parameter β.

Two types of mocks were used in this framework: EZmocks which are approximate mocks that reproduce the clustering of the eBOSS DR16 and dedicated to the estimation of the covariance matrix for each sample (LRGpCMASS, ELG, QSO). They are also used to test for the methodology choices such as binning, weighting and void finding choices. The second set of mocks are N-body simulation-based mocks designed to specifically test the model used to estimate the growth rate of structure as the simulation of the RSDs is more accurate.

For both types of simulation, the fiducial β is estimated through the theoretical prediction of f at a given cosmology and redshift and the value of the bias b, provided by companion papers of the DR16 data release: Bautista et al. (2020) and Gil-Marín et al. (2020) for the LRGpCMASS EZmocks and the Nseries, De Mattia et al. (2020) and Tamone et al. (2020) for the ELGs EZmocks and OuterRim and (Hou et al., 2020; Neveux et al., 2020) for the QSOs EZmocks and the OuterRim. The use of the estimated bias by the standard galaxy clustering is necessary as the bias is not quite a controlled parameter in the simulations.

3.3.1 Approximate mocks

The approximate mocks reproducing the data are necessary in order to infer the covariance of our measurements of the correlation function and the subsequent error in the parameter constraint. The EZmocks mocks reproduce both the clustering of the DR16 tracers but also include some observational systematics, along with having the same footprint and number density as the data. The calibration of

Table 3.4: Statistics on the distortion parameter fit on the 1000 EZmocks (950 for the ELG sample) realizations for each eBOSS tracer. The reported $\langle\beta_{\mathsf{ref}}\rangle$ and $\langle\sigma_\beta\rangle$ are the mean of the N_m individual fit ouput. The χ^2 is normalized to the number of degrees of freedom.

EZmocks	β^{fid}	$\langle\beta^{\mathsf{ref}}\rangle$	$\langle\sigma_\beta\rangle$	$\langle\chi^2\rangle$	$\lvert(\langle\beta^{\mathsf{ref}}\rangle - \beta^{\mathsf{fid}})/\beta^{\mathsf{fid}}\rvert$
LRG+CMASS	0.372	0.414	0.072	1.39	11.29%
ELG	0.63	0.521	0.101	1.14	17.3%
QSO	0.403	0.294	0.049	1.76	27%

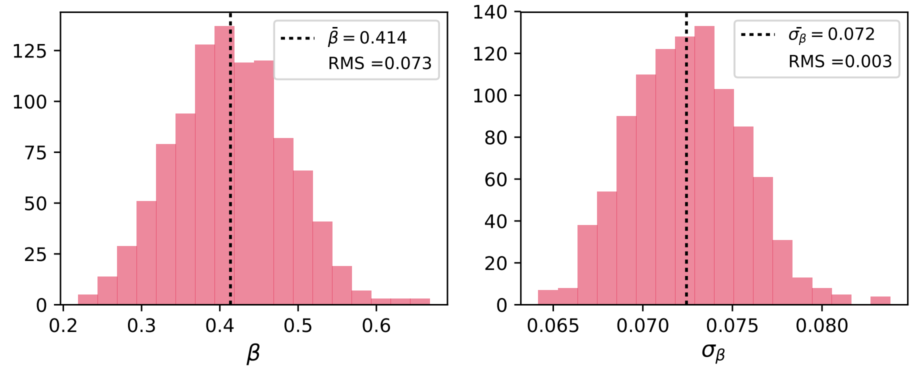
the clustering of the EZmocks on the data permeates in the comparison of the clustering of the voids in the data and those in the mocks, as can be seen in Fig. 3.15. The agreement between mocks and data thus allows us to quantify eventual systematics in our methodology.

3.3.1.1 Baseline analysis

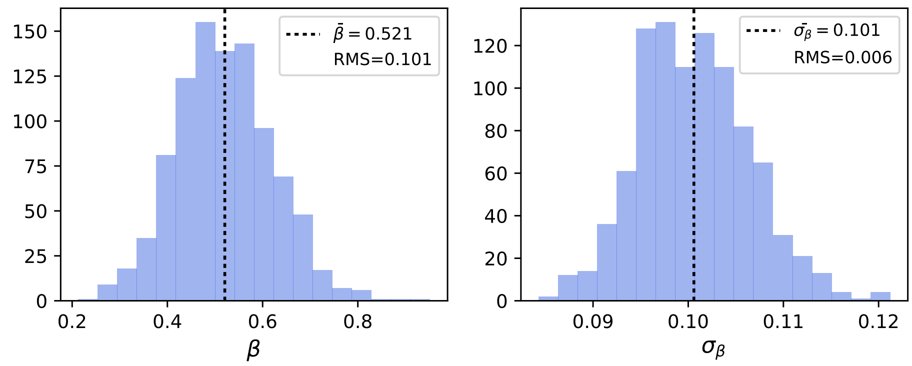
The pipeline, as depicted in the first section 3.2 is applied to each of the N_m realizations for all three tracers: ELGs, LRGs and QSOs. For each realization is obtained an estimate of β, using the covariance build from the $(N_m - 1)$ remaining mocks. The resulting distributions of the recovered β and their errors are featured in Fig. 3.10. The error distributions follow that of a Gaussian meaning that those are well estimated within the fitting procedure. Figure 3.11 presents the best-fitting residuals and associated β to the quadrupole of the void-galaxy correlation function for one realization among the thousand EZmocks.

The mean of the estimated β and σ_β are quoted in Table 3.4 as reference for systematics estimate. Results are quoted following the result on the binning analysis performed in order to optimize the parameter estimation. In the case of the LPGpCMASS and ELG samples, the recovered β values remain within or slightly higher than 1σ from the fiducial value, while the QSOs recovered value presents a more than 2σ deviation. The corresponding relative deviation is 11.3%, 17.3% and 27% for LRGs, ELGs and QSOs, respectively. It is to be noted that the EZmocks reproduce the data as well as their systematical effect which could contribute to this deviation, in addition to a possible systematic effect due to the modelling. The latter will be investigated on N-body mocks in the next section.

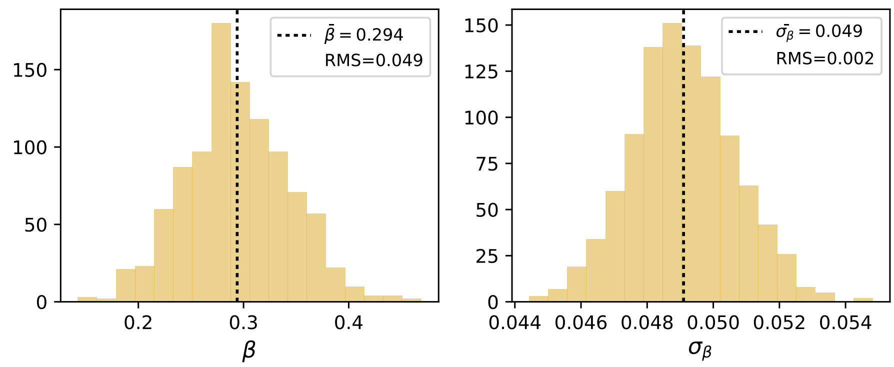
Our estimation of β relies on a ratio (difference) between the multipole contributions, as opposed to a fit of a theoretical prediction. Being a ratio of amplitudes, any impact on the estimation of the pair count can lead to a shift in the estimation of the multipoles and subsequent β parameter estimation. To this end, we check the robustness our β measurements regarding several methodology choices likely to affect the estimation of the cross-correlation function.



(a) EZMOCKS LRGpCMASS

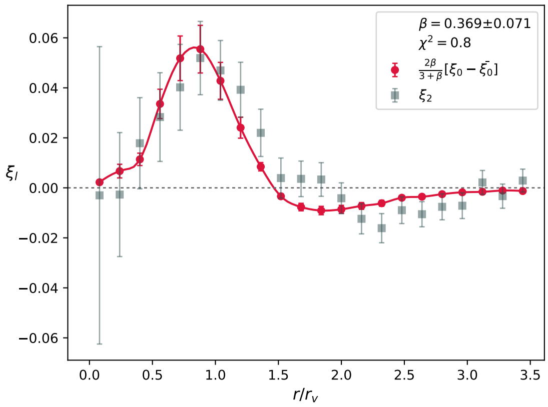


(b) EZMOCKS ELG

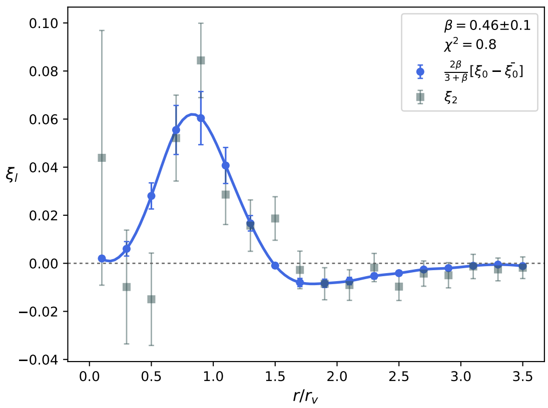


(c) EZMOCKS QSO

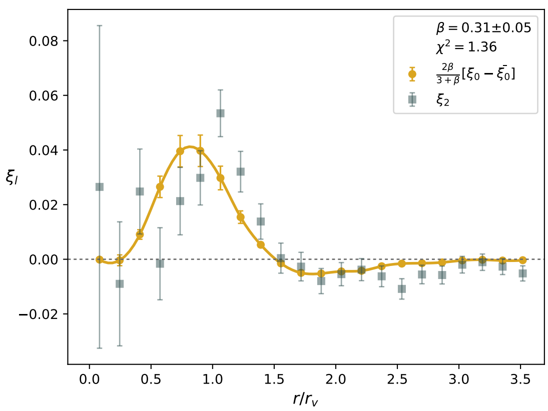
Figure 3.10: Best-fit parameters for the 1000 realizations of the EZMOCKS catalogues. Left panels display the distribution of the distortion parameter β and right panels display the distribution of the errors of β. The LRGpCMASS, ELG and QSO EZMOCKS samples are displayed in the top (a), middle (b) and bottom (c) panels, respectively.



(a) EZMOCKS LRGpCMASS



(b) EZMOCKS ELG



(c) EZMOCKS QSO

Figure 3.11: Quadrupole (ξ_2) and the best-fit of the $2\beta/(3+\beta)(\xi_0 - \bar{\xi_0})$ residuals from one EZMOCKS catalogue of the LRGpCMASS, ELG and QSO sample displayed in the top (a), middle (b) and bottom (c) panels, respectively. Error bars are the diagonal of the covariance matrix from the $N_s - 1$ remaining mocks.

3.3.1.2 Impact of the binning

The binning used to estimate the two-point correlation function may bring a change of amplitude in the estimated monopole and quadrupole. Hence, it is expected to have an impact on the estimation of β. Each of the N_m realizations is computed with several binning schemes between 14 and 22 bins at the maximum, thus affecting the fitting range.

The optimal binning scheme adopted is the one that represents a trade-off between the relative error on the parameter β defined as β/σ_β, which is used as a reference value and the minimization of the value of the χ^2. The resulting estimations are reported in Table 3.5 in terms of the χ^2, relative error, as well as the deviation from the reference value. Due to the difference in term of number density and resulting galaxy and void statistics, the number of pairs is different from one sample to another.

The β values seem to decrease along with the number of bins, which confirms the expected impact on the β estimations. The χ^2 also decreases monotonically with the number of bins, that is why it cannot be used alone as an indicator for an optimal binning scheme. It is to be noted that the ELG values quoted in the Table are different from those quoted in the baseline analysis. This discrepancy comes from the fact that the catalogue used was not generated with the same buffer particles density. However, the QSO and LRGpCMASS binning systematics remained unaffected by such change, so it was considered that this systematic check could be kept as is in the case of the ELG sample. The sensitivity to the binning scheme is also dependent on the statistics and footprint. While the LRGpCMASS and the QSO are the most populated sample, their footprint is different. The LRG is a combination of both the eBOSS footprint and the BOSS footprint, yielding an uneven number density throughout the footprint. Hence, when compared to the QSO sample that presents a similar number of object and a more homogeneous sample, the LRGpCMASS sample displays greater sensitivity to the binning. The ELG sample seems less consistent in terms of binning with deviations alternating from one order of magnitude to another, probably due to its lower statistical power.

While the selected binning scheme for each sample is used for the reference analysis of the mocks and all tests therein, the effect of the binning scheme on the estimation of β is taken into account as systematical effect as the largest deviation from the reference value, which corresponds to 0.02, 0.012 and 0.004 for the LRGpCMASS, ELGs and QSOs respectively. Those are reported in the total systematic budget as a deviation percentage from the reference β which represents a 4.8, 2.3 and 1.3 percent effect respective to the LRGpCMASS, ELGs and QSOs.

3.3.1.3 Impact of the weights

Weights are estimated for each galaxy to mitigate observational systematics such as the impact of photometry, fibre collisions and redshift selection. These weights are taken into account in both void finding and correlation function estimation.

Table 3.5: Impact of the number of bins N_b used for the estimation of the correlation function and fitting procedure. The χ^2 and relative error on β are displayed. The resulting deviations from the value of reference (in bold) for each sample are also displayed, with the error quoted as the error on the mean $(\sigma/\sqrt(N))$

EZmocks	N_b	$\langle\beta\rangle$	$\langle\sigma_\beta\rangle/\langle\beta\rangle$	$\langle\chi^2\rangle$	$\langle\beta\rangle - \langle\beta_{\mathsf{ref}}\rangle$
LRG+CMASS	16	0.434	0.177	1.57	0.020 ± 0.004
LRG+CMASS	18	0.432	0.178	1.51	0.018 ± 0.004
LRG+CMASS	20	0.43	0.178	1.48	0.016 ± 0.004
LRG+CMASS	**22**	**0.414**	**0.180**	**1.39**	**-**
LRG+CMASS	25	0.426	0.179	1.40	0.012 ± 0.004
ELG	14	0.508	0.214	1.57	0.011 ± 0.005
ELG	16	0.498	0.215	1.53	0.001 ± 0.005
ELG	**18**	**0.497**	**0.214**	**1.48**	**-**
ELG	20	0.490	0.217	1.45	-0.007 ± 0.005
ELG	22	0.485	0.219	1.42	-0.012 ± 0.005
QSO	16	0.298	0.169	2.04	0.004 ± 0.002
QSO	18	0.295	0.170	1.98	0.001 ± 0.002
QSO	20	0.293	0.169	1.89	-0.001 ± 0.002
QSO	**22**	**0.293**	**0.169**	**1.76**	**-**
QSO	25	0.292	0.170	1.66	-0.002 ± 0.002

Table 3.6: Impact of the FKP weight in regard to the reference β. Quoted error is the quadratic sum of the error on the mean of both β and β_{ref}

EZmocks	$\langle\beta\rangle - \langle\beta_{\mathsf{ref}}\rangle$
LRG+CMASS	0.006 ± 0.005
ELG	0.012 ± 0.005
QSO	0.001 ± 0.002

However, the FKP weights w_{FKP} are weights specifically designed to optimize the estimation of the amplitude of the clustering in the range of the BAO peak, in configuration or in Fourier space on the basis of the galaxies representing a Poisson sampling of the underlying matter field. As the correlation function estimation relies also on the positions of galaxies, the FKP weights are used in the correlation function. However, it is not sure how this methodology choice should affect the clustering of voids. To this end, the cross-correlation function is evaluated with and without the FKP weights to test for their impact on the estimation of β. The resulting deviations are transcribed in Table 3.6.

In a conservative approach, the deviation between the reference value and the mean β are taken as systematics and represent 1.4, 2.3 and 0.7 percent for the LRGpCMASS, ELG and QSO, respectively. In the case of QSOs, this correction represents a marginal effect below the percent level. Once again, it seems that the

Table 3.7: Impact of the correlation function estimator in regard to the reference β. Quoted error is the quadratic sum of the error on the mean of both β and β_{ref}

EZmocks	$\langle\beta\rangle - \langle\beta_{\mathsf{ref}}\rangle$
LRG+CMASS	-0.009 ± 0.004
ELG	0.017 ± 0.005
QSO	0.003 ± 0.002

impact of the FKP weights relies on the available statistics in the sample.

3.3.1.4 Impact of the estimator of the two-point correlation function

The DP estimator is our elected estimator of the two-point correlation function due to the grey zone that represents the generation of a random void catalogue in our configuration – the rescaling of the separation in the pair-count estimation. The LS estimator, however, present different properties of bias and variance in regard to the estimation of the correlation function (Landy & Szalay, 1993). The impact of the estimator is, therefore, to be taken into account in the precision on the parameter β. To this end, a simplified version of the estimator is investigated as prescribed in Hamaus et al. (2017), as the void random R_v term is negligible:

$$\xi^{LS}(r,\mu) \approx D_v D_g - D_v R_g. \tag{3.47}$$

The resulting deviation between the baseline β, estimated through the DP estimator and the resulting β estimated from the LS estimator, are recorded in Table 3.7.

The effect of the correlation estimator on the β estimation represents about 2.2% for LRG+CMASS, 3.3% for ELG and 1% for QSO.

3.3.1.5 Summary of systematics investigated on the EZmocks

Thanks to the EZmocks were tested several aspects of our methodology in the form of weights, estimator and binning choices. Table 3.8 shows the total systematic budget estimated from the EZmocks. The repercussion on the estimation of β behaves differently from one tracer to another. It is obvious that any methodology choice that affects the estimation of the two-point correlation function is responsible for these systematics.

The QSO is the most populated and homogeneous sample in terms of footprint, yielding a higher number of voids and is the less affected by systematics affecting the correlation function estimation representing a 1.7% effect on the parameter error. The LRGpCMASS sample is the most affected by the binning, followed by the estimator while the FKP is close to the percent effect. The ELG, smallest sample in terms of population of voids and galaxies, is the most affected by the change in methodology and consistently over the different systematics probed.

It seems then that the available statistics, homogeneity and footprint makes a sample more likely to be affected by our methodology choices. The more the statistics, the less the methodology choices affect the estimation of the correlation function and subsequent RSD parameter estimation.

Table 3.8: Systematic error budget estimated with the EZmocks and respective to each tracer

syst	LRGpCMASS	ELG	QSO
Binning	4.8%	2.3%	1.3%
Estimator	2.2%	3.3%	1%
FKP	1.4%	2.3%	0.7%

3.3.2 N-body mocks

Methodology choices were investigated with the EZmocks and the recovered β values are consistent with the fiducial β within 1σ except for the QSOs. In order to properly quantify the deviation from the recovered value of our measurement, N-body mocks are used. The latter represent systematic free simulations tuned to each of the tracer samples in term of effective redshift and number density. In addition, these mocks present a more accurate depiction of the RSD. The pipeline is applied in the same way as for the EZmocks, with the same binning choices.

3.3.2.1 Baseline analysis

The OuterRim simulation-based N-body mocks are specifically tailored to correspond to the number density of the DR16 samples for QSOs and ELGs. When considering the galaxy-clustering standard approach (auto-correlation or power spectrum), OuterRim simulations relay comparable clustering to the DR16 sample and their attributed EZmocks (Alam et al., 2020b; Smith et al., 2020), it is thus expected that the clustering between N-body mocks and EZmocks in the case of the void-galaxy cross-correlation function matches.

The multipoles of the correlation function between voids and galaxies for ELGs and QSOs are shown in Fig. 3.12 from the OuterRim N-body simulation, and multipoles for the LRGpCMASS sample are shown in Fig. 3.13 from Nseries simulations. These simulations differ from EZmocks as can be seen in Fig. 3.12 where the clustering of OuterRim and EZmocks is compared and show some discrepancies between the two. Therefore, EZmocks cannot be used to estimate the covariance matrix in the case of the N-body mocks. The estimation of the covariance matrix for the parameter estimation is thus performed with the N_m

available pseudo-random realizations of the respective mocks. In addition, the OuterRim mocks and EZmocks do not have the same underlying cosmology.

Another aspect of the discrepancy between the EZmocks clustering and the N-body mocks in the case of the ELG and QSO, for which the comparison is possible, is that the observational effects or the footprint of the sample may have a higher impact on the density estimation than for the galaxy clustering. The full box of the ELG OuterRim simulation presents a much higher amplitude in the density of the wall than the EZmocks, the amplitude of the quadrupole is more important as well. In the case of the QSO, the amplitude is lower for both monopole and quadrupole, the latter presenting a less defined feature.

The resulting RSD parameter β estimated and associated error distributions are shown in Fig. 3.14 for QSOs and ELGs, and in Fig. 3.13 for the LRG sample. It is to be noted that the error distribution is not exactly Gaussian due to the low amount of mocks used. This leads to higher uncertainties on the covariance estimation of the order of 67% , 23% and 20% for the OuterRim ELG, OuterRim QSO and Nseries LRG, as estimated using the Hartlap prescription in Eq. (3.43). Finally, the results on the β parameter using N-body simulations are given in Table 3.9.

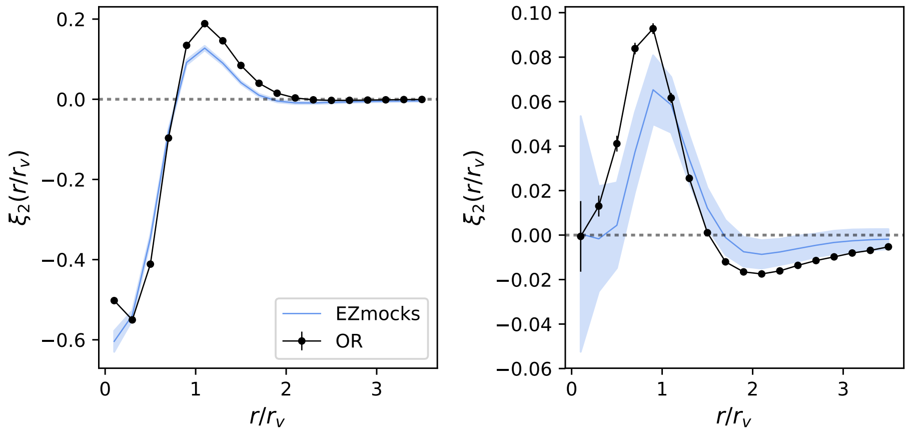
Table 3.9: Results on the analysis applied to the N-body mocks specific to each tracers. Quoted $\beta^{\mathsf{NB}}_{\mathsf{ref}}$ and $\langle\sigma_\beta\rangle$ are the mean over all N_m N-body realizations. The error on the deviation from the expected fiducial value β^{fid} is taken to be the error on the mean of the estimated $\beta^{\mathsf{NB}}_{\mathsf{ref}}$.

	$\langle\beta^{\mathsf{NB}}_{\mathsf{ref}}\rangle$	$\langle\sigma_\beta\rangle$	β^{fid}	$\langle\beta\rangle - \langle\beta^{\mathsf{fid}}\rangle$	
Nseries LRG	0.447	0.063	0.41	0.037 ± 0.007	9%
OuterRim ELG	0.629	0.027	0.686	0.057 ± 0.005	8.3%
OuterRim QSO	0.241	0.037	0.401	0.160 ± 0.004	39.9%

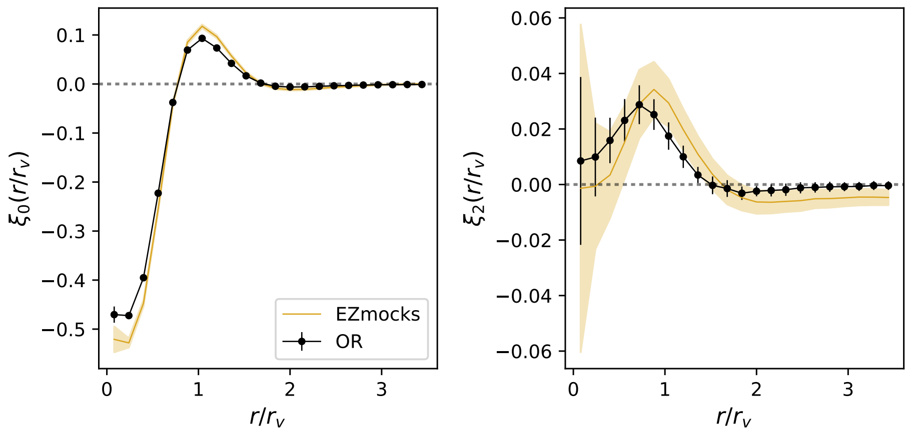
3.3.2.2 Impact of the modelling

At first, N-body mocks are used to characterize the modelling of the redshift space distortions in the form of the comparison of the recovered value compared to the fiducial expected value. In Table 3.9 are given the average β^{NB}_{ref} values extracted from OuterRim and Nseries N-body simulations, which are compared to their expected fiducial value β^{fid}. These values disagree by more than 5σ, where $\sigma = \sigma_\beta/\sqrt{N_m}$ is the error on the mean β value. This implies that the deviation from the fiducial value should be considered as a systematic error related to the RSD modelling. The reported deviations are 8.3% for the OuterRim ELGs, 9% for the Nseries and 39.9% for the OuterRim QSOs.

For the OuterRim ELGs, the 8.3% bias quoted is less than that shown for in the EZmocks of 17.3%. This may mean that some additional systematical effect displayed by the EZmocks may not be taken into account here. In the case of the



(a) ELG



(b) QSO

Figure 3.12: Averaged monopole (left panels) and quadrupole (right panels) of the void-galaxy cross-correlation function estimated in the OuterRim based N-body mocks (black) related to the ELG (top panels) and QSO (bottom panels) DR16 samples. For comparison are drawn multipoles from EZmocks for ELG (blue) and QSO (yellow). Solid lines are average values over all realizations of the considered N-body mocks and error bars or shaded regions are the 1σ errors estimated from the diagonal of the covariance matrix.

QSO, the deviation is much higher than expected, ceiling to 39.9%, superior to that hinted in the EZmocks.

The fiducial β is estimated from the predicted growth rate for a given redshift and cosmology while the linear bias of the galaxies is provided by standard galaxy clustering measurements. The consequence, as for the EZmocks, is that the expected β is not quite an independent quantity as there is no fiducial bias. Leads

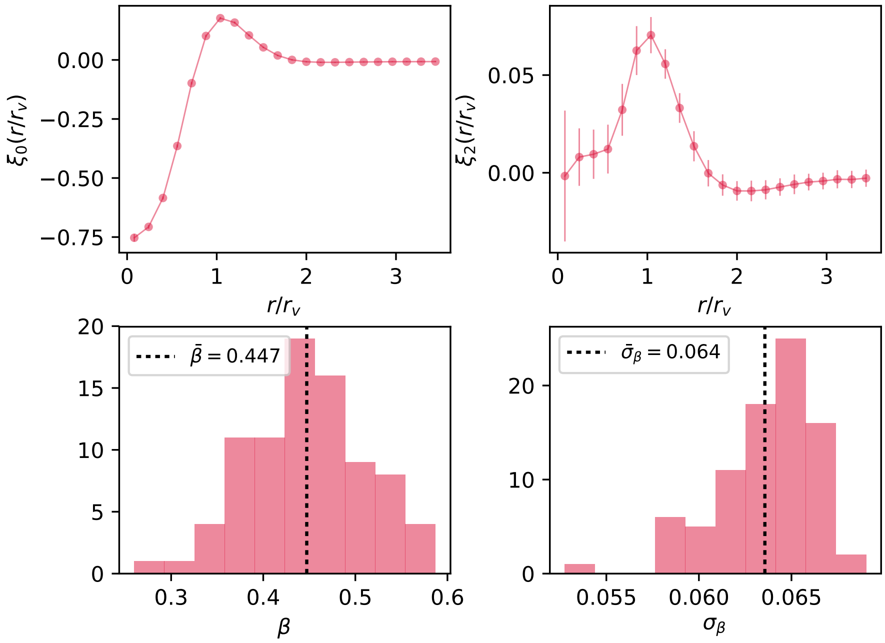


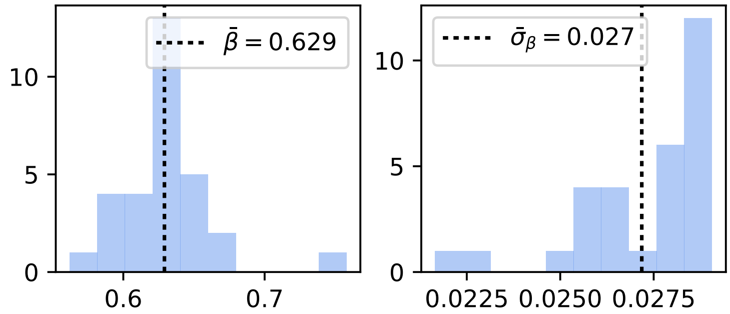
Figure 3.13: Top: Averaged monopole (left) and quadrupole (right) of the 84 NSERIES mocks. Error bars are estimated from the diagonal of the covariance matrix built with the 84 NSERIES. Bottom: Distribution of β (left) and related errors (right) estimated from the NSERIES N-body mocks.

to explain this deviation of the modelling, which represents nearly four times the deviation quoted for the LRG and the ELG, will be investigated later on. But, the modelling only cannot explain this large a deviation, especially when the modelling bias seems to be of the same magnitude when considering the ELG and the LRG samples.

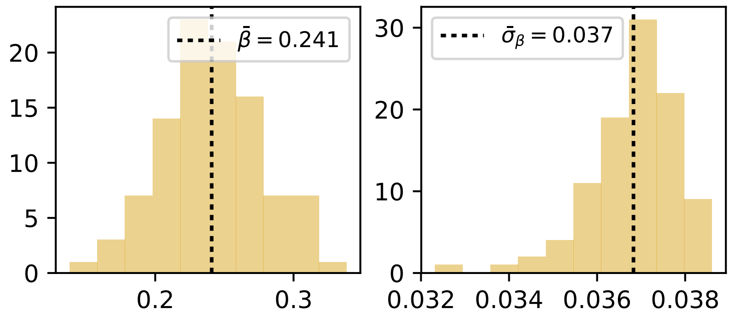
3.3.2.3 Impact of the void centre definition

Through the use of the N-body, the void centre definition is also investigated in order to validate our use of the barycentre. To this end, the resulting clustering and estimated β are compared in each of the N-body simulations to confirm that the barycentre provides the most robust estimation of the void centre in regard to our analysis. The resulting β and comparison with the reference value which we take to be those reported in Table 3.9, are displayed in Table 3.10

The use of a different definition of the barycentre incurs a significant shift in the estimate of β, especially in regard to the study of the OR mocks. While the β^{ref} values are not without systematics, as was shown in above, the change of void definition does not improve those and actually increase the already existing systematic effect, up to 50% in the case of the QSO and 24.2% for the ELGs. The LRG, however, is only affected by an order of 3 to 4% more than when considering



(a) ELG



(b) QSO

Figure 3.14: Distribution of β (left) and its associated error (right) estimated from the OuterRim based N-body mocks for the ELG (top) and QSO (bottom) samples.

the barycentre.

So, the choice of the barycentre as the void centre definition is the most optimal in this analysis of redshift space distortions, similarly confirmed in the case of the LRGs in a companion DR16 void study (see Appendix A, Nadathur et al., 2020b).

3.3.2.4 Impact of the fiducial cosmology

The impact of the fiducial cosmology is also investigated in regard to the estimation of the RSD parameter. To this end, a different cosmology $\Omega_m = 0.31$ is used in both the void finding and the estimation of the void-galaxy cross-correlation function. This test is only applied to the NSeries as they are cut-sky mocks, which makes them easier to treat in terms of the void finding. The ELG and QSO N-body are both box simulations which renders the void finding non-trivial if one were to change the cosmology. The geometry of the object positions would be an elongated box due to the AP effect resulting in a non-optimal placement of the buffer particles, which would have consequences on the void definition.

The deviation from the reference β^{ref} value and that obtained with Ω_m is of

Table 3.10: Estimated β with the circumcentre definition and deviation from the reference barycentre analysis β^{ref} and from the fiducial cosmology. The errors reported for the β are the mean of the 1σ error from the fit. Those reported for the deviations are estimated from the error on the mean ($\sigma/\sqrt(N)$).

	$\langle\beta\rangle$	$\langle\beta\rangle - \langle\beta^{\mathsf{ref}}\rangle$	$\langle\beta\rangle - \langle\beta^{\mathsf{fid}}\rangle$	$(\Delta\beta)^{\mathsf{ref}}$	$(\Delta\beta)^{\mathsf{fid}}$
NSERIES	0.465 ± 0.067	0.018 ± 0.01	0.055 ± 0.007	4%	13.4%
OUTERRIM ELG	0.52 ± 0.026	0.037 ± 0.004	0.166 ± 0.004	18.6%	24.2%
OUTERRIM QSO	0.199 ± 0.022	0.042 ± 0.005	0.20 ± 0.005	17.4%	50.4%

0.003 ± 0.010. Thus, the systematic derived from the use of the fiducial cosmology is inferior to its error. In a conservative manner, the impact of the use of this fiducial is taken to be the 1σ error, which amounts to a 2.2% relative error on the β estimation.

3.3.3 Total systematic budget

The systematic budget with approximate and N-body mocks is quoted in its entirety in Table 3.11 in terms of the relative deviation from the expected fiducial value. Each of the tracer samples has an associated systematic budget associated. The systematic estimated attributed to the change of fiducial cosmology in the NSERIES mocks was applied to each tracer sample. The total systematic budget quoted in the last line of the table corresponds to the addition in quadrature of all systematic contributions. The dominant systematic error is the one due to the RSD modelling.

Table 3.11: Total systematic budget in terms of relative errors on the β parameter obtained from tests with mock catalogues for each of the eBOSS tracer. The total systematic error is the quadratic sum of each contribution.

Type	sys in (σ_β/β) (%)	LRG	ELG	QSO
Methodology	Binning	4.8	2.3	1.3
	FKP weight	1.4	2.3	0.7
	Estimator	2.2	3.3	1.0
Void finder	Fiducial cosmology	2.2	2.2	2.2
Model	RSD modelling	9.0	8.3	39.9
Total (%)		10.8	9.76	40.0

Table 3.12: Standard deviation of the 1,000 estimated β ($\beta(\sigma)$) and σ_β in the mock density ratio of 10 and 100 configurations for all three tracer samples.

Data samples	MDR10	MDR100
–	$\beta(\sigma)/\sigma_\beta(\sigma)$	$\beta(\sigma)/\sigma_\beta(\sigma)$
LRG	0.047/0.002	0.024/0.003
ELG	0.022/0.0013	0.012/0.0007
QSO	0.012/0.0004	0.0007/0.0002

3.4 Application on DR16 data

3.4.1 Impact of the buffer thickness on the RSD parameter

Prior to extracting a final measurement of β with our final QSO, ELG and LRG voids, the impact of the number density of the generated buffer mocks in the void finding procedure is investigated. As the placement of the buffer particles is done randomly, it has an impact on the voids found and their number. To properly estimate the magnitude of this effect, the algorithm is run 1,000 times on the same galaxy sample. It can be seen that the β value is not unique for each galaxy sample but distributed around a mean value.

In this context, voids were extracted from the DR16 galaxy samples with two prescriptions. The first one corresponds to a buffer density of $10\bar{n_g}$, or mock density ratio of 10 (MDR 10), the second prescription was $100\bar{n_g}$ (MDR 100).

The resulting standard deviation of both β and their error σ_β are quoted in Table 3.12 for each data sample and each MDR configuration. As can be seen, the spread of the β distribution is affected by the number density of mocks chosen. It changes by a factor of ~ 2 from the MDR100 prescription to the MDR10 prescription. In the case of the spread of the error on the parameter $\sigma_\beta(\sigma)$, the same behaviour seems to be found for both ELGs and QSOs. Therefore, the prescription of choice was chosen to be $100\bar{n_g}$ to create the void catalogues and to extract the DR16 growth rate constraints. It is to be noted, however, that in the case of the LRG, this test was run on the LRG eBOSS only sample (without the CMASS contribution), but this behaviour seems consistent between the three tracers so it should remain the same for the LRGpCMASS sample.

No systematic error was attributed to the estimation of β from this study, as both the error on the mean β value, at the order of $0.05/\sqrt{1000} \simeq 0.001$, and the standard deviation σ_β are negligible. We consider that using the mean β estimated from the 1,000 estimated void-galaxy correlation functions relays an unbiased constraint in regard to the positions of the void finder buffer particles.

3.4.2 Clustering of the DR16 eBOSS voids

This model and subsequent study of the redshift space distortions around voids was first conducted on a subsample of the DR16 LRGs and QSOs catalogues which consisted in the DR14 of eBOSS (Hawken et al., 2020). This first foray in the void-galaxy clustering of the eBOSS LRGs and QSOs samples was shown to be highly biased by the number density of the galaxies as well as the footprint. The resulting void number densities provided very large error bars to the estimation of β in the case of the LRGs. In the case of the QSO, however, the lack of a quadrupole signal, consistent with 0 within 1σ evidenced a total lack of sensitivity of RSDs around QSO voids. While the LRGs measurement was not quite a cause of concern due to the wide range probed by the final estimated parameter, this near absence of clustering in the QSOs was puzzling.

Indeed, the lack of an RSD signal in the void-galaxy cross-correlation function was very much consistent with the signal given by random voids which carry nary a cosmological information. A careful comparison between QSO voids and voids extracted in a random subsample with the same number density than that of the data (Cousinou et al., 2019; Hawken et al., 2020) showed to be unable to discriminate in a clear-cut manner these two species of voids. This study suggested that the void finding procedure cannot quite prevent from finding 'wrong' voids and especially that these voids and the cosmological information they carry can be deeply affected by a low sampling of the galaxies in a considered comoving volume.

The work presented in this chapter consists in the application of the analysis to the complete ELG, LRG and QSO samples of the final eBOSS data release. The void-galaxy two-point cross-correlation function was estimated for each of the samples, using the voids catalogues presented in Section 3.1.4 using the Davis-Peebles estimator, as well as for all realizations of their respective approximate mocks.

Figure 3.15 displays the resulting clustering statistics for each of the tracer in terms of their monopole ξ_0 and quadrupole ξ_2 as a function of the separation between the void centre rescaled by the radius of the voids. The points correspond to the data while the shaded part shows the dispersion of the measured multipoles in the approximate mocks, the full line corresponds to the mean of the mocks. The monopole corresponds to the observed void density profile (or stacked density profile) of the considered sample which gives this very distinctive behaviour that is presented to be universal, as mentioned in Section 2.3.1.1. In this case, all of the considered voids show a non-negligible compensation wall, the most important being that of the LRGpCMASS sample. The quadrupole in its sole existence consists in itself as a proof of the anisotropy of the correlation function due to the velocities of the galaxies. The positive part of the quadrupole corresponds to the outflow of the galaxies toward the over-density surrounding the void.

In the case of all samples, there is quite a good agreement between the clustering of the data and that of the mocks. For QSOs, the signal of the quadrupole, even

Table 3.13: Final results on the RSD parameter β estimated with the separate eBOSS DR16 samples. β and its statistical error σ_{stat} are quoted as the mean of the 1000 estimation of β and subsequent error distribution for each eBOSS tracer data, The systematic error σ_{syst} corresponds to the total systematic error quoted in Table 3.11 applied to our estimate. The total error σ_{tot} corresponds to the aforementioned error added in quadrature.

Data samples	$\langle\beta\rangle$	σ_{stat}	σ_{syst}	σ_{tot}
LRG	0.415	0.075	0.045	0.087
ELG	0.665	0.107	0.065	0.125
QSO	0.313	0.049	0.125	0.134

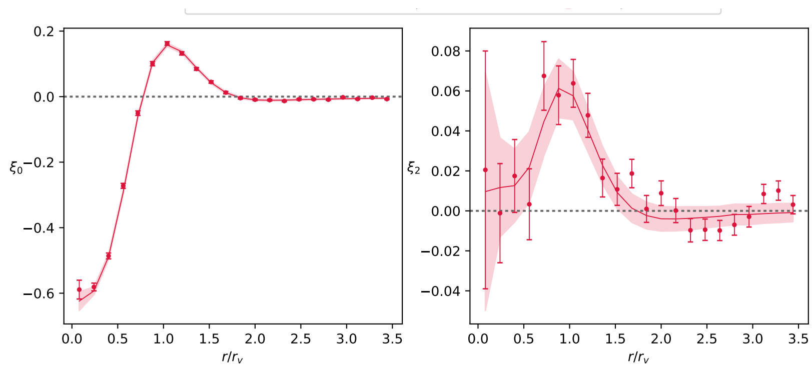
though toned down compared to that of the LRGs and the ELGs, is clearly different from 0, which represent a real improvement compared to the DR14 study. The accordance between the clustering of the data and that of the mocks is highly important.

The pipeline, as validated on mocks, was applied to all DR16 samples. For each sample, in order to take into account the random aspect of the buffer particles in the void finding, the void finder was run 1000 times on the same galaxy catalogue, producing 1000 slightly different void catalogues. The void-galaxy cross-correlation function is then estimated for each of the thousand void catalogues as well as for the parameter estimation. The best-fit of the cross-correlation function on one realization among 1000 is shown in Fig. 3.16 for each data sample, inkling to an overall good agreement of the fitted quantities. The resulting distributions of the fitted β and estimated errors are displayed in Fig. 3.17. The distributions of β and their error are effectively narrower than those provided by the EZmocks. The RMS of the β distribution is clearly negligible in regard to the estimated error on β, which allows us to conclude that this behaviour does not cause any systematical shift in the data as it is sufficiently small to be accounted for in the mocks spread.

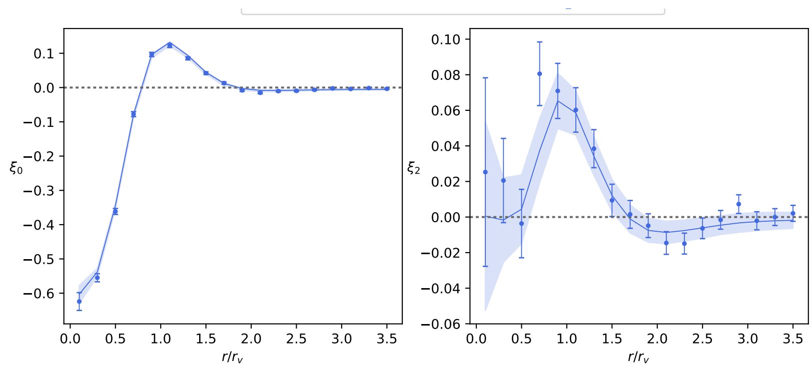
The final results of the analysis for the DR16 samples are reported in Table 3.13, where the quoted β are the mean of the thousand β estimated for each tracer and subsequent mean statistical error estimated in the fit. The systematical shifts investigated with both the EZmocks mocks and N-body mocks: binning, weights, estimator, modelling, and fiducial cosmology are added in quadrature to amount to the total presented in σ_{syst}. The final error estimate σ_{tot} is the quadratic sum of statistical and systematical errors. The modelling systematic shows to be the most dominant, especially in the case of the QSO sample.

3.4.3 Constraints on the growth rate of structure

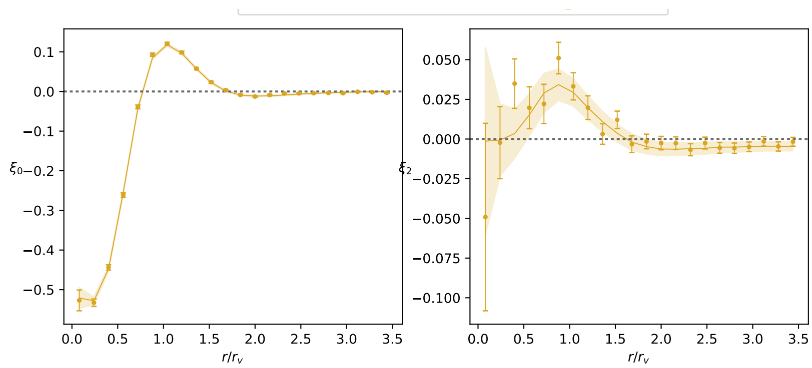
The model used for the redshift-space distortions around voids and therefore, to probe the growth rate of structure, relays a biased estimate β. This parameter is a ratio of the growth rate of structure and the bias. The relation between the



(a) LRGpCMASS

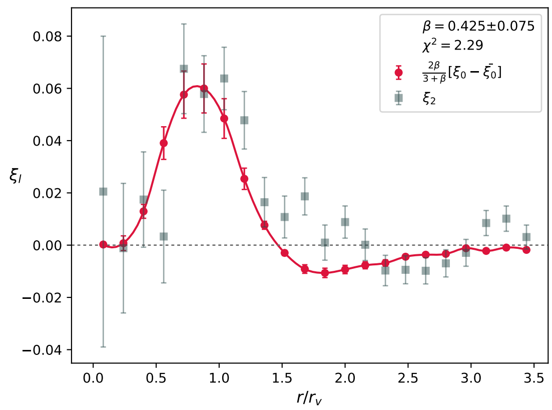


(b) ELG

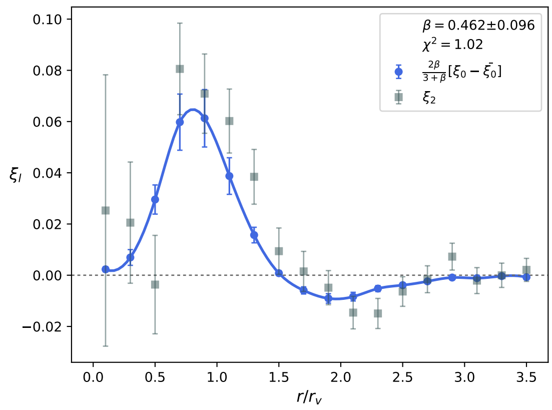


(c) QSO

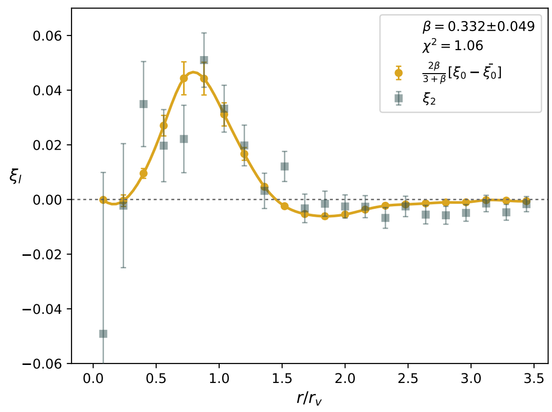
Figure 3.15: Multipoles of the DR16 void-galaxy cross-correlation functions of data compared to the mock catalogues. Left panels show the monopole and right panels show the quadrupole, both as a function of the separation distance r normalized to the effective void radius r_v. The LRG, ELG and QSO DR16 samples are displayed in the top (a), middle (b) and bottom (c) panels, respectively, for the data (circle symbol) and the mean of 1000 EZmocks realizations (solid line). The shaded region shows the standard deviation of the 1000 mock realizations, and error bars on data are the square-root of the diagonal elements of the covariance matrix.



(a) DR16 LRGpCMASS

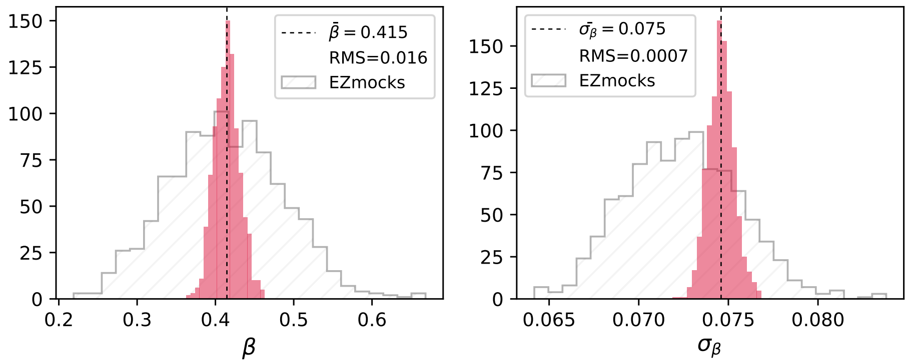


(b) DR16 ELG

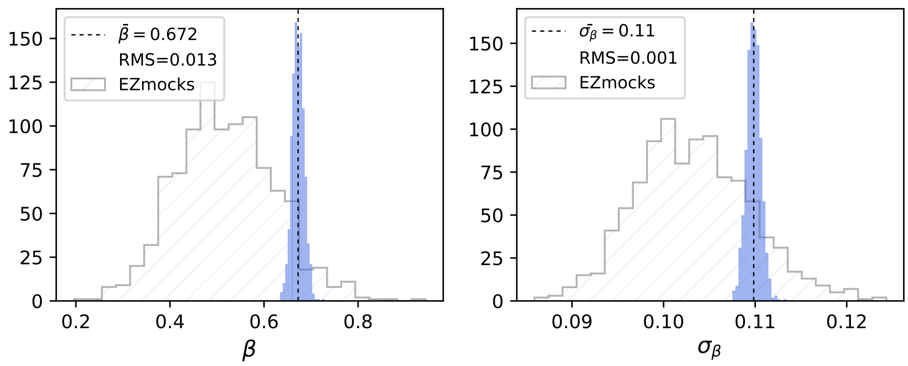


(c) DR16 QSO

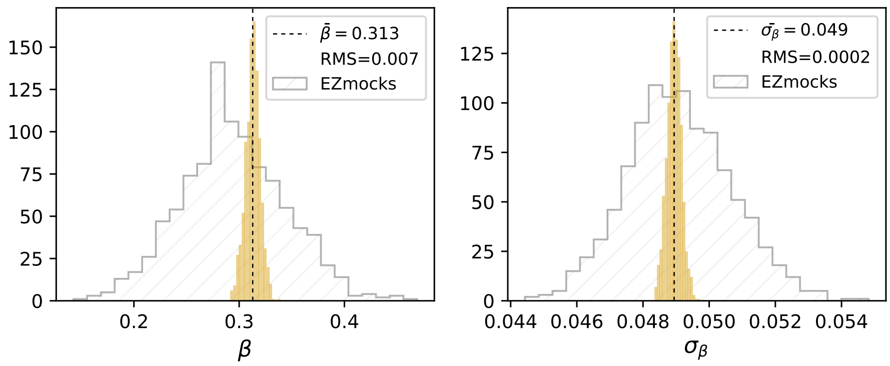
Figure 3.16: Quadrupole (ξ_2) and the best-fit of the $2\beta/(3+\beta)(\xi_0 - \bar{\xi}_0)$ from one DR16 data catalogue of the LRGpCMASS, ELG and QSO sample displayed in the top (a), middle (b) and bottom (c) panels, respectively. Error bars are the diagonal of the covariance matrix from the 1000 EZmocks realizations.



(a) DR16 LRGpCMASS



(b) DR16 ELG



(c) DR16 QSO

Figure 3.17: Best-fit parameters for the DR16 data catalogues. Left panels display the distribution of the distortion parameter β and right panels display the distribution of the errors of β. The LRG, ELG and QSO data samples are displayed in the top (a), middle (b) and bottom (c) panels, respectively.

Table 3.14: Final results on the growth rate estimate from the eBOSS DR16 void datasets. Mean values and errors on β are taken from Table 3.13. The presented errors include the systematic component. The reported value of $b_1\sigma_8$ are taken from clustering analysis in the DR16 companion papers, for the LRG sample (Bautista et al., 2020; Gil-Marín et al., 2020), the ELG sample (De Mattia et al., 2020; Tamone et al., 2020) and the QSO sample (Hou et al., 2020; Neveux et al., 2020). The growth rate constraint results from applying Eq. 3.48 to these values. The total error quoted for $f\sigma_8$ includes the galaxy bias error contribution.

Data samples	z_{eff}	β	$b_1\sigma_8$	$f\sigma_8$
LRG+CMASS	0.740	0.415 ± 0.087	1.20 ± 0.05	0.50 ± 0.11
ELG	0.847	0.665 ± 0.125	0.78 ± 0.05	0.52 ± 0.10
QSO	1.478	0.313 ± 0.134	0.96 ± 0.04	0.30 ± 0.13

void-galaxy density field is assumed to be linear with the void-matter density field, in which the bias corresponds to that of the galaxies. However, the bias is itself an unknown quantity which cannot be estimated independently of f in the modelling adopted and depends both on the tracer used to probe the density field and its effective redshift.

3.4.3.1 Estimation of the growth rate of structure

In order to translate the measured quantity in that relevant for cosmological analysis, one must remove the dependence in the bias. To this end, the measurements of the linear bias estimated with galaxy clustering standard techniques on the same samples are used to recover the growth rate at similar redshift. However, standard galaxy clustering techniques, that is the galaxy autocorrelation and power spectrum, measure a degenerated bias parameter $b\sigma_8$, where σ_8 is the standard deviation of the mass fluctuations in a sphere of radius $8h^{-1}$Mpc . The growth rate is then obtained through the following parametrization:

$$f\sigma_8 = \beta b_1 \sigma_8 \tag{3.48}$$

where $f\sigma_8$, the growth rate is also considered in terms of σ_8, allowing to constrain the quantity independently from the linear bias b (Song & Percival, 2009).

In the case of the void-galaxy correlation function estimated in redshift space, however, the bias and the growth rate cannot be yet fitted independently.

To this end, the β estimated for each tracer samples are combined with the estimated $b_1\sigma_8$ estimated by the galaxy clustering analysis in Fourier and configuration space for each DR16 sample: ELGs (De Mattia et al., 2020; Tamone et al., 2020), LRGs (Bautista et al., 2020; Gil-Marín et al., 2020) and QSOs (Hou et al., 2020; Neveux et al., 2020). The estimated β and $b_1\sigma_8$ are combined to obtain the estimations of the $f\sigma_8$ in Table 3.14 in which the errors corresponds to the total error σ_{tot} quoted in Table 3.13 and are combined with the $b_1\sigma_8$ recorded in Table 3.14 as well. The combination of the errors is done while considering that both quantities (bias and RSD parameter) are independent. A test was done by estimating the correlation between the 951 β measured with the EZMOCKS ELGs and the associated 951 $b\sigma_8$ estimated in the standard galaxy clustering framework. It was found that these two quantities are very lowly anti-correlated $r = -0.16$, which leads to an under-estimation of the $f\sigma_8$ error of the order of $\sim 4\%$. While this has not yet been tested on the other DR16 tracers such as the LRGs or the QSOs, we assume that this behaviour is similar.

The resulting $f\sigma_8$ and its errors are shown in Table 3.14, along with the estimated β and effective redshift respective to each sample. The linear bias recovered from the DR16 companion papers are also quoted.

3.4.3.2 Comparison with the literature

The growth rate of structure in the $f\sigma_8$ or β parametrisation has been probed in a wide range of redshift and tracer samples. In this section, the measurement extracted from the DR16 void-galaxy correlation functions are compared to those found in the literature.

Figure 3.18 displays the growth rate estimated in the various samples provided by the SDSS survey across its four stages. Empty markers denote the $f\sigma_8$ estimate extracted from void analyses while the associated full markers present the measurements achieved by the standard galaxy clustering analyses (auto-correlation and power spectrum). The red empty stars represent the measurement of $f\sigma_8$ at redshift $z = 0.74$, $z = 0.845$ and $z = 1.48$ for the LRG, ELG and QSO sample respectively measured in this work, to be compared to the associated galaxy clustering (GC) consensus measurements of the eBOSS DR16 companion papers. It is to be noted that our measurement of the $f\sigma_8$ is shifted in regard to the GC LRG estimate, which is due to the cut applied to our void sample. It causes the effective redshift of our sample to be higher. Our measurement is also compared to a measurement of the growth rate of structure in the DR16 eBOSS LRG sample achieved with a fundamentally different model of the void-galaxy two-point correlation function (Nadathur et al., 2020b). Overall, the agreement between our measurements and those of the eBOSS DR16 companion papers are in agreement at the level of 1σ for the LRG and QSO and at a 2σ level in the case of the ELGs. The estimated growth rates are also displayed for the BOSS DR12 sample, in three redshift bins for the standard GC consensus analysis (Alam et al., 2017) and for voids analysis. While Nadathur et al. (2019a) relays a direct estimate of $f\sigma_8$, Achitouv (2019) and Hamaus et al. (2017, 2020) apply the same or a similar model to that conducted in this chapter. As a result, their β estimate are recovered using the fiducial bias $b = 1.85$ estimated from the DR12 BOSS sample and σ_8 is computed with the Planck 2018 Λ-CDM cosmology (Aghanim et al., 2020a). This yields $\sigma_8(z = 0.32) = 0.684$ and $\sigma_8(z = 0.54) = 0.612$ for Hamaus et al. (2017),

$f\sigma_8(z = 0.32) = 0.418 \pm 0.76$ and $f\sigma_8(z = 0.54) = 0.407 \pm 0.057$ for Achitouv (2019) and $f\sigma_8(z = 0.51) = 0.620 \pm 0.105$ for Hamaus et al. (2020). Finally is plotted the $f\sigma_8$ estimated from the main galaxy sample (MGS Howlett et al., 2015) of the DR7 SDSS (Abazajian et al., 2009). Overall, the ensemble of the SDSS growth rate measurements, voids and galaxies are all consistent within $1 - 2\sigma$.

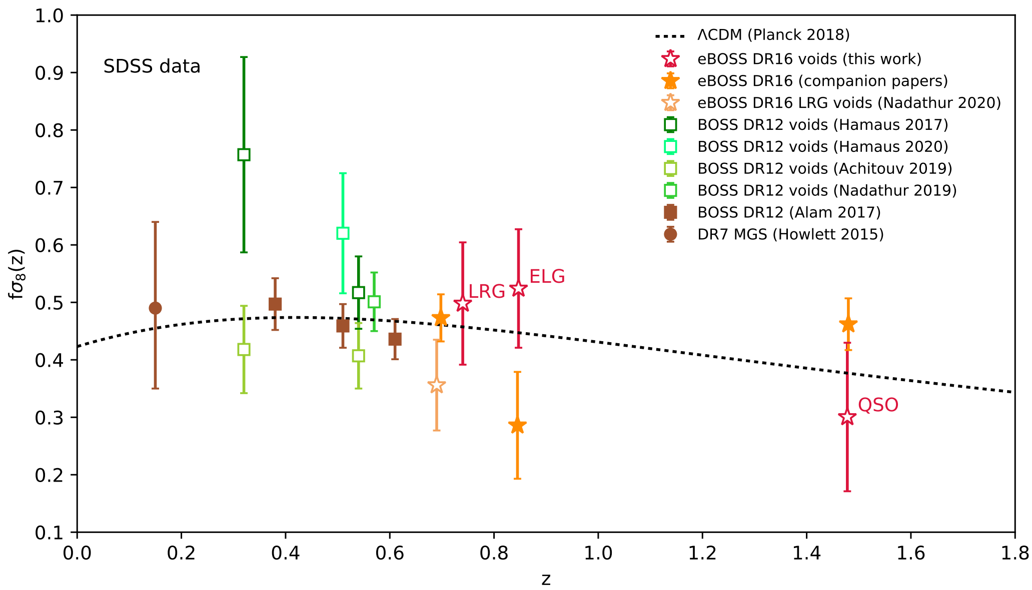


Figure 3.18: Growth rate of structure $f\sigma_8$ in terms of redshift measured in various samples of the SDSS collaboration. Empty markers display the void constraints and filled marker the galaxy clustering constraints. Red empty stars represent the estimation of the growth rate with voids for the LRGs,ELGs and QSOs in the DR16 sample presented in this chapter (Aubert et al., 2020). The orange star filled markers represent the result of the consensus measurements for the LRGs (Bautista et al., 2020; Gil-Marín et al., 2020), ELGs (De Mattia et al., 2020; Tamone et al., 2020) and QSOs (Hou et al., 2020; Neveux et al., 2020). Orange empty star represent an additional measurement of the growth rate with voids with a different analysis (Nadathur et al., 2020b). Empty square markers relate to the analysis of the BOSS DR12 (Achitouv, 2019; Hamaus et al., 2017, 2020; Nadathur et al., 2019a), filled square markers being the standard consensus DR12 measurement (Alam et al., 2017). Brown circle displays the measurements of the growth rate in the Main Galaxy Sample (MGS)(Howlett et al., 2015) in the SDSS DR7 (Abazajian et al., 2009). The dotted line is estimated Planck 2018 ,$\Omega_m = 0.31$,$\sigma_8 = 0.81$, prediction for the growth rate of structure.

Figure 3.19 compiles all the measurements of $f\sigma_8$ outside the SDSS collaboration. The upper panel displays the entirety of the measurements of $f\sigma_8$ with both galaxy clustering and void-galaxy clustering. The previous growth rate estimation on the

SDSS data samples are displayed with the same color codes as in Figure 3.18, with the exception of the DR12 BOSS CMASS sample. The comparison is thus extended to 6dFS (Beutler et al., 2012), GAMA (Blake et al., 2013), WiggleZ (Blake et al., 2011a), VIPERS (Pezzotta et al., 2017) and FastSound (Okumura et al., 2016). Along with the galaxy clustering estimation are also added the subsequent void estimation in the 6dFGS sample (Achitouv et al., 2017) and in the VIPERS survey (Hawken et al., 2017).

It is to be noted that the measurements displayed seem consistent with the Λ-CDM model Planck 2018 prediction. However, at higher redshift, $z \geq 0.6$, the spread of the estimated $f\sigma_8$ is larger.

In the lower panel of Figure 3.19 are shown the void only constraints, included in the upper panel as well. The numerous estimation of the $f\sigma_8$ values in the BOSS DR12 sample at high redshift display a spread in the variety of the model used to constrain the growth rate of structure as well as the void finding algorithm used. However, the estimated $f\sigma_8$ remains consistent between the various modelling used. It is to be noted that the recovered estimate of the growth rate relies on the fiducial bias value and predicted σ_8 at these redshifts, in the absence of a linear bias estimation. In the case of the LRG eBOSS DR16, we find good agreement between the $f\sigma_8$ recovered in this work and that extracted with a different RSD model (Nadathur et al., 2020b).

The disparity in the available growth rate measurements does not allow yet to discriminate against a Λ-CDM scenario. The growth history presented here suffers from the comparison between different analyses and sample considered as well as the lack of measurement between 0.9 and 1.4. Future surveys such as Euclid (Amendola et al., 2018; Laureijs et al., 2011) and DESI (DESI Collaboration et al., 2016a,b) are expected to provide the data enabling to fill the missing information at these redshifts.

3.4.4 Discussion

The systematic study which has been a prime focus of this work evidences a strong impact of both RSD modelling and methodology on our RSD parameter error. This systematic error can be envisioned along three major axes that will be discussed below.

3.4.4.1 On the methodology

A major aspect of the systematic studies concerned the methodology choices such as the binning choice, estimator and the FKP weights with the EZmocks respective to each tracer. These systematic effects mostly impact the estimation of the correlation and underlying pair-counts that are used to estimate it. These will, as such, affect the amplitude of the monopole and quadrupole.

A salient feature of our methodology lays in the parameter estimation which

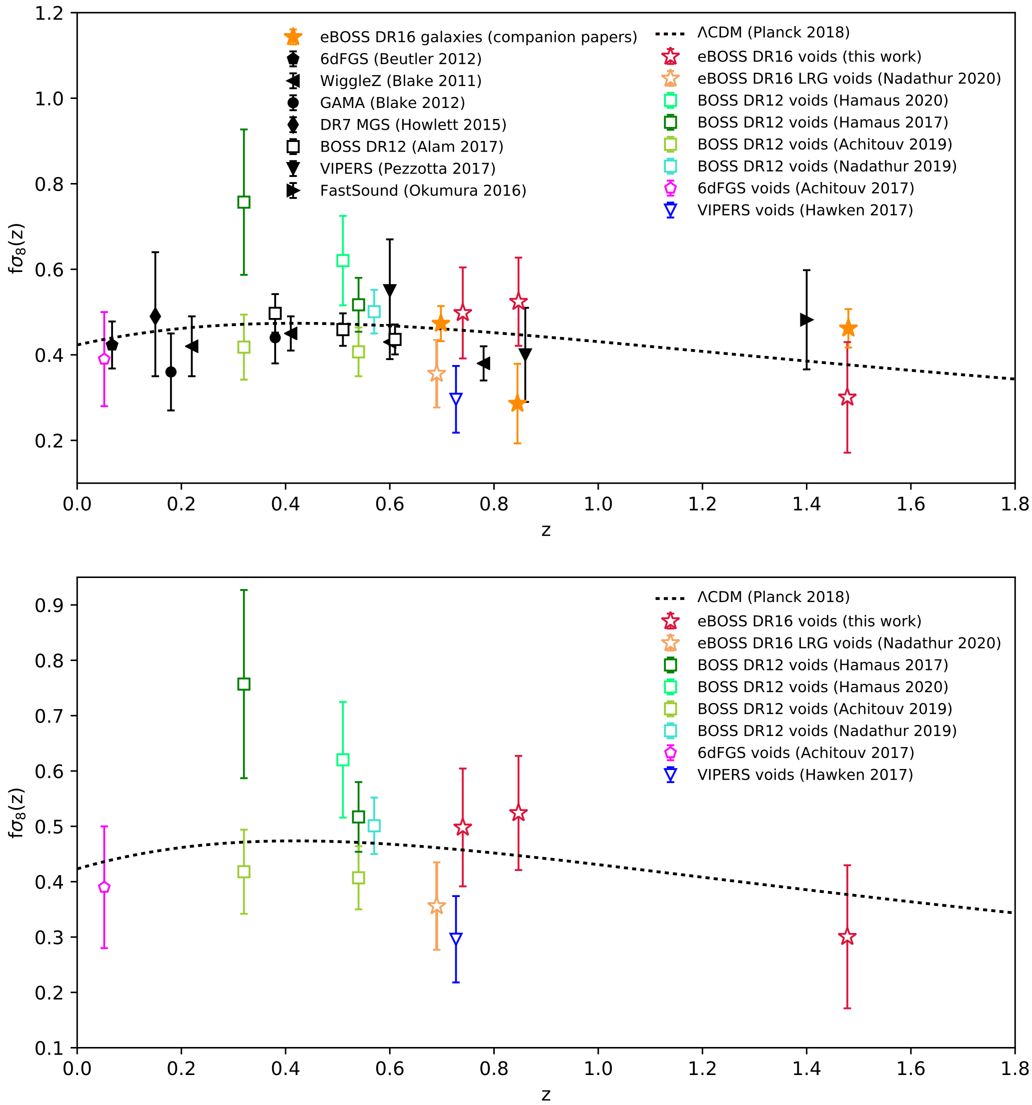


Figure 3.19: Upper panel: Comparison of the estimated growth rate of structure $f\sigma_8$ recorded in varied galaxy samples and subsequent voids. Filled black markers represent the various surveys estimation,6dF (Beutler et al., 2012), GAMA (Blake et al., 2013), WiggleZ (Blake et al., 2011a), VIPERS (Pezzotta et al., 2017) and FastSound (Okumura et al., 2016), in addition to the SDSS void (empty marker) and galaxy (black tracer) constraints presented in Fig. 3.18 and reference therein. The estimated $f\sigma_8$ with voids in 6dF (Achitouv et al., 2017, magenta hexagone) and VIPERS (Hawken et al., 2016, blue triangle) are also displayed. The dotted line is estimated Planck 2018 prediction for the growth rate of structure. Lower panel: Comparison of the estimated growth rate of structure $f\sigma_8$ using void-galaxy clustering only, the same values as in the upper panel are reported.

uses the residuals defined in Eq. 3.38 as an alternative to the ratio which directly removes the contribution of the real-space correlation function in our modelling. As a consequence, our parameter estimation relies on the direct comparison of the amplitudes of the monopole and quadrupole of the correlation function. Any modification in the latter's estimation has a non-negligible impact on the estimated RSD parameter. The methodology choices investigated induced statistical fluctuations in the estimation of the void-galaxy two-point correlation function, which resulted in a varied estimation of the β parameter. These may be mitigated with an additional parametrization of the $\xi_{vg}(r)$, the true underlying void-galaxy correlation function - at the risk of absorbing cosmological information, discarding the fitting procedure of the residuals. This kind of systematic may thus be mitigated by the use of both a theoretical modelling of the true underlying void-matter density field, which has yet to be investigated or a higher statistic. While an empirical function is available to model $\delta_{vg}(r)$, it is defined by a large number of free parameters which may absorb the cosmological signal.

These fluctuations, however, mostly concern the EZmocks ELG and LRG samples leading to a total systematic effect of 5.5% for the LRG and 4.6% for the ELGs and lowers to 1.8% in the case of the QSO sample. Hence, it seems that the sensitivity to the methodology choices may also be correlated to both uniformity and number density of the sample. The QSO sample provides the largest void and galaxy statistics among the three tracers and seems to be less affected by the methodology choices. The LRG quotes the largest deviation (5.5%) while having a superior statistics to the ELGs. The sample results in the addition of the eBOSS LRG targeted and the high redshift tail of the DR12 BOSS CMASS LRGs, this may lead to some unaccounted for uneven density patches within the survey footprint thus affecting our measurement. The systematic effect quoted by the ELGs, in this regard, may be attributed to the statistic of the sample.

3.4.4.2 On the RSD modelling

The dominating source of systematic on our parameter estimation appears from the tests of the RSD modelling on N-body simulations. While spectacular in the case of the QSO (39.9% !), the systematic shown for the LRGs and ELGs remains below the 10% level. This necessitates considering the model used for the estimation of β.

Parametrization of the fit: As mentioned earlier, the RSD model employed here allows the recovery of the RSD parameter with an amplitude ratio that we fit with the residuals defined in Eq. 3.38. This parametrization supposedly remove any dependence on the true underlying void-matter density field $\delta(r)$, assuming a linear bias relation of the form $\xi(r) \sim b\delta(r)$. This model has been used in two independent studies of the BOSS DR12 sample (Achitouv, 2019; Hamaus et al., 2017), yielding similar results consistent with a Λ-CDM scenario, as seen in the previous section. Fitting the residuals in order to estimate β might not provide an unbiased estimate as the amplitudes of the monopole and quadrupole can be derailed by statistical fluctuations.

As such, the β estimation should be reconsidered to include a parametrization of the two-point correlation function, such as the void density profile mentioned in Section 2.3.1.1. This empirical formula would allow the comparison of the measured data vector with a theoretical data vector and to have a more robust estimation of β. Such a parametrization is currently under investigation. A second way to recover a model data vector to compare to would be to recover the true correlation function. The latter can be estimated with the transverse to the line-of-sight correlation function $w_{\perp}(r)$ and the use of the inverse Abel transform as shown in (Pisani et al., 2014), as was done in Hamaus et al. (2020).

Linear bias assumption: While doubt can be cast on the redshift space distortions model adopted, another consideration has to be taken into account: the linear bias assumption. A key assumption of the model adopted is the supposed linear relation between the void-galaxy correlation function and the underlying void-matter field which is encoded in a linear **scale independent** bias. It follows that the systematic shifts attributed to the model depend on this assumption as well. However, the true bias relation in regard to the $\delta_{vg}(r)$ remains unknown.

As a result, the combined effect of galaxy sampling and void finding, which have not yet been quantified as well as the linear bias assumption may be at cause here. Such biasing could be linear and scale-independent but have an altogether different value when concerning voids. The resulting systematic shift would thus be wrongly quantified. If we consider our measurements as plotted in Figure 3.18, the additional systematics accounted in the errors of our DR16 points drive our measurements to be consistent with the reported value of the standard galaxy clustering technique (especially in the case of the QSOs).

Incomplete model: Last but not least, the model used for estimating the RSD parameter may be incomplete. Recently, Hamaus et al. (2020) considered a modification of our baseline model with the inclusion of two marginalization parameters $\mathcal{M}$ and $\mathcal{Q}$. The former accounts for the eventual contribution of Poisson voids which affects the amplitude of the correlation function and the latter accounts for void selection effects. While still within the bound of the presented model, systematics may be controlled with such parameters.

The very validity of this model has been decried in regard to its building assumptions, detailed in 3.2.2.1, that are, in summary, void-galaxy pair-count conservation between redshift space and real-space, radial distributed velocity around the void and lack of dynamic at the centre of the void itself (Chuang et al., 2017; Nadathur & Percival, 2019). At the root lies the problem of finding voids within redshift-space tracers thereby imparting a dynamical compound to the void centre.

A new model advocated by Nadathur & Percival (2019) is an expansion of our model without discarding the products $\xi(r)\delta(r)$ and $\xi(r)\Delta(r)$ in Eq. 3.21, on the basis that the terms are not negligible. The result is the appearance of a noticeable feature in the quadrupole of the correlation function in the form of a dent. Such a feature has not been noticed in our measured quadrupole. While this model, under the correct circumstances provides quite tight constraints, it does rely on calibration from N-body simulations and the estimation of the real-space correlation function through a procedure of reconstruction of the density field. This results in the analysis of a quite different subset of voids. Our RSD model and subsequent optimizations remain the most appropriate in the context of the analysis of redshift-space voids, which was similarly confirmed in Nadathur et al. (2020b).

3.4.4.3 on the QSO sample

This QSO sample represents the farthest clustering sample to this day allowing to probe deeper in time for the growth rate and ensuing implications on the properties of dark energy. However, the RSD analysis applied to both EZmocks and N-body simulation, emphasize a large systematic shift in the measurement of the RSD parameter, which cannot be solely attributed to the modelling of the RSD adopted. Indeed, the ELG and LRG samples display an 8.3% and a 9% deviation, while the QSO sample evidences a 39.9% deviation. It follows that the RSD model may account for 10% maximum, leaving the remaining 30% deviation measured to be explained.

The QSO sample has a fundamental peculiarity, compared to the galaxy samples, that could provide some hints on the reason for such a large systematic shift. Indeed, its redshift range is quite wide compared to the other two samples, spanning from 0.8 to 2.2. In combination with the number of objects available in the catalogue, the QSO sample is quite sparse.

The systematic shift is present in both type of mocks: N-body and approximation, although the OuterRim simulation results seem to be quite extreme. The major differences between both relate to the footprint covered and their nature as cubic simulations at a unique redshift $z = 1.433$ for the OuterRim and an evolutive redshift, data like, in the case of the EZmocks with a similar effective redshift of ~ 1.45. Both mocks present a similar number density and conserve the same radius properties.

Considering the EZmocks, tailored to be comparable to the data, the number density varies over the redshift range. While a mean density of the sample is considered to infer the properties of the voids, they can still exhibit a redshift dependent behaviour in conjunction with their varying number density. It seems that this redshift evolution mitigates the systematic shift in regard to the fiducial expectation, lowering it to 27%. instead of 39.9% in the case of the N-body.

The subsequent systematic shift observed in the evaluation of the β parameter could thus be interpreted as the consequence of a too low density of tracers at the redshift considered, leading to a loss of the RSD signal.A similar loss in signal was evidenced in the QSO sample of the DR14 eBOSS sample, leading to an estimation of the RSD parameter consistent with nullity (Hawken et al., 2020).

Although the DR16 QSO sample is much more populated and displays a significant quadrupole, it is still possible that the clustering amplitude of the QSO is too low to identify RSD sensitive voids, in opposition to Poisson voids which tend to have a null quadrupole (Cousinou et al., 2019). A seeming confirmation comes from the fact that the BAO reconstruction algorithm could not be applied to the QSO sample due to its sparsity (Ross et al., 2020). While the REVOLVER algorithm does not apply any smoothing scale when performing the Voronoi tessellation, it is still dependent on the tracer density.

A proper answer concerning this systematic shift is still under-investigation, but if it is indeed due to the sparsity, the RSD modelling should be re-considered to include a mitigation parameter in order to recover the underlying growth rate signal.

4 Probing cosmology with geometrical distortions of voids

4.1 The Alcock-Paczynski test

Voids are sensitive probes of cosmology through distortions processes. Indeed, as seen in Chapter 2, their definition and 'observation' is dependent on the density field evolving around them. As such, any distortions of the surrounding density field is likely to permeate the definition of voids and subsequently, their shape. A first kind of distortion was described in Chapter 3, due to the dynamics of the galaxies in the vicinity of the voids. A second one, of geometrical nature, depends directly on the present-day value of cosmological parameters: the Alcock-Paczynski effect. Void shapes, in this respect, have long been properties of interest when aiming to constrain cosmology through the use of geometric distortions.

4.1.1 Definition and theoretical approach

The Alcock-Paczynski (AP) effect is a purely geometric effect resulting from the expansion of the Universe (Alcock & Paczyński, 1979). Let us consider an object with known symmetry properties, a sphere representing an ideal case. The symmetry axis can be considered in two directions : the line-of-sight (LOS) and the direction transverse to the LOS. In a general definition, the Alcock-Paczynski effect relies on the breaking of the symmetry axis (isotropy) due to the use of a wrong cosmology when going from redshift space to comoving space.

The symmetry axis of the known object can be formalised in terms of Δz, the extent of the spherical object along the line-of-sight and $\Delta\theta$ its angular extent, transverse to the line-of-sight. It follows that Δz is quite dependent on the expansion compared to $\Delta\theta$. However, our knowledge of the shape of the object takes its root in the comoving reference frame, which yields the following relations :

$$\Delta r_\perp = D_A(z)\Delta\theta, \tag{4.1}$$

$$\Delta r_\parallel = \frac{c\Delta z}{H(z)}, \tag{4.2}$$

where $\Delta r_\perp$ and $\Delta r_\parallel$ are the extent of the object in the transverse direction and in the direction of the LOS, respectively. $D_A(z)$ is the transverse comoving distance as defined in section 1.1.4, which, in the event of a flat Universe, reduces to the radial comoving distance $D_c(z)$. $H(z)$ is the Hubble parameter or expansion rate. The redshift z for which this relation is considered is taken to be that of the object.

Assuming that the symmetry relation of the object is known in terms of the axis ratio imposes:

$$\frac{\Delta r_{\parallel}}{\Delta r_{\perp}} = \eta, \tag{4.3}$$

where η is the true axis-ratio of the object. In the case of a spherical object in real-space, this relation simply becomes :

$$\frac{\Delta r_{\parallel}}{\Delta r_{\perp}} = 1. \tag{4.4}$$

This last relation allows us to recover the general definition derived by (Alcock & Paczyński, 1979) of what is called the Alcock-Paczynski parameter F_{AP}, in the case of a spherical object:

$$\frac{\Delta z}{z\Delta\theta} = \frac{H(z)D_A(z)}{cz}. \tag{4.5}$$

In practice, as the shape of the object is known in comoving coordinates, its measurement requires to convert the observational positions $(\mathrm{RA}, \mathrm{Dec}, z)$ into their comoving counterparts. To this end, the input cosmology used in the transformation of the redshift into the comoving distance is called the **fiducial cosmology**. The difference between the true underlying cosmology of our Universe and the fiducial cosmology used to convert distances breaks the symmetry of the object, leading to a value η different from 1 in Eq. 4.3. This deviation from unity is used to verify that the fiducial cosmology does not match the true cosmology and therefore constitutes a reliable test for finding the true cosmology: this is called the AP test.

One can thus probe the cosmological parameters through the use of a fiducial cosmology when estimating $r\|$ and $r\perp$. The distortion effect spurring from the fiducial quantities can be accounted for with the a-dimensional quantities:

$$\alpha_{\perp} = \frac{\Delta r_{\perp}^{true}}{\Delta r_{\perp}^{fid}} = \frac{D_A(z)^{true}}{D_A(z)^{fid}} \tag{4.6}$$

$$\alpha_{\parallel} = \frac{\Delta r_{\parallel}^{true}}{\Delta r_{\parallel}^{fid}} = \frac{H(z)^{fid}}{H(z)^{true}}, \tag{4.7}$$

where the contributions of Δz and $\Delta\theta$ disappear. The $\alpha_{\perp}$ and $\alpha_{\parallel}$ thus indicate the distortions that occur in the transverse and parallel direction to the LOS.

When measuring the shape of a known symmetrical object in terms of its axis-ratio in a fiducial comoving coordinate system, $\frac{\Delta r_{\parallel}}{\Delta r_{\perp}}$, we actually directly quantify a *departure from symmetry* caused by the use of a fiducial cosmology. We call this quantity the ellipticity :

$$\epsilon = \frac{H(z)^{true} D_A(z)^{true}}{H(z)^{fid} D_A(z)^{fid}} \tag{4.8}$$

The application of the AP test simply needs the knowledge of the true symmetry properties of the object: $\frac{\Delta r_{\parallel}^{\text{true}}}{\Delta r_{\perp}^{\text{true}}}$, and a fiducial cosmology in order to estimate the truth value of the cosmological parameters. This straightforwardness makes it a powerful probe of cosmology.

4.1.2 Application on galaxy surveys

The proper application of the AP test thus depends on the object of interest and the estimation of the symmetry properties. A wide range of objects has been proposed as interesting Alcock-Paczynski candidates.

4.1.2.1 First forays in the Alcock-Paczynski test

When Alcock and Paczynski first proposed the eponym test, in 1979, their application generally targeted the only known spherical like structures that were observed in the Universe: clusters. However, while those were definitely observed, the lack of statistics did not allow for a consistent and unbiased constrain. The test was not considered for a few years until Phillipps (1994) suggested to apply it to quasars pairs, which through the cosmological principle should be isotropically distributed. This application was later considered for galaxy pairs (Jennings et al., 2012; Marinoni & Buzzi, 2010).

Following the same basis of isotropic and homogeneous statistical of matter, the auto-correlation function or power spectrum of galaxies were also considered as candidates for the application of the Alcock-Paczynski test (Ballinger et al., 1996; Blake et al., 2011b; Li et al., 2015; López-Corredoira, 2014; Matsubara & Suto, 1996). In addition to galaxies, a variety of tracers are available to measure the statistical properties of matter. The application of the AP test has also been considered between Ly-α correlation of neighbouring quasar pairs (Hui et al., 1999; McDonald & Miralda-Escudé, 1999), the power spectrum of clusters (Kim & Croft, 2007) as well as the power spectrum from 21cm sources (Nusser, 2005).

4.1.2.2 The Alcock-Paczynski test on voids

The first proposal to use voids as candidates for the AP test comes from Barbara Ryden (Ryden, 1995), in order to probe the deceleration parameter and thus the effect of a possible non-zero cosmological constant. This idea was born from the numerous void studies on the growth and shape of voids (see 2.3.2 and references therein) which indicate that voids tended towards sphericity. Thus, knowing both the voids location and their extent, defined by the boundary galaxies, the AP test could be applied in perfect conditions. However, several obstacles prevented the achievement of a successful constraint. On the one hand, the proposed methodology concerned the application of the AP test on individual void shapes defined from the distribution of galaxies. But the individual shapes are in fact not necessarily spherical, which does not allow to obtain constraints by this way. Moreover, the study was carried out on an early survey of galaxies with low statistics, the number of voids studied appeared to be too small and the number of detected galaxies was not large enough to accurately define the void walls. Finally, the effect of the peculiar motions of galaxies surrounding the voids is degenerate with the AP effect (Simpson & Peacock, 2010), complicating the extraction of the AP signal. The application of the AP test on objects such as voids was therefore postponed until the arrival of large galaxy surveys.

The interest for the application of the Alcock-Paczynski test on voids was rekindled with the pioneering work of Lavaux & Wandelt (2012), in which was stated that voids, on average, could be considered as standard spheres. While voids can have different shapes and sizes, the cosmological principle actually tells us that there is no reason for these voids to be oriented in a specific privileged direction. As a consequence, a large number of voids stacked onto the others should not have a privileged direction. Therefore their shape should be spherical, on average, whenever and wherever in the Universe. This stack can be considered a standard sphere. They further argued that, thanks to this property, the application of the Alcock-Paczynski test would enable to compete with the standard BAO constraints.

Following this work, the AP test was first applied on voids by Sutter et al. (2012a) according to the method developed by Lavaux & Wandelt (2012). The Alcock-Paczynski test was then applied on the DR9 release (Sutter et al., 2014c) and DR12 CMASS sample (Mao et al., 2017). Recently, the application of this test on voids found in 21cm map simulations was also investigated (Endo et al., 2020). All of the works stated above consider a stack of voids as an object.

4.1.2.3 Void stacking

In order to properly define a stack of voids, and thus, our standard sphere candidate, it is necessary to have an algorithm able to extract voids from a given sample of galaxies without imparting any shape to the considered void, in addition to returning their positions, their scale and their defining tracers. As most void identifiers rely on the assumption of a fiducial cosmology, the methodology proposed by Ryden (1995) cannot be applied formally. However, by measuring the shape of the standard sphere defined by the found voids, one can directly probe cosmological parameters and the dark energy equation of state through the relation in Eq. 4.8. Depending on the cosmology adopted, the void stack can either be elongated ($\Omega_m^{fid} < \Omega_m^{true}$) or flattened along the line of sight ($\Omega_m^{fid} > \Omega_m^{true}$).

To this end, the void stacking is a primordial step in the application of the AP test. It consists in piling up voids so that they share a common centre and transforming their coordinates so that their LOS directions are aligned. The voids boundaries are defined through their member galaxies.

Considering a void position $\mathbf{x}_v$ and one of its member particle position vector $\mathbf{x}_{p,i}$, both in comoving frame. The LOS is defined by that of the void. The positions are brought back to their position relative to the void centre:

$$\mathbf{\Delta x}_{p,i} = \mathbf{x}_{p,i} - \mathbf{x}_v. \tag{4.9}$$

These relative coordinates are then rotated so that the z-direction is aligned with the LOS, corresponding to that of the void centre: $(\Delta x_\perp, \Delta y_\perp, \Delta z_\parallel)$, where the centre of the coordinate frame corresponds to the void centre. Finally, the coordinates are rescaled by the radius of the void r_v. This last step enables to define the stack without diluting the AP signal through the stacking of voids of differing sizes. This process is reiterated for each void in the sample and each member galaxy pertaining to the considered void. Although in this chapter, we consider the application of the Alcock-Paczynski test on an object resulting on a stacking procedure of voids, it is to be noted that the study of the two-point correlation function between voids and galaxy can be considered as a stack, sensitive to the same effects.

4.2 Choice of the estimator of the Alcock-Paczynski distorsion

The distortion induced by the Alcock-Paczynski effect is oriented along the line-of-sight. As such, a spherical object should find itself flattened or elongated along this direction. In order to estimate this deviation and thus, underlying true cosmology of the universe, we consider the stack of voids as a distorted sphere along the line-of-sight: a spheroid.

As such, the ellipticity of a stack is defined as the amount by which is the stack is distorted along the line-of-sight, that is to say the ellipticity quantifies the distortion from unity where unity corresponds to a sphere[1]. This ellipticity as shown above is directly dependent on the cosmology adopted to transform redshifts into comoving coordinates. Sphericity is thus attained when the fiducial cosmology and the true underlying cosmology agree. Hence, an essential step of the Alcock-Paczynski test lies in the estimation the ellipticity of the stack and its sensitivity to the distortion of the stack along the LOS.

In this section, I present the definition of several possible estimators of the ellipticity and apply them on a toy simulation of distorted spheres. The results are then compared in order to qualify each method and decide on the most appropriate estimator of the ellipticity of a stack.

[1] Such an ellipticity is not to be mistaken with the one used in the definition of the ellipsoidal model e ; which tends to 0 when reaching sphericity. This e is generally used to qualify the shape of an individual void.

4.2.1 Ellipticity Estimators

In this part, two estimators of the ellipticity of a void stack are presented. The first estimator is based on the fitting of the galaxy positions surrounding the voids with the parametric equations of an ellipsoid. This method was initially proposed by Lavaux & Wandelt (2012) and Sutter et al. (2012a) to be applied directly on the density field. The second estimator does not consider the contour of the shape but is sensitive to the asymmetry between two perpendicular axes. This is known as the inertia tensor and has already been applied to voids in previous studies of the Alcock-Paczynski test (Mao et al., 2017; Sutter et al., 2014c).

In the following, we assume that the coordinate system has as origin the centre of all the stacked voids and that each particle of the void, i.e. each member galaxy of the void, is positioned at coordinates (x, y, z) with respect to the void centre.

4.2.1.1 Ellipsoidal model

A common way to model a void stack is that of the ellipsoid. It is considered that a stack of voids that is transformed into a fiducial real space (as opposed to true real space) will be spherical on average. The Alcock-Paczynski effect, induced by the use of a fiducial cosmology, will cause an elongation or a contraction of the structure. The void stack is supposed to have 3 directions, of which only one is aligned with the line-of-sight. As such, we can consider that the shape of the void stack will be consistent with an ellipsoid for which the parametric equation is as follows :

$$\frac{(x-\nu)^2}{a^2} + \frac{(y-\mu)^2}{b^2} + \frac{(z-\upsilon)^2}{c^2} = 1, \tag{4.10}$$

where a, b, c are the half-axis of the ellipsoid, ν, μ, υ the coordinates of the centre of the ellipsoid and x, y, z the coordinates of the points at the surface of this ellipsoid, defined as:

$$\begin{aligned} x &= a\cos\theta\sin\phi, \\ y &= b\sin\theta\sin\phi, \\ z &= c\cos\phi. \end{aligned} \tag{4.11}$$

In this framework, ν, μ, υ would be the centre of the void stack at coordinates: $(0, 0, 0)$ and the x, y, z positions that of the galaxies defining the void stack. The half-axis of the ellipsoid (a, b) should be consistent with the angular directions of the stack, while c represents the half-axis along the LOS.

4.2.1.2 Parametric definition of the ellipticity

From the parametric equation, we can define an ellipticity estimator as follows :

$$\epsilon = \sqrt{\frac{2\ c^2}{a^2 + b^2}}. \tag{4.12}$$

If we extend the model to the case of a spheroid, where both angular components have equal contributions, that is to say, $a = b$. The subsequent ellipticity reduces to:

$$\epsilon = \sqrt{\frac{c^2}{a^2}}. \tag{4.13}$$

4.2.1.3 Inertia tensor approach

While still being in an ellipsoidal model of a void stack, another non-parametric method can be applied to the positions of the galaxies in the void stack: the inertia tensor. This estimator has already been used in previous applications of the Alcock-Paczynski test (Mao et al., 2017; Sutter et al., 2014c).

The inertia tensor is by definition *an estimator of the distribution of mass (or mass tracers) around a rotation axis for a given solid.* In its general form, the use of the inertia tensor imposes no specific shape to the solid considered. The moments of inertia are computed by *the sum of the masses pertaining this system in regard to the rotation axis considered*, that is to say, the sum of the masses defining the system weighted by the distance to the considered axis.

Considering our void stack as a solid in which matter is positioned around its centre along the three dimensions (x, y, z) (or rotation axis), where z corresponds to the line-of-sight, its total inertia tensor can be expressed through the following tensor:

$$I = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix}, \tag{4.14}$$

where the diagonal elements are the *moment of inertia* of the stack, that is the sum of the mass defining the system weighted by the distance to the considered rotation axis. The off-diagonal elements are called *the product of inertia*, they measure the imbalance of the mass distribution in the body. The total contributions in the 3D inertia tensor for a void stack can thus be computed as follows:

$$\mathcal{I} = \begin{pmatrix} \sum_i m_i(y_i^2 + z_i^2) & \sum_i m_i x_i y_i & \sum_i m_i x_i z_i \\ -\sum_i m_i y_i x_i & \sum_i m_i(x_i^2 + z_i^2) & \sum_i m_i y_i z_i \\ -\sum_i m_i z_i x_i & -\sum_i m_i z_i y_i & \sum_i m_i(y_i^2 + x_i^2) \end{pmatrix}. \tag{4.15}$$

As our stack takes its root at the centre of our coordinate system, the considered mass m_i is only that of the galaxies. In practice, we consider the galaxies to be all equivalent so that $m_i = 1$, but it is also possible to consider the mass to be that of the combined systematic weights of the galaxies, as done by Mao et al. (2017) for the CMASS BOSS sample, or the weights presented in 3.1.2.3. The pertinence of the inertia tensor in the shape measurement is found in the fact that it measures the diagonal distributions of the masses, those allowing to discern a preference of a rotation axis over another.

In the case of the Alcock-Paczynski effect, preference of one axis over the other two – the z-axis or LOS (in the case of an elongation of the stack) or the x/y-axis (the angular contributions, in the case of a contraction of the stack) – is expected from the use of a wrong fiducial cosmology. As such, the inertia tensor is highly relevant in this type of study pertaining to the shape of the stack of voids.

In the framework of the ellipsoidal model of the stack, the inertia tensor eq. 4.15 is defined as purely diagonal in which each of the moments of inertia is related to a value of the half-axis of the ellipsoid (as defined in eq. 4.10):

$$I_{xx} = \frac{m}{5}(b^2 + c^2), \tag{4.16}$$
$$I_{yy} = \frac{m}{5}(a^2 + c^2), \tag{4.17}$$
$$I_{zz} = \frac{m}{5}(a^2 + b^2). \tag{4.18}$$

with $m = V\rho$. The volume of the considered ellipsoid is $V = \frac{4}{3}\pi abc$, and ρ is the constant density within the ellipsoid.The ellipticity as defined in 4.13 can thus be recovered through the following combination:

$$\epsilon = \sqrt{\frac{I_{xx} + I_{yy} - I_{zz}}{I_{zz}}}, \tag{4.19}$$

and plugging in the formulae of each moment of inertia, the above equation reduces to:

$$\epsilon = \sqrt{\frac{2\sum_i z_i^2}{\sum_i x_i^2 + y_i^2}}. \tag{4.20}$$

4.2.2 Basic simulations of void stacks

In the previous sections, I presented two formulations of the estimators of the ellipticity that can be applied to a stack of voids, or any deformed object of known symmetry properties, in order to measure a deviation caused by the Alcock-Paczynski effect. Both are based on the key assumption that a void stack deformed by the use of a wrong fiducial cosmology can be thought of as an ellipsoid or spheroid. Prior to any application to mock data or data itself, it was decided to test these estimators on a simple toy Monte Carlo model. The goal of this procedure was to test the robustness of the estimator depending on several conditions, such as the dispersion of the particles in regard to the void centre, the distribution of the particles around the voids or the number of particles defining an individual void. All these conditions may bias the estimation of the ellipticity of the void stack.

The first step is thus to investigate the sensitivity of our ellipticity estimators on an ellipsoid which reproduces some properties of the void stack. The properties are the following:

- A void stack is constituted of an ensemble of N_v individual voids,
- Each individual void is constituted of a number N_p of galaxies,
- The density of the total number of galaxies around the centre of the void stack is diffused. As such, there is a dispersion (or shell thickness) to be taken into account: σ_r

Two sets of mock stack generator were produced allowing control on the following aspects of the stack:

- The values of the half-axis a, b, c which drive the input ellipticity,
- The number of ellipsoids to simulate, each ellipsoid being related to a void,
- The number of particles defining each individual ellipsoid,
- The dispersion of the particles from the surface of the ellipsoid in %, σ_r.

Their difference relies on the way to draw the distribution of particles on the shell of the ellipsoid and their subsequent positioning around the void centre. For both simulations, we position the stack's centre at coordinates $(0, 0, 0)$, and the particles at coordinates (x, y, z) with arbitrary units.

4.2.2.1 Simulation I - Inhomogeneous distribution of points cloud

Voids galaxies are drawn along an ellipsoidal distribution around the void centre with defined half-axis (a, b, c), following the parametric equations of the ellipsoid defined in Eq. 4.10. The points are first defined by their parametric definition (4.11) for which a, b, c are fixed while ϕ and $cos\theta$ are drawn uniformly between $0 <= \phi < 2\pi$ and $0 <= \cos\theta < 1$ respectively.

For each particle, in order to place the particles in a uniform and homogeneous manner on the surface of the ellipsoid, we apply a Metropolis-Hastings method taking into account the surface element of an ellipsoid, defined as follows:

$$dS = \sin\phi\sqrt{c^2 \sin\phi^2(a^2 \sin\theta^2 + b^2 \cos\theta^2) + a^2 b^2 \cos\phi} \tag{4.21}$$

At each step, the couple (θ, ϕ) drawn for a given a, b, c, is used to compute the associated element of surface dS_{sim}. dS_{sim} is then compared to a random variable dS_{rand} ; dS_{rand} being a value drawn in a uniform distribution between 0 and the maximal surface element $dS(\phi = \pi, \theta = \pi/2, a, b, c)$. Two conditions were applied to the dS_{sim} : the first one is that it must lower than dS_{max}, the second one, in the Metropolis-Hastings framework is that $dS_{sim} > dS_{rand}$. If the conditions are not respected, the couple (θ, ϕ) is rejected and a new couple is drawn. If the couple (θ, ϕ) obeys the condition, the corresponding coordinates x_i, y_i, z_i are used to determine the position of the i^{th} particle on the surface of the ellipsoid.

After simulating the positions on the surface on the ellipsoid in a homogeneous and uniform manner, we must apply a deviation from the surface to the positions following a given standard deviation σ_r. For this reason, we compute the vector $\mathbf{N}$ normal to the tangent passing by the point $\mathbf{X}_i$. The components of this vector are defined by the derivative of the parametric equation of the ellipsoid (4.10) :

$$\mathbf{N}(x, y, z) = (\frac{2x_i}{a^2}, \frac{2y_i}{b^2}, \frac{2z_i}{c^2}) \tag{4.22}$$

After normalisation, we obtain the normed vector orthogonal to the surface of the ellipsoid $\mathbf{n}$. This vector is multiplied by a value dr drawn in the Gaussian law $G(0, a\sigma_r)$ for which the standard deviation corresponds to a percentage σ_r of the value of the axis a, b or c, depending on the simulated ellipsoid.

This deviation from the surface of the ellipsoid is then added to the parametric coordinates simulated previously:

$$\begin{aligned} x &= a\cos\theta\sin\phi + dr_x \\ y &= b\sin\theta\sin\phi + dr_y \\ z &= c\cos\phi + dr_z. \end{aligned} \tag{4.23}$$

The Metropolis-Hastings continues until the desired number of particles to define a void is reached. It is reiterated as many times as the number of voids simulated.

The output of the stack simulation is shown in Fig 4.1 for an ellipsoid of $(a = 40, b = 40, c = 45)$ defined with 4000 voids with varied N_p following the distribution of N_p of the CMASS sample. The stack thickness encoded in σ_r is varied in the range 0.01% to 0.35%, up until all of the volume is populated. The stack thickness is well taken into account although the relation is not linear. Limitations to the simulation can be perceived as the stack's thickness increases: the density seems to shift toward the coordinate centre along the z-axis. The stack is also much larger along the z coordinate axis, compared to that of the d-axis as the thickness increases. The simulation is considered inhomogeneous because the transformation applied to the coordinate system (4.23) conserve neither uniformity nor homogeneity of the point picking, rendering the stack anisotropic.

4.2.2.2 Simulation II - Homogeneous distribution of the probability density function of points cloud

This second simulation relies on the distortion of a given sphere. The initial placement of the particles is first drawn on the unit sphere. The fundamental

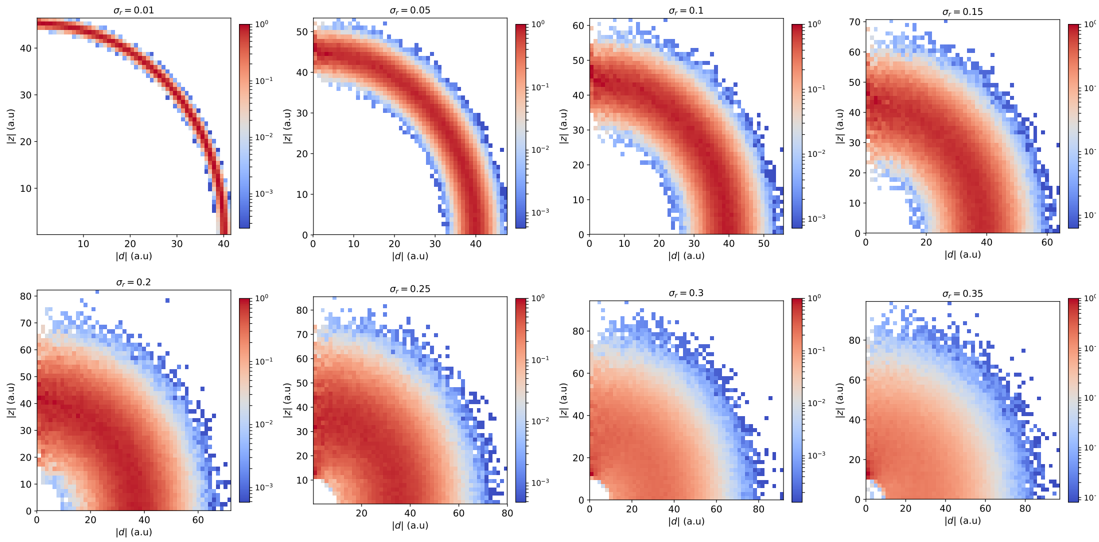


Figure 4.1: Simulated stack in the range of dispersion σ_r in the range 1% to 35% deviation using Simulation I with a Metropolis Hastings scheme and added perturbation of the point picking at the surface of the ellipsoid.

difference here is that the thickness of the stack intervenes directly in the generation of the radial component of the points considered instead of an added perturbation normal to the surface of the ellipsoid.

The volume of the simulated individual ellipsoid $V = a^3$, where a is the half-axis of the ellipsoid simulated along the transverse direction of the LOS, instead of the direct radial component. The volume of the ellipsoid is drawn in a gaussian distribution $G(a^3, \sigma_r a^3)$ where the width of the gaussian is dependent on the shell thickness parameter σ_r. We take $a = b$ so that we simulate a spheroid. The radial component of the sphere is then recovered through:

$$r = \sqrt[3]{G(a^3, \sigma_r a^3)}, \tag{4.24}$$

yielding points distributed in a shell where the density of the point picking follows a Gaussian distribution. The sphere is then distorted along the z-axis, taking the ellipticity defined in 4.12 (or 4.13 in the chosen configuration) in a linear transformation:

$$r_z = \epsilon \sqrt[3]{G(a^3, \sigma_r a^3)}. \tag{4.25}$$

It is to be noted that the σ_r quoted here is of different nature than that of the previous simulation. While it still relates to the thickness of the stack, it represents here a volume dispersion while in the first simulation it was directly related to the radius.

The resulting stacks are shown in Fig. 4.2 with the same initial conditions as the previous simulation in Fig. 4.1, that is : $N_v = 4000$ ellipsoid simulated with $(a, b, c) = (40, 40, 45)$ and an N_p distribution drawn from that of the CMASS DR12

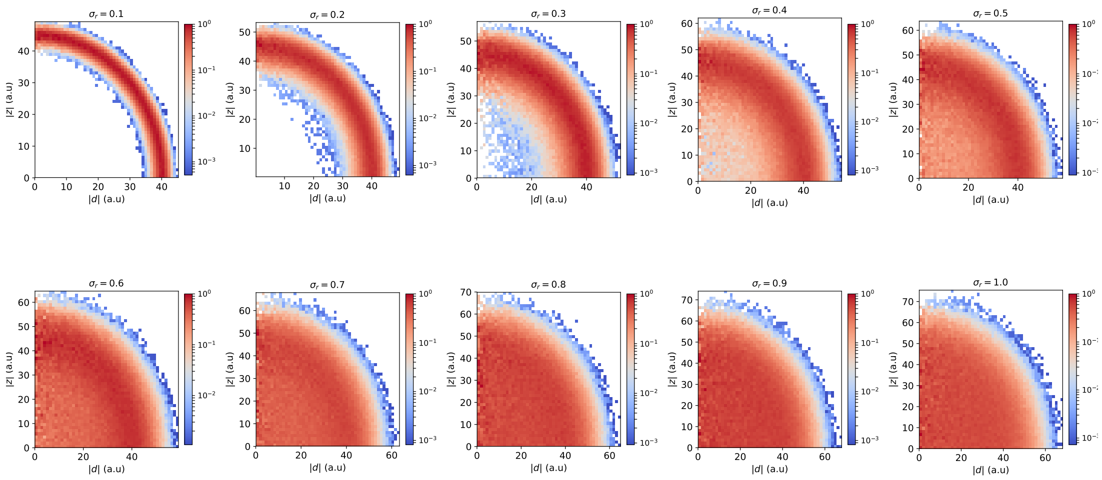


Figure 4.2: Simulated stack in the range of dispersion σ_r in the range 1% to 100% deviation using the Simulation II with a gaussian point picking on the volume of the ellipsoids

sample. The relation between the thickness of the stack σ_r and its resulting aspect is much more straightforward. The simulation is deemed homogeneous due to the density of points within the shell being controlled by the gaussian volume drawing and subsequent inference of the radial component. This allows prospecting over a range of 1% to 100% deviation of the thickness of the stack in the ellipsoid.

Both simulations are useful in assessing the robustness of our estimators. An inhomogeneous cloud of points could highlight a bias in the estimator, while the homogenous distributions would correspond to the most optimistic conditions of a stack of voids.

4.2.3 Measuring the ellipticity

The simulated stacks above are used to qualify our estimators depending on the number of particles defining the void and the thickness of the stack. The coordinates of the particles around the stack centre are known, as well as their associated property that is the number of particles defining them. A first approach was to investigate the number of particles used to define the void. Previous Alcock-Paczynski analyses (Lavaux & Wandelt, 2012; Mao et al., 2017; Sutter et al., 2012a, 2014c) considered all of the particles defining the void stack in order to estimate the ellipticity, thus giving more importance to voids defined with a large number of particles. The goal of this study is to investigate whether the deviation from sphericity of the stack can be recovered when considering a void per void approach: that is to measure the distortion on the individual voids, thus removing the influence of the void's extent and N_p.

To test both procedures, stacks are simulated with a varied thickness, depending on the simulation. The stacks are generated with 4000 individual ellipsoids, each being defined by a set number of particles, N_p. In order to test for a possible dependency of the ellipticity measurement in the N_p parameter, we vary the number of particles in the range: $N_p = [10, 15, 20, 25, 50, 75, 100]$.

4.2.3.1 Parametric fit

In order to estimate the ellipticity of the void stack with the parametric approach as given by Eq. 4.12, one must determine the values of the half-axis defining the ellipsoid. To this end, we fit the parametric equation of the ellipsoid defined in Eq. 4.10 with the following two-step procedure:

- A first global fit is applied which consists in minimizing the following expression:
  $$\chi^2 = \sum_i^{N_{tot}} (\sqrt{\frac{x_i^2}{a^2} + \frac{y_i^2}{b^2} + \frac{z_i^2}{c^2}} - 1)^2 \tag{4.26}$$
  which is nothing else than the minimization of the distance of a point to the centre of the ellipsoid $(0, 0, 0)$ for which each of its coordinates is normalized by the corresponding half-axis. The vector (x_i, y_i, z_i) corresponds to the cartesian coordinates of the i^{th} particle among the N_{tot} defining the stack. In this first fit, we obtain a primary approximation a_0, b_0, c_0 of the half-axis of the ellipsoid. These allow us to measure the global variance of the distance of the particle to the centre of the stack:
  $$\sigma^2_{stack} = \mathrm{Var}(\sqrt{\frac{x^2}{a_0^2} + \frac{y^2}{b_0^2} + \frac{z^2}{c_0^2}}). \tag{4.27}$$
  This variance takes on the role of global error in the second part of the fit.
- The second step consists in applying the fitting procedure on a void per void basis, taking into account the global dispersion of the distances of the particles to the centre of the stack. The χ^2 expression to minimize takes the following form:
  $$\chi^2 = \sum_j^{N_{part}} (\sqrt{\frac{x_j^2}{a^2} + \frac{y_j^2}{b^2} + \frac{z_j^2}{c^2}} - 1)^2/\sigma^2_{stack}, \tag{4.28}$$
  where (x_j, y_j, z_j) corresponds to the cartesian coordinates of the j^{th} particle among the N_p particles defining the void considered.

After fitting the (a, b, c) axes to each individual ellipsoid, an individual ellipticity is computed respective to each mock void following the parametric definition of the ellipticity (4.12). The stack ellipticity is then evaluated as the mean of the individual ellipticities.

Fig. 4.3 shows the deviation of the measured ellipticity of the stack in regard to that of the expected ellipticity as a function of the stack thickness squared, σ_r^2. Both panels illustrate the deviation between the expected ellipticity and the ellipticity measured through the fitting procedure presented above. Panel 4.3a displays the resulting deviation in the case of the Simulation I where a near linear relation between the deviation from the expected ellipticity and the thickness of the stack can be noted for a number of particles definition the void superior to twenty. Panel 4.3b corresponds to the Simulation II. As we can see, the more the particles define the void, the less biased is the estimate of the ellipticity. The bias seems to be stable for the stack populated by voids with a higher number of particles. It remains below 4% for voids with more or equal to twenty particles.

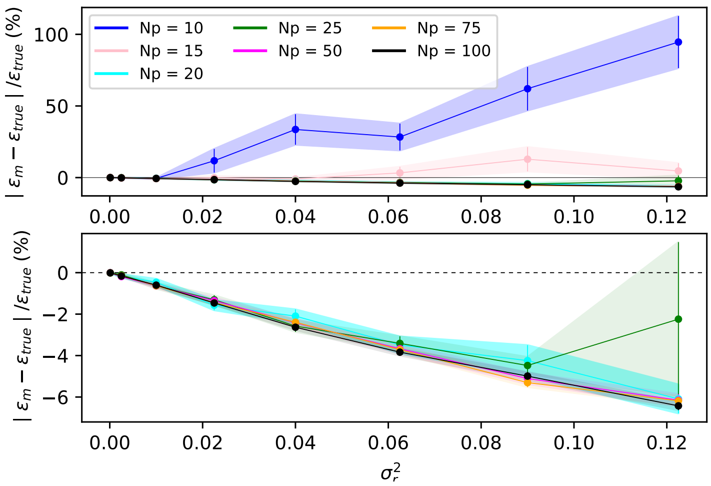
The errors estimated for this measurement corresponds to the standard deviation of the recovered ellipticity in one stack. It is to be noted that they are, as such, underestimated. Nevertheless, it surely would not affect the bias seen in with the fit of the ellipsoid.

On a void per void basis, it seems that recovery of the ellipticity is strongly unreliable when considering voids defined with a low number of particles. This sensibility to a low number of particles is somehow expected and illustrates the impact of the sampling. The less populated is the void, the less it is resolved and the resulting shape may differ significantly from the initial ellipsoid. As such, it seems that one needs to consider voids defined by a number of particles superior to 20. However, the recovery is still affected by the stack thickness and density distribution as the inhomogeneous ellipticity distribution shows an increasing bias.

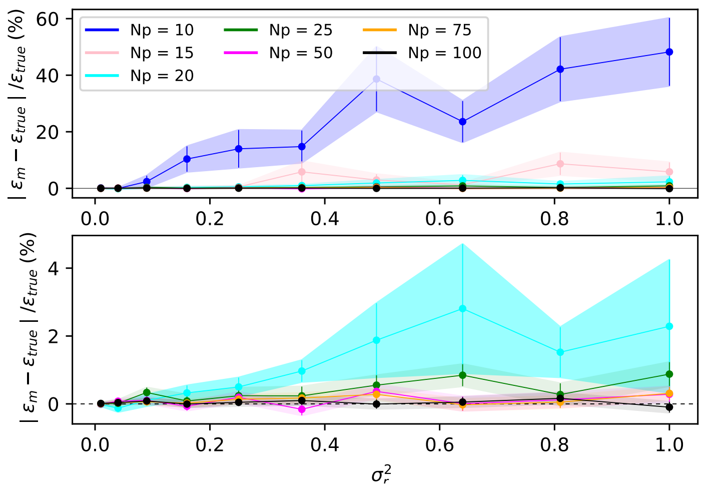
4.2.3.2 Inertia Tensor

The estimation of the ellipticity with the inertia tensor is straightforward. One only needs the (x,y,z) coordinates of the member particles respective to the individual void and measure the inertia tensor following the formula given by Eq. 4.20 on a void per void basis, allowing to recover N_v estimates of the ellipticity. The total ellipticity of the void stack is taken to be the mean of the individual ellipticities, thus negating the impact of the number of particles.

Figure 4.4 shows the deviation of the measured ellipticity of the stack in regard to the expected ellipticity as a function of the squared stack thickness σ_r^2. Both panels show the deviation between the expected ellipticity and the ellipticity measured through the inertia tensor procedure presented above. Panel. 4.3a shows the resulting deviation in the case of the Simulation I. In this case, the inertia tensor presents an offset of more than 1%, regardless of the number of particles, which decreases with the stack thickness. As the estimated errors correspond to the error on the mean of the 4000 estimated ellipticities, they are underestimated. It seems then that this bias of the inertia tensor is not quite dependent on the number of particles used to define the void. The inertia tensor systematically underestimates the ellipticity in the inhomogeneous simulation.



(a) Inhomogenous ellipticity estimation



(b) Inhomogenous ellipticity estimation

Figure 4.3: Deviation of the measured ellipticity of the stack through the ellipsoid fitting procedure in regard to that of the expected ellipticity as a function of the stack thickness square σ_r^2. Each line colour corresponds to a set number of particles used to define the void, ranging from 10 to 100. Top: deviation measured with an inhomogeneous stack generation or Simulation 1. Bottom: deviation measured with an homogeneous stack generation or Simulation 2.

In the case of simulation II, shown in Fig. 4.3b, the bias remains below 2% overall and presents an around 1% deviation from the expected value in the case of a low number of particles and below 0.5% as we consider more highly populated void.

Considering the under-estimation of the error, it is again possible, as with the Simulation I, that the number of particles has less of an impact on the estimation of the ellipticity. On a void per void basis, the use of the inertia tensor shows a higher bias in the case of Simulation I, although its evolution in regard to the stack thickness is not strong. On the other hand, the deviation in the case of Simulation 2 does not seem to be sensitive to the stack thickness, nor quite to the number of particles.

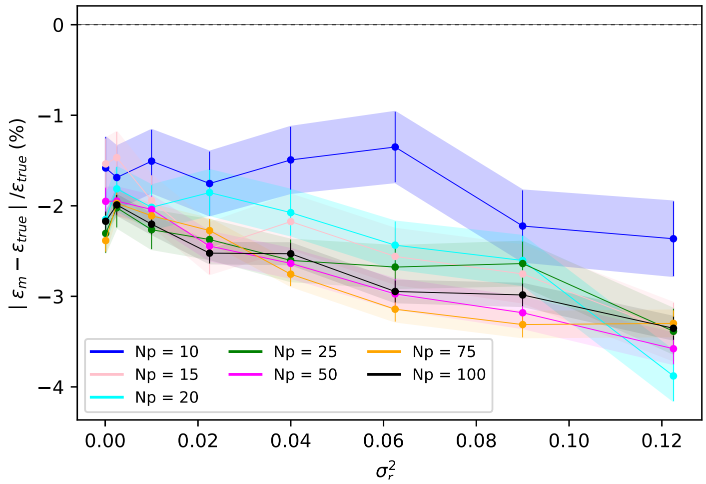
Independently of the methodology considered, the bias decreases significantly with the number of particles defining the voids and the subsequent stack. This hints to the fact that the application of the inertia tensor of the whole stack may mitigate such a bias. The inertia tensor seems less sensitive to the number of particles defining the voids, whereas the ellipsoid fitting procedure requires well-defined voids (and as such, more populated) to have a significant estimation of the half-axes.

4.2.3.3 Semi-realistic case

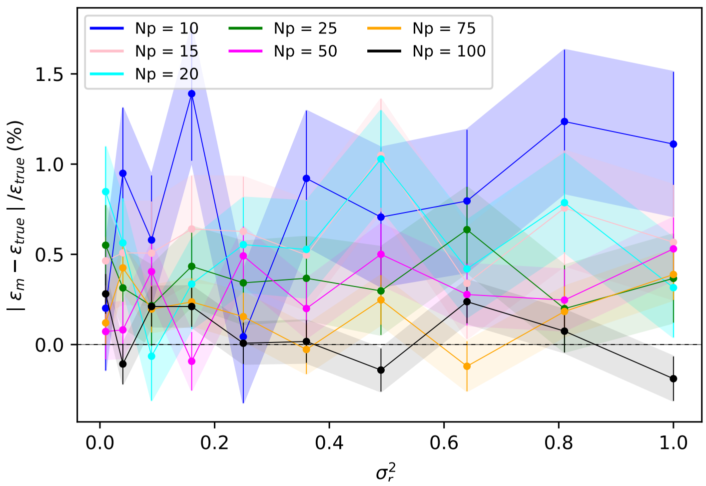
Voids, in a stack of data or mocks, are not defined by a set number of particles, the distribution of such particles is actually much wider. The estimation of the ellipticity of a void stack with a varied number of void particles was applied to the stacks displayed in Fig.4.1 and Fig.4.2, on a void per void basis, and on a whole stack approach. The latter consists in fitting the parametric equation of the ellipsoid or computing the inertia tensor on all the particles of the stack. To this end, the stack was simulated with 4000 voids in the same range of thickness, respective to Simulation I or Simulation II. To each void was associated a number N_p randomly drawn from the N_p distribution extracted from the DR12 data, after void finding. The minimal number of particles allowed in the ellipsoid definition is set to $N_p = 10$.

Figure 4.5 shows the deviation from the expected ellipticity for both types of stack simulation in regard to the stack thickness. The right panel 4.5a displays the deviation expected for ellipticity for the inhomogeneous simulation. We can see that the void-per-void fit applied to the stack becomes more and more biased up to 5% deviation from the expected ellipticity. The application of the fit to the whole stack displays a similar behaviour, reach up to 7% deviation, albeit with a reduced statistical error. In the case of the inertia tensor, whether it is applied void-per-void or to the whole stack, there is a near-constant bias slightly higher than 2% regardless of the stack thickness. As a consequence, in the case of an unknown distribution of the galaxies around the centre, the inertia tensor seems to be more appropriate, even though the presence of a 2% bias may occur. But, this could be taken into account in the systematic error budget.

The left panel 4.5b shows the deviation from the expected ellipticity in regard to the stack thickness in the case of the homogeneous simulation. In this configuration, the stack thickness ranges from 0% to 100% deviation of the thickness ellipsoid. The bias on the estimation is below 2% for every fitting method. For the fit of

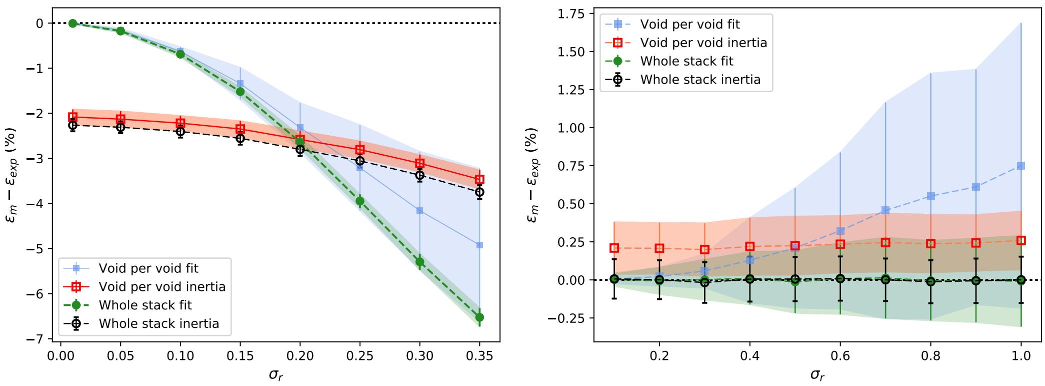


(a) Inhomogenous ellipticity estimation



(b) Homogenous ellipticity estimation

Figure 4.4: Deviation of the measured ellipticity of the stack, estimated through the inertia tensor, in regard to that of the expected ellipticity as a function of the stack thickness square σ_r^2. Each line color corresponds to a set number of particles used to define the void, ranging from 10 to 100. Top: deviation measured with an inhomogeneous stack generation or Simulation 1. Bottom: deviation measured with an homogeneous stack generation or Simulation 2.



(a) Inhomogenous ellipticity estimation **(b)** Homogenous ellipticity estimation

Figure 4.5: Deviation from the expected ellipticity in regard to the stack thickness in the case of the inhomogeneous simulation (left) and the homogeneous simulation (right). In blue is shown the void per void fit, in green the fit applied to the whole stack. The void per void inertia tensor is shown in red and the whole stack inertia in black. The displayed errors correspond to the 1σ dispersion from 500 realizations of each simulation.

the ellipsoid on a void per void basis, the estimated ellipticity seems to deviate at larger σ_r along with the error on the estimated ϵ. The fit applied to the void stack displays no bias, however, the dispersion on the measured ellipticity increases along with the stack thickness. The inertia tensor in the void per void case displays a constant bias of $\sim 0.25\%$, while its application to the whole stack incurs no bias. It is to be noted that the estimated errors remain constant in regard to the stack thickness.

From the comparison of the behaviour of the fit with that of the inertia tensor, it seems the inertia tensor has a near-constant bias that can be modelled and eventually corrected in the case of an inhomogeneous distribution of the galaxies in the stack. Its bias is thus less than that of the ellipsoid in an inhomogeneous setting. The behaviour of the ellipticity depending on the simulation conditions puts to light the fact that our point distribution may not be as homogeneous as simulation II. As such, when we consider an extreme case, the inertia tensor seems to be more appropriate than the fit of the ellipsoid to estimate the ellipticity of our stack. The application of the ellipsoid fit on the whole stack may be applicable, but, the inertia tensor seems to relay a more precise estimation at higher stack thickness.

This experiment shows two things:

– The void per void ellipticity estimation might not be the best method to estimate the ellipticity of a stack regardless of the particles. It may be biased, which, in the event of a large statistic, may prevent the AP test from making stringent constraint.

– The inertia tensor seems to be the best estimator of the ellipticity of a stack as it is more stable and less prone to varying biases.

4.3 Idealistic void stack and sensitivity to the AP effect

In the previous section, we tested two parametrizations of the ellipticity estimator and found that the less biased estimator should be the inertia tensor, making it the most appropriate to measure the ellipticity of our stack. Such performances were tested on ellipsoidal cloud of points distributions containing next to no cosmological signal.

In this section, I present a study of our full Alcock-Paczynski pipeline on a simulation of an idealistic stack of voids in a given cosmology in order to quantify the strength of the Alcock-Paczynski effect as well as test our estimator in semi-realistic conditions.

The main goal of this simulation is to test the sensitivity of a void stack to the Alcock-Paczynski signal. In addition to the void stacking procedure, the pipeline simulated has to take into account the "observational" peculiarities of our void that we obtain from our void finding algorithms as well as their void properties: the void density profile, the number of particles, the redshift considered and finally the radius.

4.3.1 Void stack generation

4.3.1.1 General assumptions

The first step of our simulation is to mimic a stack of voids. Such simulation relies on simplistic assumptions in regard to the stack: a void stack should be spherical in comoving real-space. As such, the initial distances of the particles relative to the stack centre are set to be defined by the true cosmology. These assumptions originate from the cosmological principle from which we infer that a stack of voids can be considered as a standard sphere.

4.3.1.2 Simulating the distribution of particles

Following the general assumptions of the stack, the particles are simulated in regard to the void centre. The stack generation samples points on the surface of a unit sphere in a uniform and homogeneous manner. The radial component, however, is no longer generated with the method presented either in Sim-I or Sim-II, as we want the particles around the void centre to follow an empiric formulation of the void density profile, as defined in equation 2.6 in section 2.3.1.1 that is recalled here:

$$\delta(r) = \delta_c \frac{1 - (r/r_v)^\alpha}{1 + (r/r_v)^\beta}$$

where δ_c is the density contrast, an indicator of the under-density level of the void, r_s is the scale at which the density profile crosses the average $\delta(r_s) = 0$ density threshold, and α and β govern the shape of the surrounding wall and the slope of the profile. Finally, r_v is the radius of the void.

In order to be able to take such expression as a probability distribution function, we use an *Inverse transform sampling* method which consists in linking the cumulative probability distribution function (CDF) to a random variable g. The CDF is defined as follows :

$$\mathrm{CDF}[r] = \int_{-\infty}^{-\infty} f(x)\, dx, \tag{4.29}$$

where $f(x)$ is the probability distribution function (PDF) of interest. In this case, the PDF is defined as :

$$f(x) = \frac{\delta(r) + 1}{\int_{-\infty}^{-\infty} \delta(r) + 1\, dx}, \tag{4.30}$$

where $\delta(r) + 1$ allows to have a positively defined probability, and the integral in the denominator normalizes the density profile in order to have a probability between 0 and 1. This step is necessary in order to have a cumulative function that corresponds to the requirements of the validity domain of the CDF: CDF $\in [0, 1]$. The CDF thus gives us the cumulative distribution function for a given interval of r.

The resulting CDF and associated variable r are interpolated, such that, when drawing a uniform sample of randoms value between $[0, 1]$, one can obtain the r associated.

Fig. 4.6 displays the initial void density function in its normalized form input in the simulation along with the resulting particles density profile of the simulated stack.

Voids found by ZOBOV-based algorithms such as VIDE and REVOLVER return voids member galaxies. Therefore, we possess the knowledge of the number of particles defining the voids and the subsequent number of particles distribution. To comply to our ambitions of creating a semi-realistic simulation, in line with our void finding algorithm, the number of galaxies defining each void is also sampled to correspond broadly to that found in the void catalogue output. The PDF chosen is that of a log-normal distribution which is generally defined as:

$$f(x, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp \frac{(\ln x - \mu)^2}{2\sigma^2}, \tag{4.31}$$

where x is the random variable of interest, μ the mean of the distribution and σ its

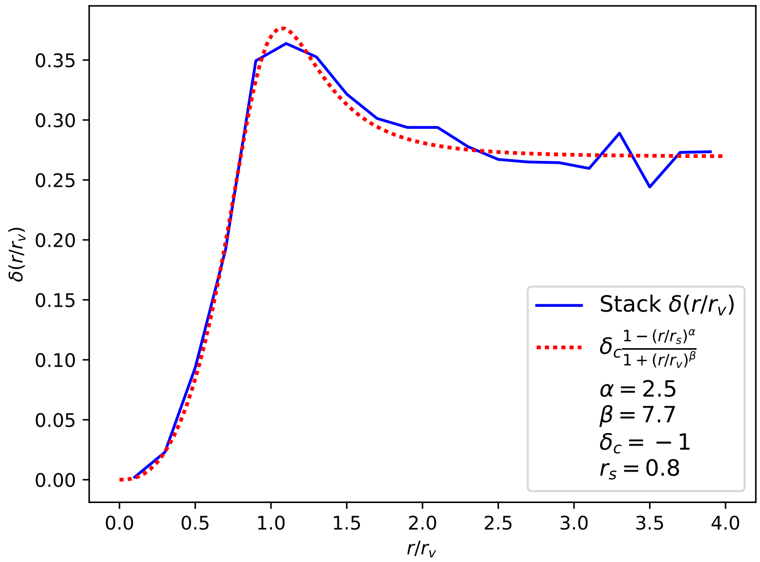


Figure 4.6: Input normalized density profile (in red) as defined by Eq. 4.3.1.2 for which the adopted parameters are given in the legend, and the recovered density profile from the simulated void particles (in blue) for one realization of the stack generator.

standard-deviation.

After comparison with catalogues found in mock data, the log-normal distribution was found to be consistent with the following parametrization:

$$F(N_p) = |Kf(0,1)| \tag{4.32}$$

where K is an arbitrary positive constant which accounts for the spread of the log-normal distribution. The distribution defined by Eq. 4.31, is centered on $\mu = 0$ with a dispersion of $\sigma = 1$ and taken as an absolute value in order to obtain integer and positive values. Fig. 4.7 shows the resulting distribution, drawn from the log-normal parametrization with an arbitrary $K = 48$ value. This distribution is in general agreement with that of one catalogue of the EZmocks.

After sampling N_p separation distances, r, particles are placed in a sphere surrounding the void with component $\cos\theta$ and ϕ. At the end of the stack generation, we thus have the (x^i_j, y^i_j, z^i_j) coordinates of the i^{th} among $N_{p,j}$ member galaxy of the j^{th} generated void among N_v, defined with respect to the void centre. A resulting stack at this stage, can be seen in Fig.4.8.

All these coordinates follow the void density distribution defined earlier. It is to be noted that the separation distances generated are not in h^{-1}Mpc but are considered as rescaled by the void radius r_v. The latter is drawn from a normal distribution distribution $\mathcal{N}(\mu = 40, \sigma = 20)$, with an added condition $R_v \geq 0$.

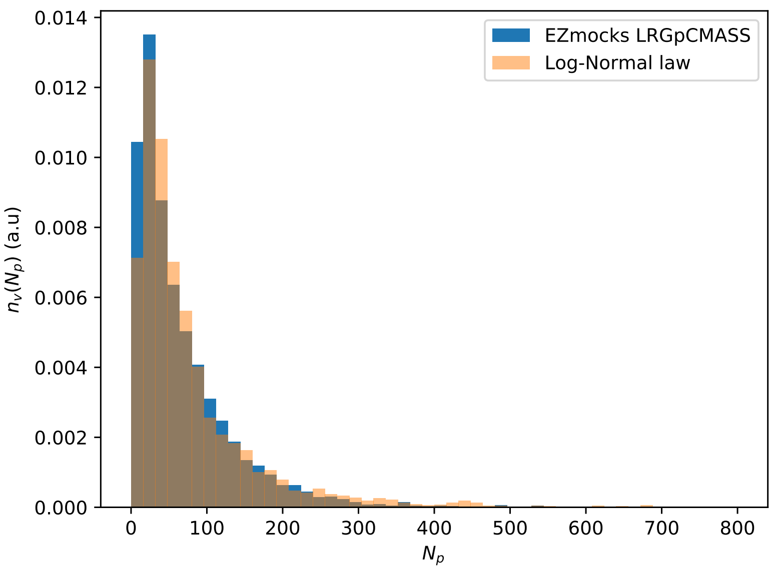


Figure 4.7: Input normalized density profile (in red) as defined by Eq. 4.3.1.2 and recovered density profile from the simulated void particles (in blue) for one realization of the stack generator.

The radius and the number of particles r_v are generally linked when considering the void finding routine, but, as the relation between those two quantities is not explicit, nor trivial, I do not take this feature into account when generating the semi-realistic stack.

4.3.1.3 Disposition in the sky and input of fiducial cosmology

From the void generation procedure is obtained a spherical stack which follows an arbitrary void density profile of the form defined in Eq.4.3.1.2. Each void is associated with a set of particles as well as a radius. In order to recover the positions of the particles in comoving space, the voids are first disposed in the observational coordinate frame $(\mathrm{RA}, \mathrm{DEC}, z)$. Each of these quantities is evaluated from a random uniform distribution of the whole sphere in the case of the angular coordinates and in a set uniform interval $[z_{min}, z_{max}]$ of redshift z.

The comoving coordinates for each void are then computed for a given flat ΛCDM cosmology that is set to be the **true** underlying cosmology of our mock stack:

$$(\Omega_m = 0.31, \Omega_\Lambda = 0.69) \tag{4.33}$$

Through the knowledge of the galaxy positions respective to the void centre along with its attributed void radius, the comoving equivalent of the galaxies position can be recovered, using the transpose rotation matrix of our void stacking procedure.

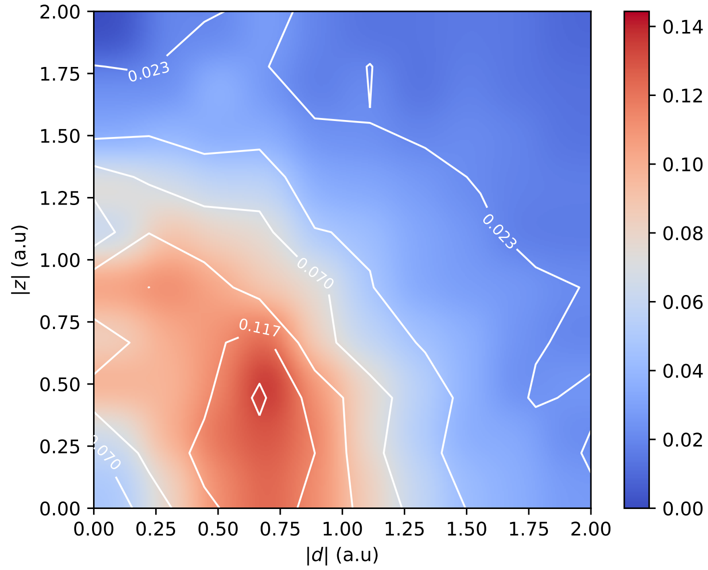


Figure 4.8: A 2D stack of one realization of the void stack generation routine. $N_v = 10000$ voids are generated.

From the comoving coordinates and our knowledge of the true cosmology defined in 4.33, the redshift of the galaxies as well as their right ascension and declination are recovered, providing the stack in $(\mathrm{RA}, \mathrm{DEC}, z)$ for galaxies and voids alike. The radius of the void is also recovered in terms of a Δz in order to modify the radius according to the fiducial cosmology of interest.

In brief, this simulation allows the generation of a mock stack of voids following a given density profile which has approximately the same characteristics as those from voids found with a ZOBOV-based algorithm. The particles positions and thus, the stack, can be controlled by picking the redshift range of interest.

Some of the assumptions used to simulate the stack generally stem from visual comparison with the output distributions procedure of our void finding algorithm. For the purpose of simulating a semi-realistic void stack and testing our Alcock-Paczynski procedure, these choices are deemed sufficient.

4.3.2 Simulating the AP test

After generation a void stack with $N_v = 5000$ in the redshift range: $0.43 - 0.7$ and the recovery of the particle positions in the sky, the Alcock-Paczynski effect is input through the stacking procedure.

Galaxies and voids are stacked in comoving space using a flat ΛCDM fiducial cosmology to convert their positions and radius to the fiducial comoving space. For each void, the line-of-sight is defined to be its position and the member galaxies are rotated so that their LOS corresponds to their respective voids. In this new coordinate frame $(\tilde{x}, \tilde{y}, \tilde{z})$ in h^{-1}Mpc , $\tilde{z}$ is the direction of the line-of-sight while $(\tilde{x}, \tilde{y})$ correspond to their transverse counterpart. Finally, the coordinates are rescaled by the radius of the void in the newly adopted cosmology. The ellipticity of the stack is then estimated using the ensemble of the particles defining it through the use of the inertia tensor estimator, as validated in the previous section. The goal of this simulation is to investigate the methodological aspects of the Alcock-Paczynski test and their influence on the parameter recovery.

The stacking procedure is applied to the stack using a fiducial cosmology which is varied in the range $\Omega_m^{fid} = 0.1$ to $\Omega_m^{fid} = 0.9$. Stacking with a different cosmology from that defined to recover the particle positions in sky coordinates enables to introduce the Alcock-Paczynski effect. For each of the considered fiducial cosmology, the simulation performs 1000 realizations of both the void stack and stacking procedure.

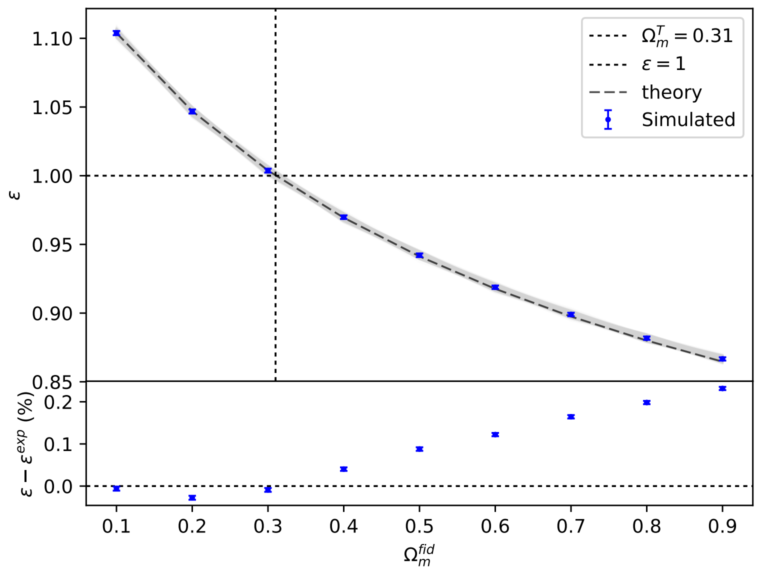


Figure 4.9: Top : Measured ellipticities through the inertia tensor in regard to the cosmology adopted. Shaded regions show the values from the 1000 realizations of the stack. Blue dots corresponds to the mean over all the realizations. Errors are taken to be the standard-deviation of the 1000 realizations. Bottom : Deviation $\epsilon^{meas} - \epsilon^{exp}$ in terms of the fiducial cosmology adopted. Errors are the error on the mean of the 1000 realizations.

The resulting signal is shown in the upper panel of Fig. 4.9 where the mean ellipticity estimated from the total of the realizations is clearly in agreement with the theoretical signal. The lower panel of Fig. 4.9 shows the deviation from the expected value at the considered fiducial cosmology. The errors quoted in the sub-panel are the errors on the mean of the realizations. Looking at the deviation from the theory in more details, however, although it is quite negligible ($\ll 1\%$), a trend can be seen at large deviations of the fiducial cosmology for which the deviation rises to 0.2%. While not quite expected, this effect could either come from the stacking procedure or from the strong AP distortion induced. The latter causing the isotropy of the stack to break. However, it is considered that the deviation at large Ω_m^{fid} can be neglected for the purpose of testing our pipeline.

Once the ellipticity is recovered using the inertia tensor, the true is estimated cosmology through a χ^2 minimisation procedure defined as follows :

$$\chi^2 = (\epsilon - \frac{H(z)^{true} D_A(z)^{true}}{H(z)^{fid} D_A(z)^{fid}})^2 / \sigma_\epsilon^2, \tag{4.34}$$

where z is taken as the mean redshift of the stack.

As we have at our disposal 1000 realizations, the fit is performed on each estimated ellipticity ϵ_i with the standard-deviation of the 999 remaining realizations taken as σ_ϵ. The resulting estimated Ω_m are shown in Fig 4.10. From the upper panel, it can be seen that the estimated Ω_m is generally consistent with the simulation true cosmology within 1σ, except for the most exotic values of the fiducial cosmology that are 0.8 and 0.9. The consequence of the deviation in the estimation of the ellipticity can be seen in the lower panel of Fig. 4.9. Indeed, this deviation permeates in the estimation of Ω_m : while the deviation remains below the percent level up until $\Omega_m^{fid} \sim 0.5$, it then increases up to 2% for the largest Ω_m. A small 0.2% error on our estimator of the ellipticity can yield a near 2% error on the cosmological parameter.

This primary analysis is considered as a baseline for the following tests on the sensitivity of the simulation to our methodology choices.

4.3.3 Sensitivity

While the simulation displays a deviation from the expected signal for larger deviations of the fiducial cosmology in regard to the true cosmology, the simulated AP effect is still present. For this reason, it is possible to test several aspects of our methodology such as the stacking procedure and orientation of the particles around the void centre, the impact of the normalization of the voids stack by the void radius as well as the position of the void centre and selection cuts.

4.3.3.1 Stacking methodology

Our void-particle re-alignment technique relies on the rotation of the galaxy positions in regard to the void centre. The latter's position is set as the line-of-sight for each individual void and their member galaxies are repositioned accordingly.

The influence of the stacking routine is then tested using another methodology of re-alignment along the LOS. It relies on the use of the (σ, π) coordinate system,

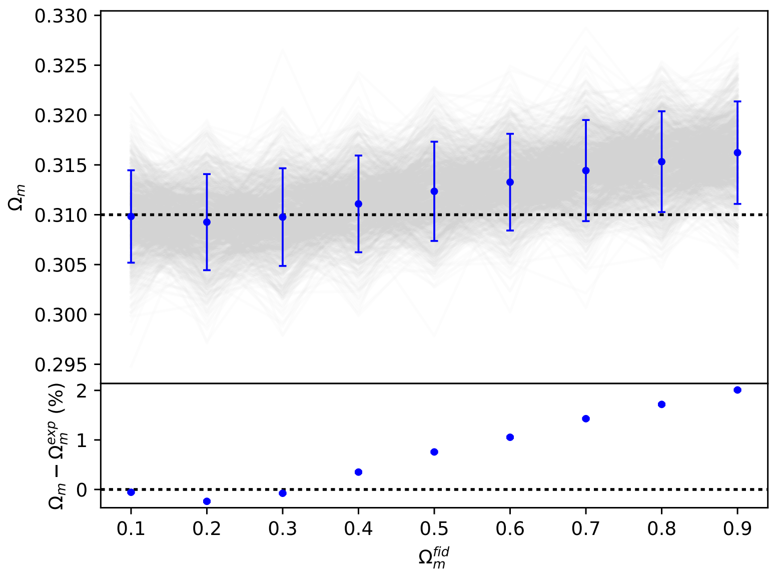


Figure 4.10: Top : Recovered Ω_m through the inertia tensor in regard to the cosmology adopted. Shaded regions shows the values from the 1000 realizations of the stack. Blue dots corresponds to the mean over all the realizations. Errors are taken to be the standard-deviation of the 1000 realizations. Bottom : Deviation $\Omega_m^{meas} - \Omega_m^{exp}$ from the true cosmology in terms of the fiducial cosmology adopted. Errors are the error on the mean of the 1000 realizations.

aka, $(r_\perp, r_\parallel)$, which are commonly used in the estimation of the 2PCF, setting the LOS in between the void centre and the considered galaxy.

To this end, the stack is generated as presented in 4.3.1 with the sole difference of using (σ, π) at the stacking step in order to recover the AP estimation. The resulting deviations from both the expected signal and baseline signal are shown in Fig. 4.11. The deviation on the whole Ω_m^{fid} range in regard to the theory is situated between 0.05% and 0.15%. When comparing with the baseline analysis on the bottom panel, it is clear that while the (σ, π) stacking methodology is more strongly biased for the low Ω_m range, which is closer to the fiducial cosmology, it tends to be less biased on the high end of the Ω_m range.

Considering this behaviour, we maintain our primary stacking method as the choice stacking methodology. Although, it seems clear that, in the event of our simulation not being at fault for this slight bias, one should not stray too far from exotic guess parameters not to bias the cosmological results.

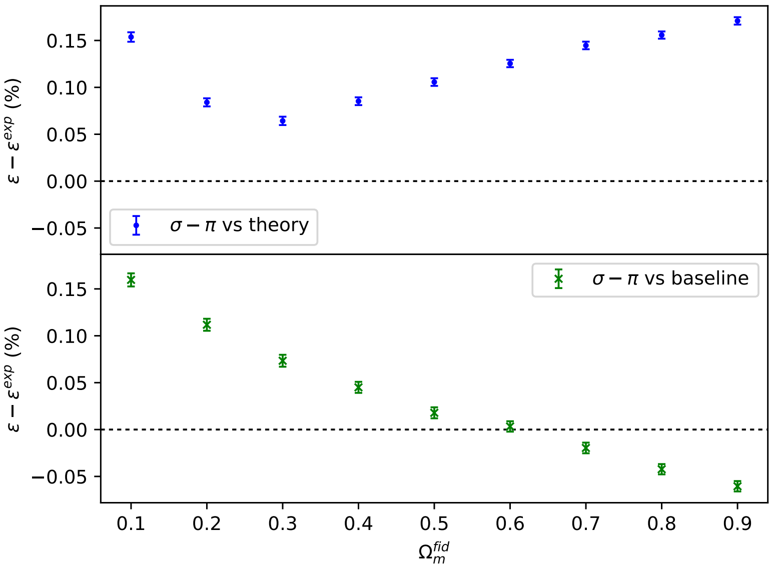


Figure 4.11: Top: Deviation of recovered signal in regard to the expected theoretical signal in % as a function of Ω_m^{fid}. Bottom : Same but in regard to the baseline measured signal. Quoted errors represent the error on the mean of the expected deviation $\sigma_\epsilon/\sqrt{N_{real}}$

4.3.3.2 Void centre definition

The void centre definition is a non-negligible quantity when considering void science. To test the impact of its definition on our measure, the void centre is re-computed as the geometrical barycentre of the member particle of the void. This re-evaluation is performed after computing the positions in the new cosmology, meaning that the void centres are then affected by the change in cosmology as well.

The resulting offset incurred by this change is plotted in Fig. 4.12, in regard to the theory and to the previous void definition. The deviation in regard to the theory remains similar to the baseline analysis and below 0.2%. Comparison with the baseline shows that the new definition of the void centre measurement does not quite affect the measurement of the ellipticity, and remains compatible within 2σ.

4.3.3.3 Stack normalisation

A last aspect of the void stacking procedure is whether to consider the positions of the galaxies in h^{-1}Mpc , thereby mixing the different scales of the void. In the baseline analysis, the separation between the voids and their respective particles is rescaled by the radius of the considered void in order to consider the profile of the void stack without losing any information on the overall shape by a loss of information from the larger voids.

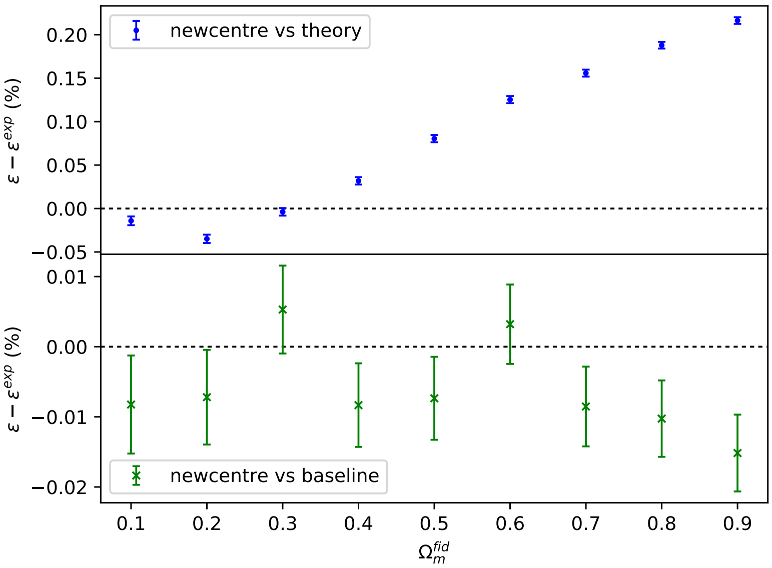


Figure 4.12: Top: Deviation of recovered signal in regard to the expected theoretical signal in % as a function of Ω_m^{fid}. Bottom : Same but in regard to the baseline measured signal. Quoted errors represent the error on the mean of the expected deviation $\sigma_\epsilon/\sqrt{N_{real}}$

The ellipticity of the stack was evaluated without re-scaling the separations to test for the impact of such a choice. The resulting deviation from theory and the baseline are shown in Fig. 4.13. Again, the bias incurred by this change does not alter the general deviation from the theory, which is confirmed when comparing with the baseline. The latter case presents, however, a stronger deviation from the baseline than the one seen with the change of definition of the void centre. At low Ω_m^{fid}, the baseline and no-normalization cases seem to agree within 2σ, at higher $\Omega_m^{fid} \geq 0.6$, a stronger deviation can be noticed.

4.3.3.4 Considered volume

It was been advertised in previous work which measured the shape of the void stack (Mao et al., 2017; Sutter et al., 2014c) that the optimal volume considered would be that of a sphere $R_{max} \leq 0.7R$. With the simulation, we investigate the impact of choosing such cut on the AP parameter measurement, depending on the cosmology. To this end, a cut value r_{cut} is applied to the radial distance of the particles. The minimal value is set to be $0.7R$ and varied up until $3R$. A final value, taken to be vary large $10R$, corresponds to the case where no particles contribution is discarded on the basis of its distance to the centre.

Fig. 4.14 displays the resulting estimation of the ellipticity of the stack ϵ as a function of the fiducial cosmology used. The use of the cut tends to flatten the

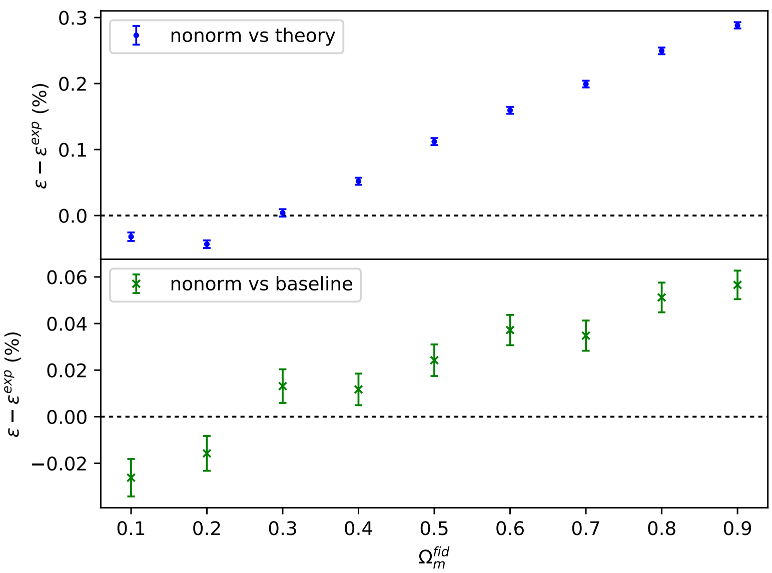


Figure 4.13: Top: Deviation of recovered signal in regard to the expected theoretical signal in % as a function of Ω_m^{fid}. Bottom : Same but in regard to the baseline measured signal. Quoted errors represent the error on the mean of the expected deviation $\sigma_\epsilon/\sqrt{N_{real}}$

estimated stack shape greatly affecting the signal recovery. The imposition of a volume in which is measured the signal can cause a deviation of more than one percent, in perfect conditions, on the recovered signal (depending on the fiducial cosmology adopted). Regardless of the cut adopted, the ellipticities measured all converge to the true cosmology expectation $\epsilon = 1$ thus enabling the recovery of the true cosmology.

4.4 Application of the AP test on data like catalogues

In the previous sections, we validated our pipeline which aims to measure the Alcock-Paczynski effect on a void stack alone. Before any application on the data and various samples (such as the eBOSS DR16 samples), the procedure is tested on a low number of Patchy mocks $N_m = 100$ reproducing the clustering and statistics of the north galactic cap of the BOSS DR12 CMASS sample. This test aims to confront the AP procedure on realistic data and to characterise its sensitivity to the AP signal before any attempt to constrain the cosmology.

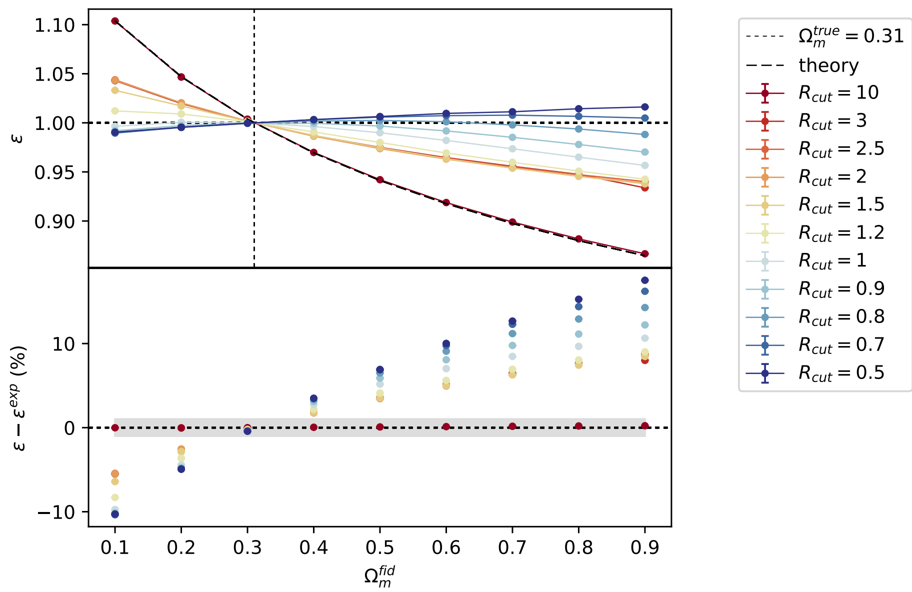


Figure 4.14: Top: Estimated ellipticity in function of the fiducial cosmology with varying volume cuts, circles denotes the estimated measurement for a given cut. Bottom: Deviation in percent from the expected ellipticity, the gray area represents the zone within 1% deviation.

4.4.1 Reference analysis

We find voids with the void finding algorithms presented in section 2.2.3, with a flat Λ-CDM with $\Omega_m = 0.31$. The void finder returns the voids barycentre positions along with their associated particles. The AP pipeline is applied as explained previously. The ellipticity of the stack is then estimated through Eq. 4.20, as it was shown to be the most robust estimate out of the envisioned estimators.

Through the simulation, it was possible to check several methodology choices to validate the recovery signal. In this case, we consider the standard stacking approach (as opposed to the (σ, π) approach), without R_{cut} as it seems to bias the estimation of the ellipticity.

Two major aspects of the void stack definition could not be tested in the simulation: the impact of N_p, the number of particles defining the void and R_v the void radius. Both quantities, in reality, display a dependent behaviour : small voids tend to be defined by a low amount of particles, and conversely, the larger voids are those that are more defined. This relation can be seen in Fig. 4.15 which displays the distribution of the number of particles as a function of their corresponding radius. The figure was truncated at $N_p = 800$ for clarity but the largest voids counts 1200 member particles. This dependency is expected as the volume is defined as the sum of the individual volumes enclosing the particles during the void finding stage. This relation needs to be investigated as the largest voids, although they should tend toward sphericity (Nadathur, 2016), could bias the shape of the stack in regard to the less defined voids. In reverse, voids defined by a low number of particles are more numerous and tend to have more complex shapes.

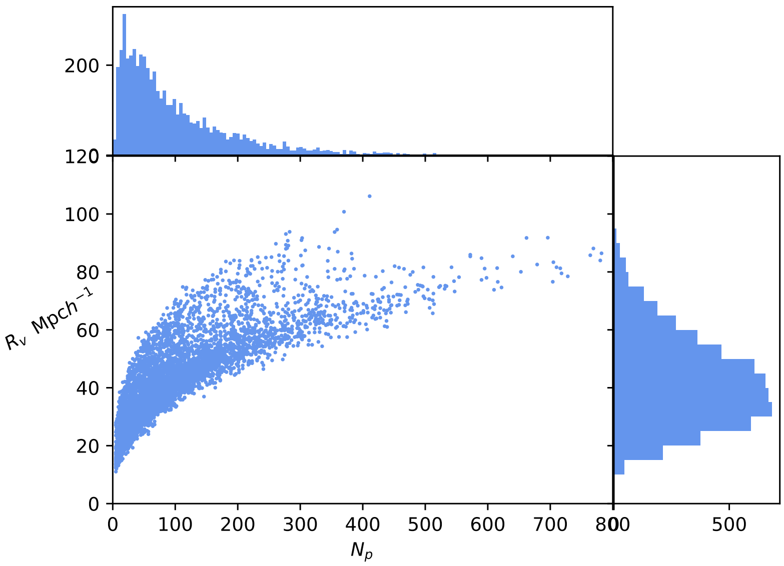


Figure 4.15: Distribution of the number of particles N_p in regard to the corresponding void radius R_v for one mock among the 100 used in this investigation

The pipeline is run on all available voids in each of the N_m mocks, without applying any cuts. The baseline ellipticity is averaged over all mocks and the 1σ dispersion is recovered, yielding:

$$\langle \epsilon \rangle = 0.9904 \pm 0.0046. \tag{4.35}$$

This ellipticity will be used to compare to the recovered ellipticity in the two following subsections.

4.4.1.1 Impact of the number of particles

For the same baseline, that is the catalogues found with $\Omega_m^{fid} = 0.31$, the pipeline is run using several sets of cuts in order to study the impact of the number of particles of the void considered in the stack. No additional cut is applied in this context.

First were applied open cuts on the sample, that is, without bounding the number of particles between both lower and upper boundary. A larger number of cuts are investigated in the lower boundary as it represents most of the statistics, while the upper boundary is chosen loosely. Table 4.1 quotes the impact of the cuts on both the void statistics and estimated ellipticity.

The low boundary cut does not have a major impact on the estimation of ϵ. In comparison with the reference ellipticity (4.35), the subsequent deviation remains below the 1σ error added in quadrature. This is also shown with the deviation of the ellipticity in regard to the expected ϵ which remains stable between -0.83% and -0.87%.

While the cut can affect quite significantly the number of voids stacked, $N_p > 100$ cuts out 62% of the voids, the ellipticity of the stack doesn't seem affected. The dispersion of the ϵ in the mocks does not show a large increase of the error considering the loss of statistics presented by the application of the N_p^{cut}. On the contrary, the impact of adding an upper boundary condition to the stacking of populated voids does not greatly affect the statistics of the stacked voids. Nevertheless, the effect of adding such a boundary can be noted in the estimation of the ellipticity that increases along with the number of cut voids. However, this variation remains within 1σ of the baseline analysis. The effect of the number of particles can be considered negligible.

Table 4.1: Statistics recovered while investigating the impact of an open cut (no upper or lower boundary) on the number of particles defining a void. N_p cut gives the value(s) for which a void is excluded for the stack. N_v^{cut} is the number of cut voids regarding in comparison to the whole sample. The min and max radius of the void are given in h^{-1}Mpc . $\langle\epsilon\rangle$ is the mean ellipticity recovered from the 100 catalogues along with the 1σ dispersion. The deviation from the estimated baseline is quoted in $\langle\epsilon_{\text{base}}\rangle$, with the errors added in quadrature and $\Delta\epsilon$ relay the deviation from the expected ellipticity.

N_p^{cut}	N_v^{cut}	$R_{\min} - R_{\max}$	$\langle\epsilon\rangle$	$\langle\epsilon\rangle$ - $\langle\epsilon_{\text{base}}\rangle$	$\Delta\epsilon^{exp}$
			$N_p \geq N_p^{cut}$		
10	3.3%	$13.8 - 195.3$	0.9904 ± 0.0046	0.0 ± 0.0065	-0.85%
20	12.3%	$17.2 - 126.5$	0.9905 ± 0.0047	0.0001 ± 0.0065	-0.85%
30	21.3%	$20.1 - 126.5$	0.9906 ± 0.0047	0.0002 ± 0.0065	-0.83%
40	29.6%	$22.5 - 125.1$	0.9906 ± 0.0048	0.0002 ± 0.0066	-0.83%
50	36.9%	$24.7 - 125.1$	0.9906 ± 0.0049	0.0002 ± 0.0067	-0.83%
100	62.8%	$33.3 - 125.1$	0.9903 ± 0.0059	-0.0001 ± 0.0075	-0.87%
			$N_p \leq N_p^{cut}$		
600	0.4%	$11.8 - 191.9$	0.9908 ± 0.0048	0.0004 ± 0.0066	-0.81%
400	2.3%	$11.8 - 185.4$	0.9915 ± 0.0042	0.0011 ± 0.0062	-0.74%
200	13.99%	$11.8 - 171.5$	0.9922 ± 0.0045	0.0018 ± 0.0064	-0.67%

In addition to imposing either a lower or an upper boundary to the number of particles, closed cuts are investigated as well. For the lower boundary are considered the minimal number of particles which do not cut more than $\sim 20\%$ of voids out of the sample while are considered all the upper boundaries investigated above. The resulting statistics are shown in Table 4.2, in which are considered all association of a minimal and maximal number of particles.

Considering a low and high boundary to the number of particles defining the stack voids affects the radius range of the considered voids. The variation of the estimated ellipticity remains quite consistent when considering the same upper boundary but varying the lower boundary. This implies that the inclusion of highly populated voids can drive the estimation of the ellipticity. The dispersion of the estimated ellipticity over all $N_{\rm mocks} = 100$ seems impacted by the voids selected through the lower boundary, which is consistent with the diminution of the statistic incurred by such choices. As for the previous cuts considered, the deviation from the estimated baseline for the whole stack remains within 1σ. The effect of the cuts does not greatly influence the estimation of the ellipticity.

Table 4.2: Same as Table 4.1 but considering a closed cut with an upper and lower boundary that are quoted in N_p^{cut}.

N_p^{cut}	N_v^{cut}	$R_{\rm min} - R_{\rm max}$	$\langle\epsilon\rangle$	$\langle\epsilon\rangle$ - $\langle\epsilon_{\rm base}\rangle$	$\Delta\epsilon$
10-200	18%	$14.1 - 170.5$	0.9923 ± 0.0045	0.002 ± 0.0064	-0.66%
10-400	6.4%	$14.1 - 184.5$	0.9915 ± 0.0042	0.0012 ± 0.0062	-0.74%
10-600	4.5%	$14.1 - 191.1$	0.9909 ± 0.0048	0.0005 ± 0.0066	-0.81%
20-200	27%	$17.5 - 100$	0.9925 ± 0.0047	0.0021 ± 0.0065	-0.64%
20-400	16%	$17.5 - 115$	0.9917 ± 0.0043	0.0013 ± 0.0063	-0.73%
20-600	14%	$17.5 - 122$	0.991 ± 0.0049	0.0006 ± 0.0067	-0.8%
30-200	36%	$20.3 - 191.9$	0.9927 ± 0.0047	0.0023 ± 0.0066	0.62%
30-400	24%	$20.3 - 185.4$	0.9918 ± 0.0044	0.0014 ± 0.0063	0.72%
30-600	23%	$20.3 - 171.5$	0.991 ± 0.0049	0.0007 ± 0.0067	0.79%

4.4.1.2 Impact of the radii considered

While the number of particles is at the core of the definition of the stack, it is the impact of the radius that is generally considered regarding the void definition.

Firstly, the radius is common to any void finder which is not the case for the number of particles. Secondly, the size of voids is more indicative of their likeliness to be affected by systematical effects such as the peculiar velocities (both linear and non-linear)(Hamaus et al., 2014a; Pontzen et al., 2016) and to their tendency toward sphericity. Thirdly, considering a large void that is defined with a (reasonably) low amount of particles, it is more likely to truly correspond to an under-density in the sample. In the opposite case, a small void defined by a large number of particles may very well be a local minimum within an over-dense region.

It is generally accepted that voids below the mean tracer separation (m.p.s), defined as $m.p.s = (\bar{n})^{-1/3}$, are more likely to be subjected to non-linear effects which could bias the measurement. In this part, we investigate the impact of the sizes of the stacked voids on the ellipticity considered. The mps is estimated to be $\sim 17\ h^{-1}$Mpc .

Table 4.3 displays the open cuts considered which consist in discarding the small radius with three selection cuts: the mean radius $\bar{R}_v$, the median radius $\tilde{R}_v$ and twice the m.p.s. The statistic ,when discarding the small voids, is seriously depleted, as expected in the case of cutting at the mean or median values of the radius. The range of particles corresponds to the widest range available in the sample. All cuts result in discarding voids below $N_p < 10$. Overall, the estimated ellipticities display a similar behaviour: the deviation from the baseline is within 1σ and the deviation from the expected ellipticity is about 1%.

Table 4.3: Statistics recovered while investigating the impact of an open cut (no upper or lower boundary) on the radius of the stacked voids. The first column R_v cut corresponds to the adopted cuts, N_p gives the minimum and maximum values of the number of particles defining the voids. In parenthesis are given the evaluated values of the cut in h^{-1}Mpc averaged over the 100 mocks.

R_v^{cut}	N_v^{cut}	$N_p^{\min} - N_p^{\max}$	$\langle\epsilon\rangle$	$\langle\epsilon\rangle$ - $\langle\epsilon_{\rm base}\rangle$	$\Delta\epsilon$
$> 2mps$ (34)	32.8%	$10 - 1027$	0.9889 ± 0.0049	-0.0015 ± 0.0067	-1%
$> \bar{R}_v(42)$	54.3%	$19 - 1027$	0.9870 ± 0.0054	-0.0031 ± 0.0071	-1.2%
$> \tilde{R}_v(40)$	50%	$17 - 1027$	0.9873 ± 0.0054	-0.0034 ± 0.0071	-1.2%

As for the selection in terms of particles, we apply more stringent cuts to the void selection prior to the stacking procedure. Two bins are considered : a bin of thickness $1\sigma_{R_v}$ which is about 15 h^{-1}Mpc and $2\sigma_{R_v}$ which corresponds to 30 h^{-1}Mpc . The quoted results from the selection cuts are reported in Table 4.4. The severity of the cuts applied has repercussions on the dispersion of the ellipticities, most notably in the 15 h^{-1}Mpc sized bin. The estimation of the ellipticity seems also quite affected by the different cuts, but, compared with the baseline analysis (without any cuts), there is no significant deviation.

To conclude, it seems that the quantity of interest to stack the voids should be the radius. While the number of particles may have some impact on the measurement, as it translates a certain level of definition of the voids, discarding voids in terms of their radius may be more relevant in this analysis. Indeed, cutting out small radii may relieve our stack from voids subjected to non-linear effects, more likely to be found at small scales regardless of the number of particles defining them.

Considering the radius, we consider three cuts as candidates for the estimation of the ellipticity of the void: $2m.p.s > R_v > \sigma_{R_v}(34-49)$, $\bar{R}_v > R_v > \sigma_{R_v}(42-57)$ and$\tilde{}R_v > R_v > \sigma_{R_v}(40-55)$. While they tend to discard a lot of the statistic, they seem to be the closest to the expected Alcokc-Paczynski distortion value.

Table 4.4: Same as Table 4.3 but in the case of a binning approach on the radius of the stacked voids. The first column R_v cut corresponds to the adopted cuts, N_p gives the minimum and maximum values of the particles defining the voids. The quoted error for the ellipticity corresponds to the 1σ dispersion in the mocks. As an indication $\bar{R}_v = 42$, $\tilde{R}_v = 40$,$2mps = 34$ and $\sigma_r = 15$, in h^{-1}Mpc , evaluated with the average over N_{mocks}.

R_v^{cut}	N_v^{cut}	$N_p^{\min} - N_p^{\max}$	$\langle\epsilon\rangle$	$\langle\epsilon\rangle$ - $\langle\epsilon_{\text{base}}\rangle$	$\Delta\epsilon$
$2mps - \sigma_{R_v}$	72%	$10 - 259$	0.9993 ± 0.0087	0.0089 ± 0.0098	0.04%
$2mps - 2\sigma_{R_v}$	46%	$10 - 472$	0.9955 ± 0.0052	0.0051 ± 0.0069	-0.34%
$\bar{R}_v - \sigma_{R_v}$	71%	$20 - 369$	0.9952 ± 0.0063	0.0048 ± 0.0074	-0.37%
$\bar{R}_v - 2\sigma_{R_v}$	58%	$20 - 605$	0.9903 ± 0.0059	-0.0001 ± 0.0078	-0.86%
$\tilde{R}_v - \sigma_{R_v}$	54%	$17 - 345$	0.9958 ± 0.0061	0.0054 ± 0.0076	-0.31%
$\tilde{R}_v - 2\sigma_{R_v}$	69%	$17 - 578$	0.9908 ± 0.0057	0.0004 ± 0.0074	-0.81%

4.4.2 Fiducial cosmology matters

While testing the validity of the pipeline, it was shown previously that the ellipticity recovery was well below the percent regardless of the fiducial cosmology used. Although the deviation from the expected ϵ features an increase at larger values of the fiducial cosmology, those are considered not to be dominant in regard to the other systematics of the mock data sample.

Before performing any parameter determination from the reference analysis and picking the cut of interest among the three candidates chosen above, the sensitivity in regard to the change of fiducial cosmology is investigated. The formulation of the AP test implies that a signature deviation should be recovered in respect to a given redshift and a given fiducial cosmology.

Due to several systematic effects, such as the effect of redshift space distortions in the vicinity of the stacked voids, it is expected of the signal to appear flattened to some extent (Cai et al., 2016). Such a flattening was not quite seen in the case of the reference analysis.

4.4.2.1 Baseline analysis

The void finder is run on the same $N_{\mathsf{mocks}} = 100$ used to perform the estimation of the previous section in a wide range of fiducial cosmology $0.1 \leq \Omega_m^{fid} \leq 0.9$. The purpose of such an investigation is to ascertain whether the parameter measured in the baseline analysis is a valid ellipticity measurement in the application of the AP test. If yes, then the pipeline should detect a significant variation of the ellipticity as a function of the fiducial cosmology.

Table 4.5 reports the ellipticities extracted from the pipeline ran on all 100 mocks for nine input fiducial cosmologies in both void finder and stacking procedure. The input ellipticities display no variation at all in regard to the fiducial cosmology.

With the toy simulation, the expected signal was easily recovered with nearly

Ω_m^{fid}	$\langle \epsilon^{baseline} \rangle$
0.1	0.9883 ± 0.0049
0.2	0.9898 ± 0.0050
0.31	$\mathbf{0.9904 \pm 0.0046}$
0.4	0.9906 ± 0.0047
0.5	0.9911 ± 0.0049
0.6	0.9912 ± 0.0046
0.7	0.9914 ± 0.0046
0.8	0.9916 ± 0.0047
0.9	0.9916 ± 0.0049

Table 4.5: Ellipticities in the baseline analysis in regard to the input fiducial cosmology Ω_m^{fid}. The reported values and error are taken as the average ellipticity over all hundred mocks and the error is the 1σ standard-deviation.

negligible effects at large Ω_m^{fid}. In this case, no hint of variation can be seen. With the analysis applied in the previous section, it was shown that the impact of considering the radii and number of particles did not affect the measurement significantly, as such, the use of cuts in this case would not help to discern any AP signal.

However, some aspects were tested with the toy simulation that we re-consider here that may allow us to unearth an AP signal:

- Void centre: The void finder REVOLVER provides two void centre definitions, the barycentre and the circumcentre. In the RSD chapter, it was shown that the circumcentre was more sensitive to the effect of the change in fiducial cosmology.

- Stacking methodology: In our pipeline, the galaxies are rotated so that the z-direction is aligned with the LOS, which corresponds to that of the voids. The $\sigma - \pi$ (or $r_\perp - r_\parallel$) is another way to estimate the positions of the stack particles.

- Radial cut: We've seen in the simulation that constraining the volume considered to measure the ellipticity may bias such a measurement. However, when considering several cuts and varying cosmology, all converge toward the true underlying cosmology.

The effect of the first two can be seen in Fig. 4.16, it is clear that neither the stacking procedure nor the definition of the void centre is at cause in the absence of an Alcock-Paczynski signal. The signal is flat in all the configuration despite the presence of a significant offset from the other two configurations in the case of the void centre definition.

The third item which corresponds to the use of a radial cut to bound the estimation of the ellipticity in a given volume is displayed in Fig. 4.17. The use

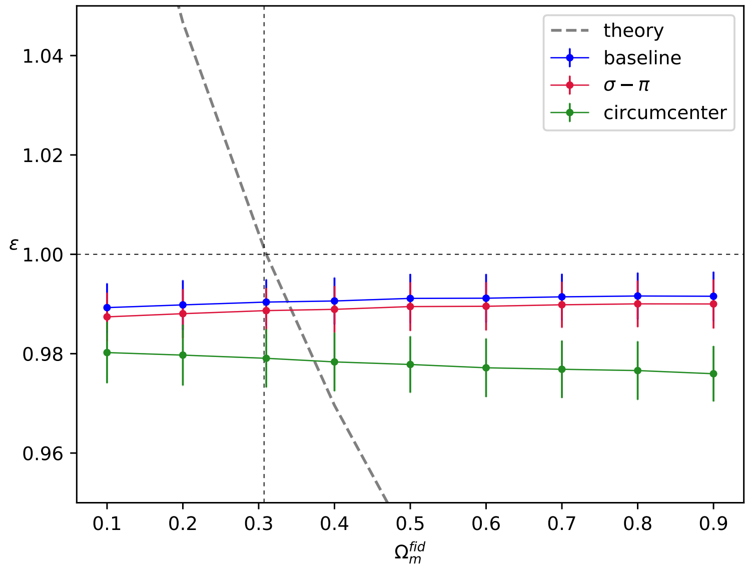


Figure 4.16: Estimated ellipticity from the mean of the mocks for several configurations: the blue full line is the baseline analysis whose values are reported in Table 4.5, in red is the $\sigma - \pi$ stacking method and in green corresponds the pipeline ran with a different void centre definition. The grey dashed line displays the theory for comparison. The quoted errors display the 1σ dispersion of the mocks.

of a radial cut greatly affects the estimation of the ellipticity of the void stack. Nevertheless, no sign of convergence toward a specific ellipticity can be seen. A variation in regard to the fiducial cosmology becomes more noticeable as the radial cut passes the $rcut \leq 1$ threshold. The variation remains of the order of 1.5%, which is too low to consider it likely to be an AP signal.

The consequence of this analysis is that, despite the promise shown by the toy simulation in terms of constraining power, the application of the Alcock-Paczynski test to mocks (and subsequently to data) is not quite possible solely using the real-space theory. The effect of peculiar velocities may be too important for such an analysis.

4.4.2.2 Alcock-Paczynski signal in the literature

While the qualification of both the estimator and the pipeline was thoroughly tested against a toy simulation, a major bias in the present analysis was the assumption that the peculiar velocities in the stack would not significantly influence the measurement. As a consequence, it has to be admitted that while the standard sphere assumption may be true when considering voids in real-space, that may not be the case in redshift-space.

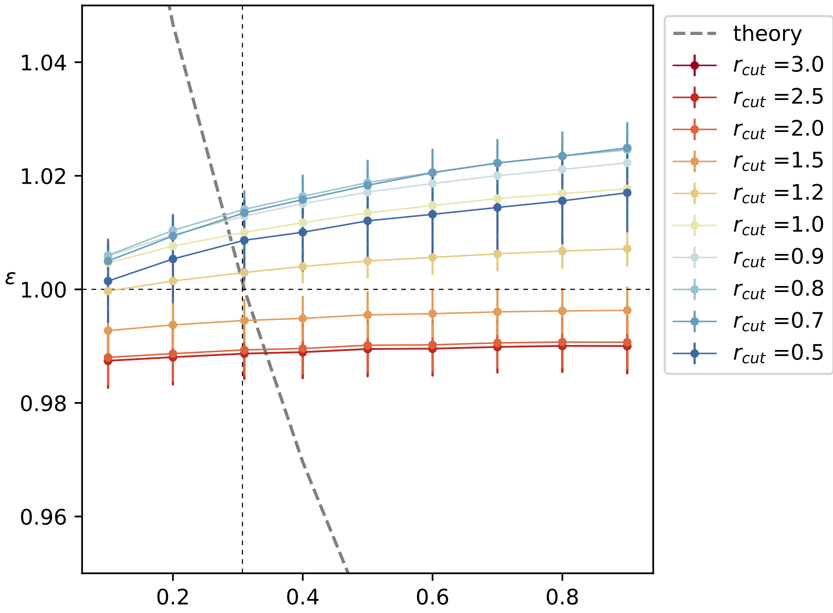


Figure 4.17: Estimated ellipticity from the mean of the mocks for several radial cuts (listed in the legend). The dashed line shows the theoretical expectations. The quoted error corresponds to the 1σ dispersion of the mocks.

Indeed, the elongation/contraction due to both linear and non-linear velocities may change the shape of individual voids. In a similar analysis, Mao et al. (2017) considered the fact that the peculiar velocities may change the shape of the volume. As a result, they use the inertia tensor estimator in concurrence with an added parameter e_{cut}, used to change the volume in which the ellipticity is measured. Provided a set of coordinates (x, y, z) where z corresponds to the LOS, the ellipticity will be considered in a volume of radius:

$$r = \sqrt{x^2 + y^2 + (\frac{z}{e_{cut}})^2}. \tag{4.36}$$

This parameter is varied up on a wide range of values and the ellipticity of the stack is taken to be the value e_{cut} at which the estimated ellipticity through the use of the inertia tensor ϵ is equal to e_{cut}. This enables to probe a variety of ellipsoidal volumes, instead of considering only spherical volumes, in the event where a radial cut is applied.

For comparison to our analysis, we apply the same pipeline on the hundred N_m mocks used previously and display the result in Fig. 4.18. The baseline analysis and that including the e_{cut} shows no significant sensitivity to the AP signal.

As for the baseline, we investigate the impact of the radial cut on the estimation

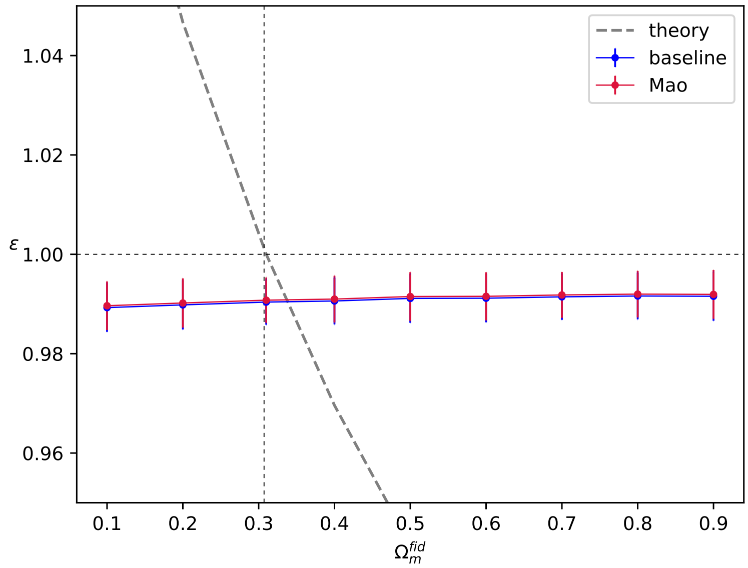


Figure 4.18: Estimation of the ellipticity using the inertia tensor. In blue is the baseline analysis reported in Table 4.5, in red the analysis performed with the prescription of Mao et al. (2017).

of the ellipticity of the stack in concurrence with the e_{cut} parameter. The resulting ellipticities shown in Figure 4.19 display quite a different behaviour depending on the cut. At high r_{cut}, the behaviour is consistent with that found in the baseline analysis. In the range $r_{cut} = 1.2$ to $r_{cut} = 0.9$, the signal appears very noisy which may denote a disappearance of the signal in this range. At low $r_{cut} \leq 0.8$, a significant AP signal is recovered and appears quite flattened, which is consistent with the previous findings of Mao et al. (2017) and Sutter et al. (2014c).

The AP signal is recovered through the use of the inertia tensor in combination with the e_{cut} value. The errors displayed are quite important in comparison to the baseline analysis, probably due to the interpolation scheme used to recover the estimated ellipticity of the void stack. The flattened appearance can be attributed to the peculiar velocities which in the interior of the voids tend to contract the shape along the LOS (Cai et al., 2016). A similar dampening of the ellipticity of the void stack was noted by Endo et al. (2020) when considering the shape in redshift-space, even though their sensitivity to the AP signal in this configuration is surprising as they report no significant cut on the volume considered while measuring the inertia tensor.

However, as it is, it is not so straightforward to constrain the cosmological parameters. Mao et al. (2017) circumvented this problem through the use of mock catalogues in both real and redshift space, applying the same r_{cut} which enabled to map the flattening of the Alcock-Paczynski signal from redshift to real space.

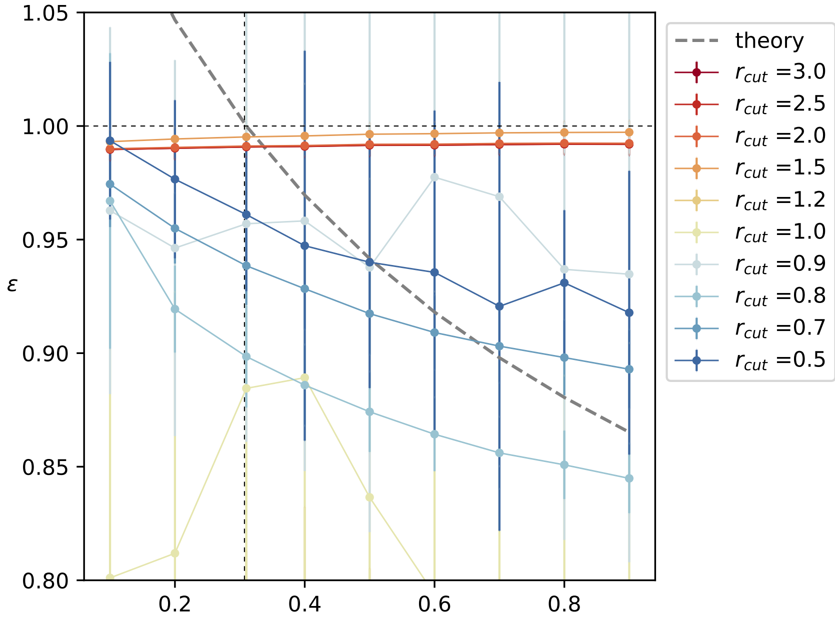


Figure 4.19: Estimation of the ellipticity using the inertia tensor and the e_{cut} parameter. The dashed line represent the theoretical expectations. The errors quoted come from the 1σ deviation over all the mocks. The $r_{cut} = 1.2$ is below the y-axis range and has been cut out for visibility.

Endo et al., 2020 proposed a calibrating function, required to be run on simulation prior to the application of the Alcock-Paczynski test.

The main cause of concern in the application of the AP test lies in its seemingly volume-dependent effect. For one, the disappearance of the signal when considering the entirety of the stack is a surprise which has never been noted before. The volume dependence at low radii can be understood as the impact of the peculiar velocities in the vicinity of the void that, through linear theory, are directly dependent in the volume-averaged density profile. It could be that changing the considered volume of the void stack may change the intensity of the effect of the peculiar velocities, leading to a different amount of flattening. But this effect has to be further investigated before extracting a proper cosmological constraint.

While the Alcock-Paczynski test on voids was hailed as a groundbreaking cosmological probe (Lavaux & Wandelt, 2012), it appears that its application is still crippled by the existence of peculiar velocities in the vicinity of the void. A possible way to get around the lack of modelling of the redshift-space Alcock-Paczynski test could be the use of the two-point correlation function between voids and galaxies. The latter can be considered as a stacking methodology, and, in combination with a modelling of the redshift-space distortions around voids, could allow disentangling the AP effect from the RSDs.

Conclusion

In recent years with the increase of statistics of both redshift and photometric surveys, voids have been pushed into the headlights as promising probes of cosmology. The analysis of their defining properties such as their shape, size and distribution of matter around them have shown to be sensitive to several key questions of the cosmological context. Among them, the main mystery lies in the existence of a dark sector, composed of dark matter and dark energy, the latter responsible for the late-time acceleration of the expansion of the Universe. Voids have been considered to be interesting environments to probe the late-time cosmic acceleration due to their nature of under-dense objects. They should thus be dominated by this component of the Universe which appears to be everything but massive.

This thesis aimed to investigate the potential of voids to constrain the nature of the late-time cosmic acceleration. This nature is of two forms: a time-varying dark energy component that would remain within the context of GR and a possible modification of the laws of gravity formulated in Einstein equations.

The first aspect of my work consisted in testing General Relativity in the framework of the eBOSS collaboration. To this end, I extracted voids from three galaxy samples: Emission Line Galaxies, Luminous Red Galaxies and Quasi-Stellar Objects at three different epochs using void finding algorithms. These void finding algorithms were compared in order to validate the void finding procedure prior to its application on data. This study allowed to define selection cuts to mitigate the edge contamination brought by the void finding algorithm itself. The voids found were then used to measure the void-galaxy cross-correlation function in order to study the redshift-space distortions (RSD) and extract a constraint on the RSD parameter β. Through the use of mocks, I estimated the impact of several methodology choices, as well as the RSD model used, on the accuracy of the measurements. The subsequent systematic effects evaluated proved to be dominant when considering the RSD model. While this cast doubts on its validity, its formulation remains, for now, the best option when considering voids found in redshift space. Nevertheless, such strong systematic effects show that a lot of systematics have yet to be properly identified and mitigated in the context of void science.

In a second analysis, I investigated the application of the Alcock-Paczynski (AP) test on void stacks. Voids, due to the cosmological principle, are assumed to be spherical in average. This statistical property enables to consider voids stacks as standard spheres on which the AP test can be performed by measuring the axis-ratio of the stack. Candidate estimators of this ellipticity were tested against toy simulation to find the most optimal: the inertia tensor. The sensitivity of the estimator was then tested with a more realistic simulation of a distorted void stack, devised to encode some of the void properties extracted with the VIDE and REVOLVER algorithms: the number of galaxies defining a void and a void density profile. The simulated AP effect was recovered with high accuracy. When confronted with mock data, the devised inertia tensor did not perform well. A modification of its parametrization allowed to recover the AP signal revealing a volume-dependent behaviour when applying the analysis in redshift space. Such behaviour needs to be properly modelled and understood prior to the application of the Alcock-Paczynski test without calibration methods. It also remains to be seen if the signal recovery displays a volume dependence in real-space as well.

The work carried out in this thesis, probing dynamical and geometrical distortions around voids emphasised several limitations of the void science. The main limitation can be found in the lack of theoretical modelling of the void two-point statistics. Such a model would enable both to gain a strong constraining power on cosmology but also to investigate the validity of the assumptions adopted when analyzing the void clustering statistics such as the linearity of the bias and its estimation.

A second limitation consists in the estimation of the systematics pertaining to voids. While most void works have been carried out on N-Body simulations, the study of the void-galaxy correlation has been extensively studied in real conditions. However, the impact of the void finding process and the systematics of the galaxies on parameter estimation has yet to be properly studied while it is paramount to an unbiased cosmological constraint. I consider the work presented in this thesis as a first step toward such characterization and plan to extend this further with, for example, the improvement void stack. With more realistic features, it could provide a test ground to ascertain systematical effects.

With the coming large-scale spectroscopic surveys such as DESI and Euclid, tens of thousands of voids are expected to be found. It follows that the analysis techniques have to be refined and the systematic to be mastered to allow voids to reach their full potential as discriminating cosmological probes.

Résumé Substantiel

La cosmologie moderne

La cosmologie porte sur l'étude des lois qui gouvernent l'Univers que nous habitons dans le but de comprendre son origine et son évolution. A l'heure actuelle, notre description de l'Univers est guidée par un modèle standard de la cosmologie. Ce dernier repose sur des fondements théoriques et observationnels.

Un Univers en expansion

Les fondements du modèle standard de la cosmologie reposent sur un cadre théorique fourni par la Relativité Générale formulée par Albert Einstein. Celle-ci est contenue dans un ensemble d'équations qui lient les contributions énergétiques de notre Univers à sa géométrie. Afin de retrouver la formulation de notre modèle standard, deux fondements sont nécessaires :

- Notre Univers obéit au principe cosmologique. Ce dernier énonce qu'il n'y a ni position ni direction privilégiée dans l'Univers. Il est donc homogène et isotrope.
- Notre Univers est en expansion, comme l'ont découvert Vesto Slipher, Edwin Hubble et George Lemaître dans la première moitié du XX$^{\text{e}}$ siècle.

Ces deux fondements peuvent être encodés dans la définition de la métrique de notre Univers et permettent la résolution des équations d'Einstein. L'évolution de l'Univers peut alors être décrite par un ensemble de paramètres qui décrivent la contribution énergétique de l'Univers ainsi que leur expansion.

Le modèle Λ-CDM

L'observation de plusieurs phénomènes astrophysiques comme les vitesses de rotation des galaxies et les supernovae de type Ia ont permis des avancées majeures dans l'inventaire du contenu de l'Univers. Il est ainsi apparu que notre Univers était dominé par deux substances : la matière noire, une substance massive mais invisible à nos détecteurs et l'énergie noire qui serait responsable de l'accélération de l'expansion de l'Univers qui est subie actuellement.

Les observations les plus récentes de plusieurs sondes cosmologiques ont convergé vers une description de l'Univers avec le contenu suivant :

$$\Omega_m = 0.3111 \pm 0.0056 \quad (4.37)$$
$$\Omega_\Lambda = 0.6889 \pm 0.0056 \quad (4.38)$$
$$\Omega_k = 0.001 \pm 0.002. \quad (4.39)$$

L'Univers peut donc être considéré comme plat et dominé par le secteur sombre. L'existence de ce secteur sombre reste néanmoins un mystère non-résolu dans la cosmologie. Si des réponses peuvent être cherchées du côté de la physique des particules dans le cas de la matière noire, la nature de l'énergie noire est considérée comme une conséquence cosmologique. Il pourrait s'agir d'une quantité évoluant dans le temps tout en respectant le cadre donné par la Relativité Générale : l'énergie noire dynamique ou bien, elle pourrait être une conséquence d'une modification des lois de la gravité à grande échelle. Un des enjeux majeurs de la cosmologie moderne est donc de caractériser la nature de l'énergie noire.

L'apport des structure cosmiques

Les structures cosmiques sont une preuve de la particularité de la distribution de matière à grande échelle. Cette dernière ressemble à une toile cosmique formée de filaments, de noeuds, de murs et de vides cosmiques. Afin de comprendre notre Univers et plus spécifiquement, l'énergie noire, l'observation et la modélisation de la toile cosmique est primordiale car elle contient à la fois des informations portant sur les paramètres cosmologiques mais aussi, des possibles traces d'une modification des lois de la gravité.

Un cadre peut être donné dans la théorie linéaire en considérant que cette toile prend sa source dans les fluctuations de densité de l'Univers jeune. La croissance des structures peut être alors comprise comme une conséquence des processus concurrent que sont l'intéraction gravitationnelle et l'expansion de l'Univers au fil du temps.

L'information sur la croissance des structures est encodée dans le taux de croissance des structure :

$$f := \frac{d \ln D_+}{d \ln a}, \qquad f \approx \Omega_m(a)^\gamma \quad (4.40)$$

Ce dernier est donc la dérivée logarithmique de la fonction de croissance D_+ en fonction du facteur d'échelle de l'Univers a et nous donne le taux auquel les structures grandissent. Il est aussi paramétré simplement par l'équation située à droite. Dans ce cas, l'indice γ est prédit dans le contexte de la Relativité Générale avec une valeur de 0.545. La mesure du taux de croissance des structures est donc un test de la validité de la théorie d'Einstein et permettrait de trouver des réponses sur l'origine de l'accélération de l'expansion de l'Univers.

Les vides cosmiques dans la structure à grande échelle

C'est à la fin des années soixante-dix que la distribution spécifique de la matière à grande échelle a été découverte. Celle-ci a été accompagnée de l'observation de vastes régions dépourvues de galaxies : les vides cosmiques. Dès lors, un des enjeu de la cosmologie a été de comprendre comment ces grandes étendues de vides se sont formées ainsi que les processus qui leur ont permis d'atteindre des tailles si importantes.

L'existence de ces vides cosmiques provoque un changement de paradigme sur la vision de la distribution de la matière à grande échelle et remet en cause les modèles de formation des structures déjà observables à l'époque comme les amas ou groupes de galaxies. Elle mène aussi au développement des premiers relevés spectroscopiques de galaxies qui sont maintenant au premier plan de la cosmologie actuelle.

Extraction et définition

Le développement des relevés de galaxies a permis de tracer la toile cosmique avec une précision croissante et de mettre en lumière les zones sous-denses de notre univers.

Pour ce faire, il est nécessaire de développer des techniques d'extraction de vides dans les échantillons de galaxies car les vides, étant dépourvus de matière, sont des objets qui ne sont pas observables de manière directe.

Ces algorithmes sont généralement appelés void finder[2]. La pluralité de ces techniques se rassemblent autour de trois aspects fondamentaux: premièrement, l'estimation de champs de densité, deuxièmement, l'identification des zones de sous-densités et troisièmement, la caractérisation de ces sous-densités en tant que vides cosmiques.

Ces algorithmes ont été développés avec différentes définitions du vide cosmique et peuvent être classés dans trois catégories. Tout d'abord les algorithmes à critère de densité qui entre en jeu dans la définition du vide, une sous-densité n'étant considérée comme un vide que si son niveau de sous-densité est inférieure à un seuil donné. Ensuite, les algorithmes géométriques, dont la construction de la sous-densité repose sur des transformations géométriques comme la tessellation de Voronoi ou bien une partition de l'espace comme des sphères ou des grilles. Finalement, des algorithmes dynamiques qui reconstruisent et identifient les sous-densités à partir du champ de vitesse.

Chaque type d'algorithme a ses avantages et inconvénients, en fonction du type de données, de la géométrie du sondage ou de l'utilisation cosmologique que l'on souhaite en faire. Plusieurs études comparatives de void finder ont été menées,

[2] en français : trouveur de vides

elles confirment que les vides trouvés sont localisés autour des mêmes régions (sous-denses), même si leur forme, centre et taille peuvent être différents. Il en résulte que la définition des vides reste intimement liée à l'algorithme utilisé pour le trouver.

Durant ma thèse, deux algorithmes ont été utilisés pour trouver des vides: VIDE et REVOLVER. Ces deux algorithmes sont basés sur l'algorithme ZOBOV (ZOne Bordering on Voidness), qui se compose de deux étapes. Dans un premier temps, la tessellation de Voronoi permet de reconstruire le champ de densité localement en construisant une cellule autour de chaque galaxie. L'inverse du volume de cette cellule représente une estimation locale de la densité. En rassemblant les cellules adjacentes les plus volumineuses, il est donc possible de construire une première approximation des sous-densités locales du champ de matière tracé par les galaxies. Dans un second temps, la transformée de Watershed est appliquée ce qui équivaut à mesurer les différents niveaux de sous-densités permettant de rassembler les zones et de définir ainsi les vides cosmiques.

Les algorithmes VIDE et REVOLVER présentent l'avantage d'être adaptés aux échantillons de données qui présentent une géométrie généralement complexe en vertu des masques et sélections appliquées. Les contours des relevés sont non-triviaux et sont pris en compte avant d'appliquer la tessellation et transformation de Watershed afin de garantir l'extraction de vides à l'intérieur du relevé.

A l'issue de ces deux procédés, il est alors possible de définir les vides et d'en extraire les propriétés comme la position de leur centre, leur taille et leur forme.

Les vides cosmiques

Les vides cosmiques se caractérisent par leur qualité de zones sous-denses. Etant des objets occupant un volume très important dans les structures à grande échelle, ils sont définis par certaines propriétés.

Croissance des vides

Comme pour le reste des grandes structures, l'origine des vides est attribuée à l'existence de fluctuations de densité dans l'Univers primordial. Ces sous-densités auraient crû pour devenir les vides cosmiques observés aujourd'hui. Au fil de l'expansion, les dépressions dans le champ de densité deviennent de moins en moins denses et s'étendent. Cela a pour conséquence de générer un profil de la répartition de la matière autour du vide, le profil de densité du vide, caractéristique des vides cosmiques. Marqué par une sous-densité minimale en son centre, le vide présente un profil de densité qui croît jusqu'à former un mur avant de diminuer vers la densité moyenne de l'Univers.

Les vides cosmiques participent au cours de leur croissance à celle de la structure à grande échelle qui les abrite. Ils se forment de manière hiérarchique. Les grands vides sont attribués à la fusion de proto-vides et leur expansion concurrente est considérée comme la source des murs se formant entre les vides.

Propriétés

Au-delà de leur nature sous-dense, les vides cosmiques se caractérisent par certaines de leur propriétés. Ces dernières évoluent dans le temps de concert avec l'expansion. Deux propriétés principales sont utilisées pour définir les vides : leur taille et leur forme.

La taille des vides est généralement indiquée par convention comme un rayon. En réalité, il sous-entend le volume enclos dans la sous-densité, en se plaçant dans le cadre où le volume est assimilé à celui d'une sphère. La taille des vides a toujours représenté un intérêt majeur en cosmologie car selon les tailles atteintes aux différentes époques, il est alors possible de discriminer entre différents modèles cosmologiques.

La forme des vides représente également un intérêt en cosmologie. Dans les premières études des vides cosmiques, il a été considéré que les vides évolueraient au cours du temps vers un état de sphéricité. En réalité, les vides ne sont pas tout à fait sphériques et, bien que l'estimation de leur forme soit dépendante de l'algorithme utilisé pour définir les vides, il a été montré que leur forme pouvait dépendre de la cosmologie.

Les vides et la cosmologie

Bien que les propriétés des vides puissent être utilisées pour contraindre la cosmologie, les vides cosmiques peuvent être analysés sous plusieurs aspects afin d'obtenir des informations sur les problématiques de la cosmologie actuelle comme l'énergie noire. De par leur nature, ils sont considérés comme des environnements idéaux pour étudier les propriétés de l'énergie noire.

Ainsi, les vides sont aussi sensibles à l'étude du *clustering*[3] des galaxies autour d'eux, mais peuvent aussi être analysés au regard des effets de lentillages gravitationnels ou en considérant des effets de distorsions géométriques, connus sous le nom d'effet Alcock-Paczynski dont il sera parlé plus en détail plus loin.

Enfin les contraintes cosmologiques avec les vides cosmiques sont complémentaires aux techniques standard utilisées en cosmologie avec les galaxies et peuvent ainsi aider à briser des dégénérescences entre les paramètres cosmologiques.

Contraindre la cosmologie avec les distorsions dynamiques des vides

Les vides cosmiques étant sensibles à l'effet de la dynamique des galaxies environnantes, ils permettent de contraindre le taux de croissance des structures. Au

[3] Le *clustering* est la manière dont les objets se rassemblent du fait de leur propriété massive

cours de ma thèse, j'ai étudié les vides cosmiques issus des échantillons de galaxies observées par le relevé *extended Baryon Oscillation Spectroscopic Survey* (eBOSS) dans le cadre du relevé *Sloan Digital Sky Survey* (SDSS).

Sloan Digital Sky Survey : Cartographier l'Univers

Le relevé SDSS a démarré sous l'égide de James Gunn afin de relever les positions des galaxies de manière systématique et de cartographier la distribution de la matière dans la structure à grande échelle. Depuis le début des années 2000, le relevé a connu plusieurs phases qui ont largement contribué à notre compréhension actuelle de la cosmologie. Les premières étapes ont permis la détection de l'empreinte des oscillations acoustiques de baryons dans la statistique à deux points des galaxies. Puis succéda le projet Baryonic Oscillation Spectroscopic survey (BOSS) dédié à la mesure précise de cet effet dans des échantillons de galaxies lumineuses rouges permettant d'atteindre une contrainte fine des paramètres cosmologiques. Finalement, la dernière étape de ce relevé dédié à la cosmologie a été le sondage eBOSS démarré en 2014 et achevé en 2020 avec la publication des dernières données (la Data Release) DR16.

Le relevé eBOSS

Durant ces cinq années ont été relevés des milliers de spectres pour quatre populations d'objets différentes. Les galaxies comptent trois populations: les Galaxies Lumineuses Rouges (LRG), les Galaxies à Raies d'Emissions (ELG) et les objects quasi-stellaires (QSO) ou quasars. Ces échantillons sont situés à des redshifts de $z_{LRG} = 0.6$, $z_{ELG} = 0.8$ et $z_{QSO} = 1.45$ en moyenne et possèdent des propriétés différentes. Les LRG sont des objets plus massifs et particulièrement nombreux à bas redshift. Les ELG sont des galaxies très actives du point de vue de la formation d'étoiles en leur sein et sont très nombreuses aux redshifts supérieurs à 1. Enfin, les QSOs sont les objets les plus lointains et se répartissent en deux populations au sein du relevé eBOSS, les quasars destinés au *clustering* et les quasars destinés à l'étude de la forêt Lyman-α.

Les échantillons d'intérêt pour l'étude des vides sont les distributions discrètes des galaxies. Leurs positions sont utilisées pour extraire les vides. Avec celles-ci sont aussi fournis des poids pour pondérer les systématiques observationnelles, qu'elles soient d'origine astronomique avec le poids photométrique w_{sys}, instrumentale avec le poids dû aux collisions de fibres w_{cp}, ou encore issue de la détermination du redshift des galaxies w_{noz}. Enfin, un poids est attribué aux galaxies à partir de leur densité en redshift afin d'optimiser l'estimation des statistiques à deux points dans le cadre de la mesure des oscillations acoustiques de baryons. En plus des échantillons de données sont fournis des *mocks* qui reproduisent les propriétés statistiques des différents échantillons : les EZmocks dans le but d'estimer la variance de la statistique à deux points. Un second type de *mocks* a aussi été développé à partir de simulations N-corps dans le but de caractériser les analyses appliquées aux données.

Extraction des vides dans eBOSS

A partir des échantillons de galaxies sont extraits les vides cosmiques. Une première étude a été menée pour comparer les vides trouvés avec les algorithmes VIDE et REVOLVER. A partir de 500 mocks ont été extraits des vides afin de qualifier l'impact de la génération de fausses particules qui enclosent le volume du relevé. Ces dernières sont primordiales pour estimer le volume du relevé et *a fortiori*, le volume des cellules de Voronoi. Les deux algorithmes ne génèrent pas ces particules de la même manière. Une étude comparative de ces deux algorithmes permet alors de procéder à des choix de sélection des vides.

L'impact majoritaire étudié est la disposition des fausses particules générées par les void finders autour des bornes en redshift qui sont de 0.55 à 1.055 dans les mocks considérés. Deux choix ont été étudiés, une distribution plus large que les bornes en redshift entre 0.4 et 1.2 appelé WIDE et une distribution étroite ayant les mêmes bornes que l'échantillon 0.55 et 1.05 appelée NARROW. Cette étude a révélé l'impact non négligeable de la disposition de ces fausses particules sur l'estimation des volumes et des positions des galaxies. Une distribution WIDE a tendance à mener à une contamination des vides jusqu'au coeur de l'échantillon tandis qu'une distribution NARROW entraine une contamination des vides situés près des bornes en redshift seulement. La contamination se trouve généralement dans les petits vides dans le cas NARROW, tandis que dans le cas WIDE, il peut aussi être noté une propension à trouver des vides plus grands.

Au-delà du comportement différent entre VIDE et REVOLVER face à la disposition des fausses particules, les vides trouvés au coeur de l'échantillon dans le cas NARROW sont similaires, avec une propension pour l'algorithme VIDE à trouver des petits vides supplémentaires.

Afin d'enlever la possible contamination sur les bornes en redshift de la distribution de vides, quatre coupures ont été étudiées dans les deux configurations WIDE et NARROW afin de couper les vides dont le volume dépasse les bornes en redshifts. La première configuration s'est montrée plus sensible aux coupures que la seconde. Parmis les quatres coupures étudiées, celle selectionnée correspond au meilleur compromis entre l'atténuation de la contamination et le nombre de vides coupés.

L'algorithme choisit pour extraire les vides dans les échantillons DR16 est REVOLVER. Bien que VIDE et REVOLVER soient semblables, ce dernier présente l'avantage de prendre en compte les poids de correction des galaxies. Des choix supplémentaires de sélections ont été opérés, notamment sur le nombre de particules ainsi que la proximité des vides par rapports aux bords du survey (angulaire également).

Les vides ont été extraits dans chacun des échantillons de galaxies : ELG, LRG et QSO ainsi que leur mocks respectif. Les propriétés des vides et échantillons de

Table 4.6: Propriétés des catalogues de vides identifiés dans les catalogues eBOSS DR16 et leur EZmocks associés. N_g représente le nombre de galaxies, N_v et $N_{v,cut}$ sont le nombre moyen de vides extraits avant et après coupure respectivement accompagné de leur écart-type calculés avec les 1000 mocks et catalogues de données. z_{eff} est le redshift effectif du catalogue de vides après sélection et s_{max} est la séparation maximale entre une paire vide-galaxie utilisée dans l'estimation de la fonction de corrélation.

Echantillon	N_g	N_v	$N_{v,cut}$	z_{eff}	s_{max}	Aire (deg^2)
EZmocks						
ELG	173,736	$2,210 \pm 35$	$1,895 \pm 37$	0.847	3.60	1,170
LRGpCMASS	380,190	$4,305 \pm 54$	$2,850 \pm 47$	0.740	3.52	9,493
QSO	343,700	$5,449 \pm 53$	$4,321 \pm 52$	1.478	3.52	4,808
Data sample						
ELG	173,736	$2,097 \pm 5$	$1,801 \pm 5$	0.847	3.60	1,170
LRGpCMASS	377,458	$4,228 \pm 11$	$2,814 \pm 12$	0.740	3.52	9,493
QSO	343,708	$5,451 \pm 8$	$4,347 \pm 9$	1.478	3.52	4,808

galaxies sont résumés dans le tableau 4.6

Distorsions dans l'espace des redshifts et modélisation

Les vides et les galaxies sont soumis aux effets dynamiques à l'origine du processus de formation et de croissance des structures. De ce fait, la position des galaxies en redshift se retrouve contaminée par ces effets dynamiques, introduisant un décalage supplémentaire dans le redshift. Ces effets se manifestent par l'introduction de distorsions du champ de matière mesurées, aussi connues sous le nom de *redshift space distortions* (RSD).

Ces effets dynamiques dépendent de l'échelle sondée et sont généralement catégorisées en deux contributions :

- Linéaire : La contribution linéaire des RSD est attribuée au processus au coeur de la formation des structure. Elle correspond à des vitesses dirigées vers le centre des structures, dans le cas de structures sur-denses et des vitesses dirigées vers l'extérieur des structures dans le cas des sous-densité comme les vides cosmiques. Ces vitesses sont une conséquence du processus de croissance des structures. Elles modifient l'aspect visuel des structures qui peuvent apparaître comme allongée ou contractée le long de la ligne de visée : l'effet Kaiser.

- Non-Linéaire : Les effets non-linéaires interviennent généralement dans les processus d'effondrement non-linéaire des sur-densités. Cela introduit des déformations à plus petites échelles sous formes d'élongations, aussi connues sous le nom d'effet *Finger-of-god*.

Les vides ont l'avantage d'être considérés comme des objets plus linéaires que les galaxies ou autre marqueurs de sur-densité. Par conséquent, les contributions dynamiques principales sont linéaires. Ils sont donc d'intérêt majeur pour la cosmologie, car ces vitesses dépendent du taux de croissance des structures f, prédit dans le contexte de la Relativité Générale.

Dans le cadre des vides, la mesure du taux de croissance des structures se fait grâce à la fonction de corrélation croisée entre les vides et les galaxies qui consiste à compter le nombre de paires vides-galaxies à une séparation s dans un volume donné. En considérant une conservation du nombre de paires vides-galaxies entre l'espace réel et l'espace des redshifts, il est possible de modéliser la fonction de corrélation théorique en faisant intervenir le taux de croissance des structures :

$$\xi(s,\mu) = \xi(r) + \frac{\beta}{3}\bar{\xi}(r) + f\mu^2(\xi(r) - \bar{\xi}(r)). \tag{4.41}$$

Dans ce cas, la relation entre le champ vide-matière estimé $\xi(r)$ et le vrai champ vide-matière sous-jacent est considérée comme étant linéaire. La décomposition de cette fonction en polynôme de Legendre permet de faire apparaitre une relation simple entre le monopole, qui estime le profil de densité du vide et le quadrupole, qui contient l'information dynamique autour du vide:

$$\begin{aligned} G(\beta) &= \frac{\xi_2^s(r)}{\xi_0^s(r) - \bar{\xi}_0^s(r)} && (4.42) \\ &= \frac{2\beta}{3+\beta}, && (4.43) \end{aligned}$$

où β est f/b, le quotient du taux de croissance des structure et le biais des galaxies. L'analyse porte donc sur l'estimation de ce β et l'extraction d'une contrainte sur le taux de croissance des structures dans les trois échantillons de galaxies : ELG, LRG et QSO.

Etude des effets systématiques

Avant de pratiquer l'estimation du β dans les données, les choix de méthodologies ainsi d'éventuels effets systématiques sont testés sur les EZmocks et les mocks basés sur des simulations à N-corps correspondant à chaque traceur.

Avec les EZmocks sont testés les choix méthodologiques que sont le *binning*[4], l'estimateur de la fonction de corrélation croisée ainsi que l'utilisation des poids FKP. Les deux premiers peuvent influer sur l'amplitude des monopoles et des quadrupoles de la fonction de corrélation et peuvent modifier le signal RSD. Les poids, quant à eux, sont utilisés pour optimiser le signal dans l'analyse standard de la fonction de corrélation à deux points des galaxies. L'utilisation de ces poids dans

[4]Echantillonage choisi pour l'estimation de la fonction de corrélation

le calcul de la fonction de corrélation à deux points entre les vides et les galaxies n'est donc pas optimisé. Le choix a été fait de les utiliser et l'impact de ce choix est étudié.

Avec les simulations à N-corps sont étudiés les effets systématiques liés au modèle utilisé pour contraindre le taux de croissance des structures ainsi que l'impact de la définition du centre du vide et finalement, le choix de la cosmologie fiducielle. Il est apparu que le choix du barycentre pour contraindre le taux de croissance des structures avec le modèle considéré est optimal.

Les déviations entre les valeurs attendues dans les simulations et les paramètres RSD β estimés sont utilisés comme systématiques. Dans le cas des EZmocks, c'est la déviation par rapport à l'analyse de référence qui est reportée comme effet systématiques. Ces effets et leurs intensités sont reportés dans le tableau 4.7.

Table 4.7: Budget systématique total en terme des erreurs relatives sur le paramètre β obtenues à partir des tests dans les mocks pour chacun des traceurs eBOSS. L'erreur systématique totale est la somme quadratique de chaque contribution.

Type	sys in (σ_β/β) (%)	LRG	ELG	QSO
Méthodologie	Binning	4.8	2.3	1.3
	Poids FKP	1.4	2.3	0.7
	Estimateur	2.2	3.3	1.0
Void finder	Cosmologie Fiducielle	2.2	2.2	2.2
Modèle	Modèle RSD	9.0	8.3	39.9
Total (%)		10.8	9.76	40.0

Il apparaît que les effets systématiques les plus importants se trouvent dans l'estimation du paramètre β pour chaque échantillon avec un effet dominant dans le cas de l'échantillon des quasars d'une hauteur de 40%.

Analyse des DR16

Après avoir estimé les différents effets systématiques dans les mocks, la fonction de corrélation vide-galaxie est calculée pour chacun des échantillons considérés. Pour chaque échantillon, mille catalogues de vides sont extraits afin de tenir compte de l'impact des fausses particules générées lors de la recherche des vides. Le paramètre β est alors estimé pour chaque fonction de corrélation et c'est la valeur moyenne des mille β mesurés qui est considérée comme notre mesure pour chaque échantillon. Ces valeurs sont reportées dans le tableau 4.8.

Afin de pouvoir estimer le taux de croissance des structures, il est nécessaire de connaître le biais linaire associés à chaque échantillon de galaxies b. En pratique, ce biais est estimé dans les analyses de l'autocorrelation des galaxies. Le taux de

Table 4.8: Résultats finaux sur le paramètre RSD β estimé dans les différents échantillons eBOSS DR16. β et son erreur statistiques σ_{stat} sont reportés comme étant la moyenne des 1000 estimations de β et ses erreurs dans chaque échantillon. Les erreurs systématiques σ_{syst} correspondent à celles reportées dans la budget systématique total appliqué à notre estimation de β. L'erreur totale σ_{tot} est la somme quadratique des contributions systématique et statistique.

Echantillon	$\langle\beta\rangle$	σ_{stat}	σ_{syst}	σ_{tot}
LRG	0.415	0.075	0.045	0.087
ELG	0.665	0.107	0.065	0.125
QSO	0.313	0.049	0.125	0.134

Table 4.9: Résultats finaux sur l'estimation du taux de croissance des structure dans les vides de eBOSS DR16. Les β et leur erreur sont les moyennes des 1000 estimations de β et leur erreur. Les erreurs cités prennent en compte l'erreur systématique. Les valeurs reportées de $b_1\sigma_8$ proviennent des analyses de *clustering* des échantillons DR16 pour les LRGs (Bautista et al., 2020; Gil-Marín et al., 2020), les ELGs (De Mattia et al., 2020; Tamone et al., 2020) et les QSOs (Hou et al., 2020; Neveux et al., 2020). La contrainte sur le taux de croissance des structures résulte de la combinaison entre le β et $b\sigma_8$ et son erreur inclut l'erreur du biais des galaxies.

Echantillon	z_{eff}	β	$b_1\sigma_8$	$f\sigma_8$
LRG+CMASS	0.740	0.415 ± 0.087	1.20 ± 0.05	0.50 ± 0.11
ELG	0.847	0.665 ± 0.125	0.78 ± 0.05	0.52 ± 0.10
QSO	1.478	0.313 ± 0.134	0.96 ± 0.04	0.30 ± 0.13

croissance des structures est donc obtenu en utilisant les mesures issues de l'analyse de la statistique à deux points de chaque traceur au sein de la collaboration eBOSS. Le paramètre β est alors transformé en $f\sigma_8$ ou f est le taux de croissance des structures et σ_8. Le taux de croissance estimé à partir des vides cosmiques est cité dans le tableau 4.9.

Les contraintes finales du taux de croissance des structures présentées précédemment sont reportées dans la figure 4.20 avec un ensemble non-exhaustif des contraintes sur $f\sigma_8$ réalisées ces dernières années. A l'heure actuelle, la dispersion de ces estimations ne permet pas de mettre en évidence une déviation par rapport au modèle standard de la cosmologie. Les mesures effectuées à travers l'étude des vides restent en accord avec les estimations provenant des techniques d'analyses standard.

Pour conclure cette partie, bien que l'analyse des vides soit prometteuse pour obtenir des contraintes complémentaires du taux de croissance des structures, la présence d'éventuels effets systématiques se montre non-négligeable. Dans le cas des quasars, l'origine d'une déviation si importante par rapport à la valeur attendue

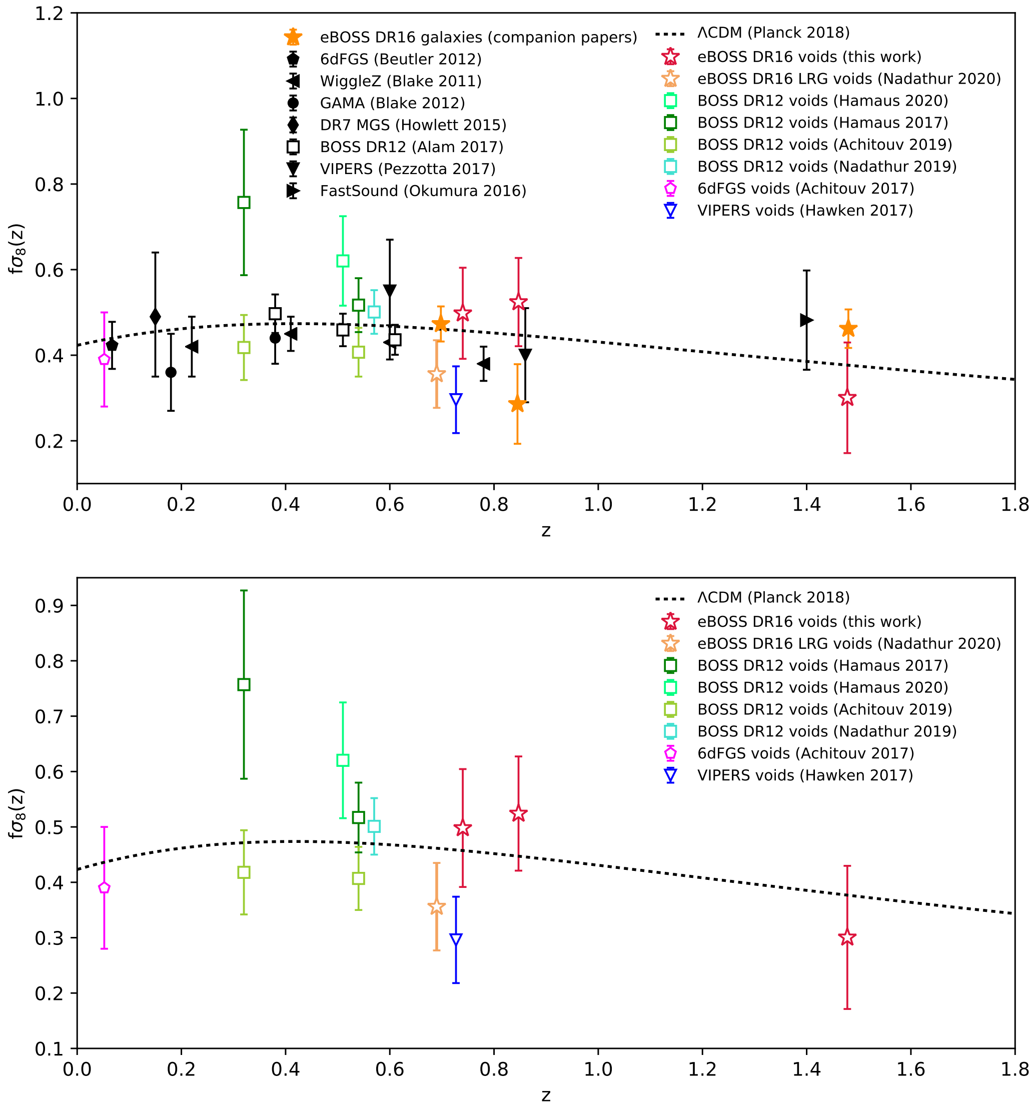


Figure 4.20: Panel supérieur: Comparaison des taux de croissance des structure $f\sigma_8$ estimé dans différents échantillon de galaxies et vides. Les marqueurs noirs pleins représentent les différents relevés, 6dF (Beutler et al., 2012), GAMA (Blake et al., 2013), WiggleZ (Blake et al., 2011a), VIPERS (Pezzotta et al., 2017) et FastSound (Okumura et al., 2016). Les marqueurs vides représentes les contraintes provenant de vides et les marqueur pleins, les galaxies. La ligne pointillée est la prédiction du taux de croissance des structure issue des résultats Planck 2018 Panel inférieur: Comparaison des estimations de taux de croissance des structure $f\sigma_8$ à partir de l'étude de la fonction de corrélation vide-galaxie seulement.

dans la simulation reste incertaine. Elle ne peut néanmoins pas être attribuée au modèle de RSD adopté. Afin d'obtenir les contraintes cosmologiques les plus fines possible, il est donc primordial de mener une étude poussée des effets systématiques.

Contraindre la cosmologie avec les distorsions géométriques des vides

Les RSD représentent une distorsion provenant de la dynamique en jeu dans la croissance des structures et permet de déceler une éventuelle modification des lois de la gravité. Un autre aspect de la contrainte cosmologique de l'énergie noire repose sur l'estimation de son équation d'état. Une manière d'achever cette contrainte est de mesurer les distances des objets qui en dépendent comme dans le cas du test Alcock-Paczynski.

Le test Alcock-Paczynski

Le test Alcock-Paczynski (AP) est un test purement géométrique de la cosmologie. Il repose sur la connaissance de la forme d'un objet possédant des propriétés de symétrie, dont un exemple parfait est une sphère. L'étendue de la sphère peut alors être considérée le long de deux axes majeurs : la ligne de visée et la direction transverse à la ligne de visée qui sont définis comme suit :

$$\Delta r_\perp = D_A(z)\Delta\theta, \tag{4.44}$$

$$\Delta r_\parallel = \frac{c\Delta z}{H(z)}, \tag{4.45}$$

où $D_A(z)$ est la distance de diamètre angulaire, $H(z)$ est le paramètre de Hubble. $\Delta\theta$ correspond à l'étendue angulaire de l'objet considéré et Δz à son étendue le long de la ligne de visée.

La vraie cosmologie sous-jacente de notre Univers n'étant pas connu, il est nécessaire d'utiliser une cosmologie fiducielle pour estimer les distances et les propriétés des vides. Cela a pour conséquence de déformer la relation de symétrie de la sphère : l'effet Alcock-Paczynski.

Cette déformation est alors quantifiable par le rapport des axes de symétrie de la sphère, corrigée par l'usage de la cosmologie fiducielle :

$$\epsilon = \frac{H(z)^{true} D_A(z)^{true}}{H(z)^{fid} D_A(z)^{fid}}. \tag{4.46}$$

ϵ est généralement appelé ellipticité.

Le test Alcock-Paczynski consiste donc simplement à mesurer les axes de symétrie d'un objet dans une cosmologie donnée pour en déduire la cosmologie sous-jacente de notre Univers.

Les vides, bien qu'étant individuellement de formes disparates, peuvent être considérées, lorsqu'ils sont empilés les uns sur les autres en grand nombre, comme étant un objet sphérique. Ils deviennent de cette manière des objets de choix pour la contrainte cosmologique avec le test Alcock-Paczynski.

Détermination du meilleur estimateur

Pour appliquer le test AP, il est primordial de définir un estimateur de l'ellipticité robuste et non biaisé. Pour cela, deux approches ont été testées. La première consiste à considérer un empilement (*stack*) de vides comme étant une ellipsoide du fait de l'effet AP. Une ellipsoide est définie par trois axes de symétrie de valeur a, b et c. En considérant l'axe aligné avec c comme étant colinéaire avec la ligne de visée. L'ellipticité peut être alors définit comme :

$$\epsilon = \sqrt{\frac{2c^2}{a^2 + b^2}} \tag{4.47}$$

Cette ellipticité peut ainsi être mesurée en ajustant les galaxies définissant un stack de vides avec les axes d'une ellipsoides.

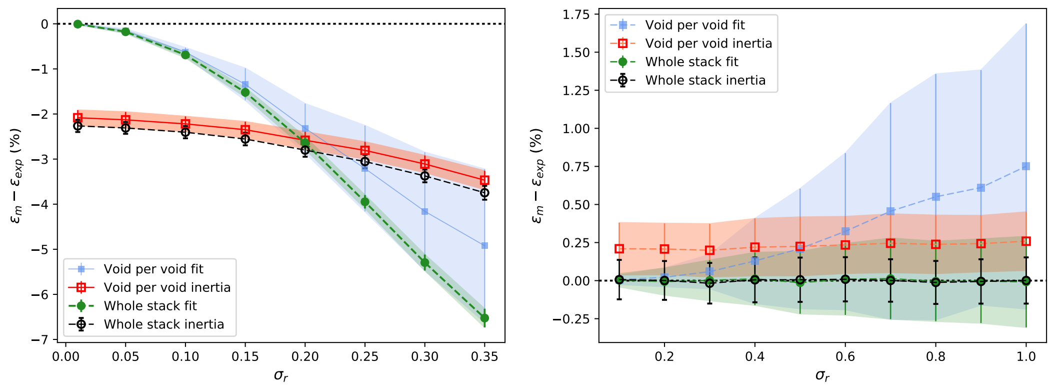
La deuxième approche repose sur l'utilisation du tenseur d'inertie. Ce dernier permet d'estimer les axes préférentiels de symétrie d'un solide. En considérant le stack comme une ellipsoide, l'ellipticité peut alors être déterminée comme une association des moments d'inertie du stack qui se réduit à :

$$\epsilon = \sqrt{\frac{\sum_i 2z_i^2}{\sum_i x_i^2 + y_i^2}}. \tag{4.48}$$

Dans cette expression, (x_i, y_i, z_i) donnent les coordonnées de la ième particules définissant le stack.

Pour tester ces estimateurs, deux simulations de stack ont été développées. Les deux simulations génèrent un nombre fini d'ellipsoide pour reproduire un stack de vides. Les simulations diffèrent dans la disposition des galaxies sur le pourtour des ellipsoides. L'une place les particules sur la surface d'une ellipsoide et impose une épaisseur en introduisant une déviation gaussienne de la position des particules par rapport à la surface de l'ellipsoide. Ce processus, cependant, cause une génération des points non-homogène dans la coquille d'ellipsoide. La seconde méthode repose sur la génération d'une coquille sphérique dont la distribution en volume est gaussienne. Les coordonnées radiales sont ensuite inférées de la distribution de volume et une distorsion selon la coordonnées z est ajoutée. Ce processus permet une population du stack plus homogène.

En testant les performances des deux estimateurs sur ces simulations, c'est le tenseur d'inertie qui semble le plus à même de fournir une estimation de l'ellipticité la moins biaisée, comme on peut le voir dans la figure 4.21.



(a) Inhomogenous ellipticity estimation **(b)** Homogenous ellipticity estimation

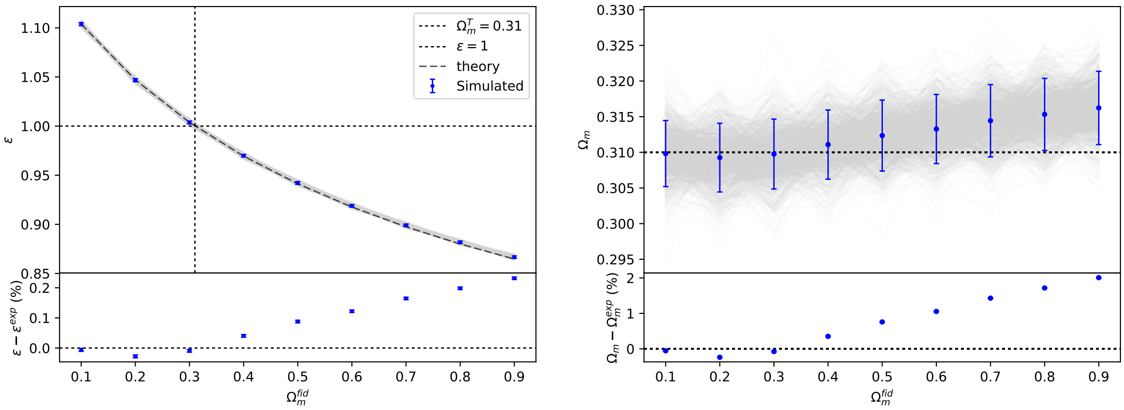
Figure 4.21: Déviation de l'ellipiticé mesurée par rapport à l'ellipticité attendue en fonction de l'épaisseur du stack considéré dans le cas de la simulation non-homogène (à gauche) et homogène (à droite). En bleu est montré l'estimation de l'ellipticité issue de l'ajustement vide par vide de l'équation de l'ellipsoide, en vert le même ajustement appliqué à l'ensemble des particules du stack. Le tenseur d'inertie estimé vide par vide est montré en rouge et le tenseur d'inertie appliqué sur l'ensemble du stack est en noir. Les erreurs affichées correspondent à la dispersion à 1σ estimée à partir de 500 réalisations de chaque simulation.

Sensibilité à l'effet AP

Après avoir fait le choix du tenseur d'inertie comme estimateur de référence pour mesurer l'ellipticité du stack, on procède à la caractérisation de sa sensibilité à l'effet AP.

Pour ce faire, une simulation idéaliste d'un stack de vides est mise en place. Celle-ci repose sur plusieurs hypothèses : le stack en espace réel est sphérique, la distribution de densité suit le profil de densité caractéristique des vides et les vides simulés sont définis par un rayon et un nombre de particules semblables aux vides extrait grâce aux algorithmes VIDE et REVOLVER. Après génération du stack, les positions des vides et des galaxies sont transformées dans l'espace observable (redshift, ascension droit et déclinaison) à partir d'une cosmologie dite 'vraie'. L'effet Alcock-Paczynski est alors introduit par l'utilisation d'une cosmologie différente dans la procédure d'empilement des vides. La Figure 4.22 montre la sensibilité de l'estimateur à l'effet AP. Une déviation du signal peut être notée à l'utilisation d'une cosmologie fiducielle supérieure, elle reste néanmoins bien inférieure au pourcent et négligeable au voisinage de la vraie cosmologie.

La sensibilité à l'effet AP peut être alors testé suivant plusieurs configurations en changeant la méthode d'empilement, en recalculant le centre des vides avec les



(a) Ellipticity **(b)** Cosmologie

Figure 4.22: Haut : Ellipticités (à gauche) et Cosmologie (à droite) mesurées en fonction de la cosmologie fiducielle adoptée. Les régions ombragée montre les valeurs issues des 1000 réalisations du stack. Les points bleus correspondent à la moyenne sur l'ensemble des réalisations. Les erreurs sont prises comme étant l'écart-type des 1000 réalisations. Bas : Déviation des quantités mesurées par rapports aux quantités attendues en fonction de la cosmologie fiducielle attendue.

coordonnées comobile estimées à partir de la cosmologie fiducielles. Dans ces cas-là, la déviation par rapport au signal attendue reste dans le même ordre de grandeur et bien inférieur au pourcent. Le volume considéré pour mesurer l'ellipticité est aussi varié. Dans ce cas, la mesure du signal est grandement affectée mais une convergeance des mesures vers celle de la vraie cosmologie peut être notée.

Application sur des mocks

Après avoir confirmé la sensibilité de notre procédure d'estimation de l'ellipticité d'un stack de vides, celle-ci est appliquée sur 500 mocks. Dans un premier temps, l'impact de la sélection en termes de leur nombre de galaxies et leur taille est étudié. Ces deux grandeurs influent peu sur l'estimation du signal avec des variations inférieures au pourcent.

La sensibilité au signal AP est alors étudiée dans les mocks en extrayant les vides cosmiques en faisant varier la cosmologie fiducielle. Les ellipticités mesurées dans ce cas ne présentent aucune variation en fonction de la cosmologie fiducielle adoptée. Le changement de définition de vide ou de méthode de stacking ne montrent aucune influence sur la récupération du signal AP. Enfin, le changement de volume considéré n'atteste aucune convergence vers une valeur précise de la cosmologie.

Afin de retrouver le signal AP, l'estimateur de l'ellipticité est modifié pour modifier le volume considéré. Au lieu de mesurer l'ellipticité dans un volume sphérique, elle est estimée dans une série de volumes ellipsoidaux. Bien qu'un signal AP soit détecté, il n'est pas en accord avec le signal théorique. De plus, la variation de la valeur du volume considéré influe grandement sur ce signal et change son amplitude. Il s'en suit que le signal Alcock-Paczynski dépend du volume considéré, mais est aussi grandement affecté par les effets dynamiques des distorsions dans l'espace des redshift.

La contrainte cosmologique par le biais du test Alcock-Paczynski seulement nécessite alors une modélisation des effets de distorsions dans l'espace des redshifts. De plus, la dépendance de ce signal en terme du volume considéré doit être comprise.

Conclusion

Pour conclure, ma thèse a porté sur l'étude des vides cosmiques dans le but de contraindre la cosmologie sous deux aspects différents. L'un porte sur l'étude de la dynamique des galaxies autour des vides qui permettraient de trouver des traces d'une éventuelle modification des lois de la gravité, l'autre porte sur l'étude de la forme de stacks de vides qui seraient sensible à l'effet Alcock-Paczynski. Ce dernier permet de contraindre directement les paramètres cosmologiques ainsi que l'équation d'état de l'énergie noire.

La contrainte du taux de croissance des structures grâce aux vides est prometteuse, cependant, des effets systématiques sont à déplorer et nécessite des tests accrus sur la validité des vides et des modèles appliqués. Concernant le test Alcock-Paczynski, au-delà de l'effet des distorsions dans l'espace des redshifts, la dépendance en volume du signal AP représente un second challenge. Ces effets doivent être proprement modélisé et compris avant de pouvoir appliquer une analyse Alcock-Paczynski qui ne repose sur aucune calibration.

Bibliography

Abazajian, K. N., J. K. Adelman-McCarthy, M. A. Agüeros, S. S. Allam, C. Allende Prieto, et al. (2009). The Seventh Data Release of the Sloan Digital Sky Survey. *The Astrophysical Journal Supplement Series* **182**: 543–558. DOI: 10.1088/0067-0049/182/2/543 (cit. on pp. 68, 117).

Abell, G. O. (1958). The Distribution of Rich Clusters of Galaxies. *The Astrophysical Journal Supplement Series* **3**: 211. DOI: 10.1086/190036 (cit. on p. 39).

Abell, G. O. (1961). Evidence regarding second-order clustering of galaxies and interactions between clusters of galaxies. *The Astronomical Journal* **66**: 607. DOI: 10.1086/108472 (cit. on p. 40).

Achitouv, I. (2019). New constraints on the linear growth rate using cosmic voids in the SDSS DR12 datasets. *\prd* **100**:12, 123513. DOI: 10.1103/PhysRevD.100.123513 (cit. on pp. 64, 89, 92, 116, 117, 120).

Achitouv, I., C. Blake, P. Carter, J. Koda, F. Beutler (2017). Consistency of the growth rate in different environments with the 6-degree Field Galaxy Survey: Measurement of the void-galaxy and galaxy-galaxy correlation functions. *\prd* **95**:8, 083502. DOI: 10.1103/PhysRevD.95.083502 (cit. on pp. 64, 89, 118, 119).

Ade, P. A. R., N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, et al. (2014). Planck2013 results. XVI. Cosmological parameters. *\aap* **571**: A16. DOI: 10.1051/0004-6361/201321591 (cit. on p. 74).

Aghanim, N., Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, et al. (2020a). Planck 2018 results - I. Overview and the cosmological legacy of Planck. en. *Astronomy & Astrophysics* **641**: A1. DOI: 10.1051/0004-6361/201833880 (cit. on pp. 25, 27, 36, 63, 116).

Aghanim, N., Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, et al. (2020b). Planck 2018 results - V. CMB power spectra and likelihoods. en. *Astronomy & Astrophysics* **641**: A5. DOI: 10.1051/0004-6361/201936386 (cit. on p. 26).

Aghanim, N., Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, et al. (2020c). Planck 2018 results - VI. Cosmological parameters. en. *Astronomy & Astrophysics* **641**: A6. DOI: 10.1051/0004-6361/201833910 (cit. on pp. 18, 25, 37).

Ahumada, R., C. A. Prieto, A. Almeida, F. Anders, S. F. Anderson, et al. (2020). The 16th Data Release of the Sloan Digital Sky Surveys: First Release from the APOGEE-2 Southern Survey and Full Release of eBOSS Spectra. en. *The Astrophysical Journal Supplement Series* **249**:1, 3. DOI: 10.3847/1538-4365/ab929e (cit. on p. 69).

Aiola, S., E. Calabrese, L. Maurin, S. Naess, B. L. Schmitt, et al. (2020). The Atacama Cosmology Telescope: DR4 Maps and Cosmological Parameters. *arXiv* *e-prints* **2007**: arXiv:2007.07288. URL: http://adsabs.harvard.edu/abs/2020arXiv200707288A (cit. on p. 27).

Alam, S., F. D. Albareti, C. Allende Prieto, F. Anders, S. F. Anderson, et al. (2015). The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III. *ArXiv e-prints* (cit. on p. 69).

Alam, S., M. Ata, S. Bailey, F. Beutler, D. Bizyaev, et al. (2017). The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. *\mnras* **470**: 2617–2652. DOI: 10.1093/mnras/stx721 (cit. on pp. 26, 69, 75, 116, 117).

Alam, S., M. Aubert, S. Avila, C. Balland, J. E. Bautista, et al. (2020a). The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological Implications from two Decades of Spectroscopic Surveys at the Apache Point observatory. en. URL: https://inspirehep.net/literature/1807779 (cit. on p. 69).

Alam, S., A. de Mattia, A. Tamone, S. Ávila, J. A. Peacock, et al. (2020b). The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: N-body Mock Challenge for the eBOSS Emission Line Galaxy Sample. *arXiv e-prints* **2007**: arXiv:2007.09004. URL: http://adsabs.harvard.edu/abs/2020arXiv200709004A (cit. on pp. 26, 74, 103).

Albareti, F. D. et al. (2017). The Thirteenth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-IV Survey MApping Nearby Galaxies at Apache Point Observatory. *\apjs* **233**: 25. URL: https://doi.org/10.3847/1538-4365/aa8992 (cit. on p. 70).

Alcock, C., B. Paczyński (1979). An evolution free test for non-zero cosmological constant. en. *Nature* **281**:5730, 358–359. DOI: 10.1038/281358a0 (cit. on pp. 63, 124, 125).

Alpher, R. A., H. Bethe, G. Gamow (1948). The Origin of Chemical Elements. *Physical Review* **73**: 803–804. DOI: 10.1103/PhysRev.73.803 (cit. on p. 24).

Amendola, L. et al. (2018). Cosmology and fundamental physics with the Euclid satellite. *Living Rev. Rel.* **21**:1, 2. DOI: 10.1007/s41114-017-0010-3 (cit. on pp. 28, 71, 118).

Amendola, L., J. A. Frieman, I. Waga (1999). Weak gravitational lensing by voids. *Monthly Notices of the Royal Astronomical Society* **309**: 465–473. DOI: 10.1046/j.1365-8711.1999.02841.x (cit. on p. 62).

Asaka, T., M. Shaposhnikov (2005). The νMSM, dark matter and baryon asymmetry of the universe. en. *Physics Letters B* **620**:1-2, 17–26. DOI: 10.1016/j.physletb.2005.06.020 (cit. on p. 21).

Aubert, M., M.-C. Cousinou, S. Escoffier, A. J. Hawken, S. Nadathur, et al. (2020). The Completed SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Growth rate of structure measurement from cosmic voids. *arXiv e-prints* **2007**: arXiv:2007.09013. URL: http://adsabs.harvard.edu/abs/2020arXiv200709013A (cit. on pp. 64, 117).

Bahcall, J. N., P. C. Joss (1976). Is the local supercluster a physical association. *The Astrophysical Journal* **203**: 23–32. DOI: 10.1086/154043 (cit. on p. 39).
Baker, T., J. Clampitt, B. Jain, M. Trodden (2018). Void lensing as a test of gravity. *Physical Review D* **98**: 023511. DOI: 10.1103/PhysRevD.98.023511 (cit. on p. 62).
Ballinger, W. E., J. A. Peacock, A. F. Heavens (1996). Measuring the cosmological constant with redshift surveys. *Monthly Notices of the Royal Astronomical Society* **282**: 877. DOI: 10.1093/mnras/282.3.877 (cit. on p. 126).
Bardeen, J. M., J. R. Bond, N. Kaiser, A. S. Szalay (1986). The statistics of peaks of Gaussian random fields. *The Astrophysical Journal* **304**: 15–61. DOI: 10.1086/164143 (cit. on pp. 55, 60).
Bartelmann, M., P. Schneider (2001). Weak gravitational lensing. en. *Physics Reports* **340**:4-5, 291–472. DOI: 10.1016/S0370-1573(00)00082-X (cit. on p. 20).
Bautista, J. E., R. Paviot, M. Vargas Magaña, S. de la Torre, S. Fromenteau, et al. (2020). The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: measurement of the BAO and growth rate of structure of the luminous red galaxy sample from the anisotropic correlation function between redshifts 0.6 and 1. *arXiv e-prints* **2007**: arXiv:2007.08993. URL: http://adsabs.harvard.edu/abs/2020arXiv200708993B (cit. on pp. 27, 71, 96, 115, 117, 177).
Bennett, C. L., D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, et al. (2013). Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. *The Astrophysical Journal Supplement Series* **208**: 20. DOI: 10.1088/0067-0049/208/2/20 (cit. on p. 25).
Bernardeau, F. (2012). *Cosmologie - Des fondements théoriques aux observations: Des fondements théoriques aux observations*. fr. EDP Sciences (cit. on p. 29).
Bertschinger, E. (1985). The self-similar evolution of holes in an Einstein-de Sitter universe. *The Astrophysical Journal Supplement Series* **58**: 1–37. DOI: 10.1086/191027 (cit. on pp. 56, 57, 60).
Beutler, F., C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, et al. (2012). The 6dF Galaxy Survey: z≈ 0 measurements of the growth rate and $\sigma 8$. *\mnras* **423**:4, 3430–3444. DOI: 10.1111/j.1365-2966.2012.21136.x (cit. on pp. 118, 119, 178).
Biswas, R., E. Alizadeh, B. D. Wandelt (2010). Voids as a precision probe of dark energy. en. *Physical Review D* **82**:2, 023002. DOI: 10.1103/PhysRevD.82.023002 (cit. on p. 65).
Blake, C., I. K. Baldry, J. Bland-Hawthorn, L. Christodoulou, M. Colless, et al. (2013). Galaxy And Mass Assembly (GAMA): improved cosmic growth measurements using multiple tracers of large-scale structure. en. *Monthly Notices of the Royal Astronomical Society* **436**:4, 3089–3105. DOI: 10.1093/mnras/stt1791 (cit. on pp. 118, 119, 178).
Blake, C., S. Brough, M. Colless, C. Contreras, W. Couch, et al. (2011a). The WiggleZ Dark Energy Survey: the growth rate of cosmic structure since redshift

z=0.9. *Monthly Notices of the Royal Astronomical Society* **415**: 2876–2891. DOI: 10.1111/j.1365-2966.2011.18903.x (cit. on pp. 118, 119, 178).
Blake, C., E. A. Kazin, F. Beutler, T. M. Davis, D. Parkinson, et al. (2011b). The WiggleZ Dark Energy Survey: mapping the distance-redshift relation with baryon acoustic oscillations. *Monthly Notices of the Royal Astronomical Society* **418**: 1707–1724. DOI: 10.1111/j.1365-2966.2011.19592.x (cit. on p. 126).
Blanton, M. R., M. A. Bershady, B. Abolfathi, F. D. Albareti, C. Allende Prieto, et al. (2017). Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe. *The Astronomical Journal* **154**: 28. DOI: 10.3847/1538-3881/aa7567 (cit. on p. 69).
Blanton, M. R., D. W. Hogg, N. A. Bahcall, J. Brinkmann, M. Britton, et al. (2003). The Galaxy Luminosity Function and Luminosity Density at Redshift z = 0.1*. en. *The Astrophysical Journal* **592**:2, 819. DOI: 10.1086/375776 (cit. on p. 29).
Blumenthal, G. R., L. N. da Costa, D. S. Goldwirth, M. Lecar, T. Piran (1992). The largest possible voids. *The Astrophysical Journal* **388**: 234–241. DOI: 10.1086/171147 (cit. on p. 59).
Blumenthal, G. R., S. M. Faber, J. R. Primack, M. J. Rees (1984). Formation of galaxies and large-scale structure with cold dark matter. en. *Nature* **311**:5986, 517–525. DOI: 10.1038/311517a0 (cit. on p. 21).
Bos, E. G. P., R. van de Weygaert, K. Dolag, V. Pettorino (2012). The darkness that shaped the void: dark energy and cosmic voids. en. *Monthly Notices of the Royal Astronomical Society* **426**:1, 440–461. DOI: 10.1111/j.1365-2966.2012.21478.x (cit. on p. 61).
Bothun, G. (2000). Modern cosmological observations and problems. en. URL: https://inspirehep.net/literature/486371 (cit. on p. 29).
Broadhurst, T. J., R. S. Ellis, D. C. Koo, A. S. Szalay (1990). Large-scale distribution of galaxies at the Galactic poles. en. *Nature* **343**:6260, 726–728. DOI: 10.1038/343726a0 (cit. on p. 26).
Cai, Y.-C., M. Neyrinck, Q. Mao, J. A. Peacock, I. Szapudi, et al. (2017). The lensing and temperature imprints of voids on the cosmic microwave background. *\mnras* **466**: 3364–3375. DOI: 10.1093/mnras/stw3299 (cit. on p. 63).
Cai, Y.-C., A. Taylor, J. A. Peacock, N. Padilla (2016). Redshift-space distortions around voids. en. *Monthly Notices of the Royal Astronomical Society* **462**:3, 2465–2477. DOI: 10.1093/mnras/stw1809 (cit. on pp. 91, 93–95, 158, 162).
Cautun, M., E. Paillas, Y.-C. Cai, S. Bose, J. Armijo, et al. (2018). The Santiago-Harvard-Edinburgh-Durham void comparison - I. SHEDding light on chameleon gravity tests. *Monthly Notices of the Royal Astronomical Society* **476**: 3195–3217. DOI: 10.1093/mnras/sty463 (cit. on pp. 46, 53, 60).
Ceccarelli, L., D. Paz, M. Lares, N. Padilla, D. G. Lambas (2013). Clues on void evolution - I. Large-scale galaxy distributions around voids. *Monthly Notices of the Royal Astronomical Society* **434**: 1435–1442. DOI: 10.1093/mnras/stt1097 (cit. on p. 88).

Chan, K. C., N. Hamaus, V. Desjacques (2014). Large-scale clustering of cosmic voids. *Physical Review D* **90**: 103521. DOI: 10.1103/PhysRevD.90.103521 (cit. on p. 54).

Chan, K. C., Y. Li, M. Biagetti, N. Hamaus (2020). Measurement of Void Bias Using Separate Universe Simulations. *The Astrophysical Journal* **889**: 89. DOI: 10.3847/1538-4357/ab64ec (cit. on p. 54).

Chantavat, T., U. Sawangwit, P. M. Sutter, B. D. Wandelt (2016). Cosmological parameter constraints from CMB lensing with cosmic voids. *Physical Review D* **93**: 043523. DOI: 10.1103/PhysRevD.93.043523 (cit. on p. 63).

Chevallier, M., D. Polarski (2001). Accelerating Universes with Scaling Dark Matter. *International Journal of Modern Physics D* **10**: 213–223. DOI: 10.1142/S0218271801000822 (cit. on p. 23).

Chuang, C.-H., F.-S. Kitaura, Y. Liang, A. Font-Ribera, C. Zhao, et al. (2017). Linear redshift space distortions for cosmic voids based on galaxies in redshift space. *\prd* **95**:6, 063528. DOI: 10.1103/PhysRevD.95.063528 (cit. on pp. 64, 121).

Chuang, C.-H., F.-S. Kitaura, F. Prada, C. Zhao, G. Yepes (2015). EZmocks: extending the Zel'dovich approximation to generate mock galaxy catalogues with accurate clustering statistics. *\mnras* **446**: 2621–2628. DOI: 10.1093/mnras/stu2301 (cit. on p. 74).

Clampitt, J., Y.-C. Cai, B. Li (2013). Voids in modified gravity: excursion set predictions. *Monthly Notices of the Royal Astronomical Society* **431**: 749–766. DOI: 10.1093/mnras/stt219 (cit. on p. 60).

Clampitt, J., B. Jain (2015). Lensing measurements of the mass distribution in SDSS voids. *Monthly Notices of the Royal Astronomical Society* **454**: 3357–3365. DOI: 10.1093/mnras/stv2215 (cit. on pp. 44, 62).

Clampitt, J., B. Jain, C. Sánchez (2016). Clustering and bias measurements of SDSS voids. *Monthly Notices of the Royal Astronomical Society* **456**: 4425–4431. DOI: 10.1093/mnras/stv2933 (cit. on p. 54).

Clowe, D., M. Bradač, A. H. Gonzalez, M. Markevitch, S. W. Randall, et al. (2006). A Direct Empirical Proof of the Existence of Dark Matter. en. *The Astrophysical Journal* **648**:2, L109–L113. DOI: 10.1086/508162 (cit. on pp. 20, 21).

Colberg, J. M., F. Pearce, C. Foster, E. Platen, R. Brunino, et al. (2008). The Aspen-Amsterdam void finder comparison project. *\mnras* **387**:2, 933–944. DOI: 10.1111/j.1365-2966.2008.13307.x (cit. on pp. 44, 45, 47, 53).

Contarini, S., T. Ronconi, F. Marulli, L. Moscardini, A. Veropalumbo, et al. (2019). Cosmological exploitation of the size function of cosmic voids identified in the distribution of biased tracers. *\mnras* **488**:3, 3526–3540. DOI: 10.1093/mnras/stz1989 (cit. on pp. 55, 60).

Corbelli, E., P. Salucci (2000). The extended rotation curve and the dark matter halo of M33. en. *Monthly Notices of the Royal Astronomical Society* **311**:2, 441–447. DOI: 10.1046/j.1365-8711.2000.03075.x (cit. on p. 19).

Courtois, H. M., Y. Hoffman, R. B. Tully, S. Gottlöber (2011). THREE-DIMENSIONAL VELOCITY AND DENSITY RECONSTRUCTIONS OF THE LOCAL UNIVERSE WITH COSMICFLOWS-1. en. *The Astrophysical Journal* **744**:1, 43. DOI: 10.1088/0004-637X/744/1/43 (cit. on p. 44).

Cousinou, M. C., A. Pisani, A. Tilquin, N. Hamaus, A. J. Hawken, et al. (2019). Multivariate analysis of cosmic void characteristics. *Astronomy and Computing* **27**: 53. DOI: 10.1016/j.ascom.2019.03.001 (cit. on pp. 54, 110, 123).

Davies, C. T., M. Cautun, B. Li (2018). Weak lensing by voids in weak lensing maps. *Monthly Notices of the Royal Astronomical Society* **480**: L101–L105. DOI: 10.1093/mnrasl/sly135 (cit. on p. 63).

Davis, M., P. J. E. Peebles (1983). A survey of galaxy redshifts. V - The two-point position and velocity correlations. *\apj* **267**: 465–482 (cit. on p. 92).

Dawson, K. S., D. J. Schlegel, C. P. Ahn, S. F. Anderson, É. Aubourg, et al. (2013). The Baryon Oscillation Spectroscopic Survey of SDSS-III. *\aj* **145**: 10. DOI: 10.1088/0004-6256/145/1/10 (cit. on pp. 69, 72).

Dawson, K. S., J.-P. Kneib, W. J. Percival, S. Alam, F. D. Albareti, et al. (2016). The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data. *The Astronomical Journal* **151**: 44. DOI: 10.3847/0004-6256/151/2/44 (cit. on p. 69).

De Felice, A., S. Tsujikawa (2010). f(R) theories. en. *Living Rev.Rel.* **13**: 3. DOI: 10.12942/lrr-2010-3 (cit. on p. 24).

De Mattia, A. et al. (2020). The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey emission line galaxy sample: measurement of the BAO and growth rate of structure from the anisotropic power spectrum between redshift 0.6 and 1.1. *submitted* (cit. on pp. 96, 115, 117, 177).

De Bernardis, P., P. A. R. Ade, J. J. Bock, J. R. Bond, J. Borrill, et al. (2000). A flat Universe from high-resolution maps of the cosmic microwave background radiation. *Nature* **404**: 955–959. DOI: 10.1038/35010035 (cit. on p. 25).

De Lapparent, V., M. J. Geller, J. P. Huchra (1986). A slice of the universe. *The Astrophysical Journal Letters* **302**: L1–L5. DOI: 10.1086/184625 (cit. on pp. 41, 42).

De Vaucouleurs, G. (1975). Supergalactic studies. I - Supergalactic distribution of the nearest galaxies. *The Astrophysical Journal* **202**: 319–326. DOI: 10.1086/153979 (cit. on p. 39).

De Vaucouleurs, G. (1976). Is the local supercluster a random clumping accident. *The Astrophysical Journal* **203**: 33–38. DOI: 10.1086/154044 (cit. on p. 39).

Deffayet, C., G. Dvali, G. Gabadadze (2002). Accelerated universe from gravity leaking to extra dimensions. en. *Physical Review D* **65**:4, 044023. DOI: 10.1103/PhysRevD.65.044023 (cit. on p. 24).

DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, et al. (2016a). The DESI Experiment Part I: Science,Targeting, and Survey Design. *ArXiv e-prints* (cit. on pp. 28, 71, 118).

DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, et al. (2016b). The DESI Experiment Part II: Instrument Design. *ArXiv e-prints* (cit. on pp. 28, 71, 118).

Dey, A., D. J. Schlegel, D. Lang, R. Blum, K. Burleigh, et al. (2019). Overview of the DESI Legacy Imaging Surveys. *aj* **157**: 168. DOI: 10.3847/1538-3881/ab089d (cit. on p. 71).

Dicke, R. H., P. J. E. Peebles, P. G. Roll, D. T. Wilkinson (1965). Cosmic Black-Body Radiation. *The Astrophysical Journal* **142**: 414–419. DOI: 10.1086/148306 (cit. on p. 24).

Dodelson, S. (2003). Modern Cosmology. en. URL: https://inspirehep.net/literature/640063 (cit. on p. 88).

Dubinski, J., L. N. da Costa, D. S. Goldwirth, M. Lecar, T. Piran (1993). Void evolution and the large-scale structure. *The Astrophysical Journal* **410**: 458–468. DOI: 10.1086/172762 (cit. on pp. 56, 57, 60).

Dvali, G., G. Gabadadze, M. Porrati (2000). 4D gravity on a brane in 5D Minkowski space. en. *Physics Letters B* **485**:1-3, 208–214. DOI: 10.1016/S0370-2693(00)00669-9 (cit. on p. 24).

Eisenstein, D. J., D. H. Weinberg, E. Agol, H. Aihara, C. Allende Prieto, et al. (2011). SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems. *aj* **142**: 72. DOI: 10.1088/0004-6256/142/3/72 (cit. on p. 69).

Eisenstein, D. J., J. Annis, J. E. Gunn, A. S. Szalay, A. J. Connolly, et al. (2001). Spectroscopic Target Selection for the Sloan Digital Sky Survey: The Luminous Red Galaxy Sample. *The Astronomical Journal* **122**: 2267–2280. DOI: 10.1086/323717 (cit. on p. 71).

Eisenstein, D. J., W. Hu (1998). Baryonic Features in the Matter Transfer Function. *The Astrophysical Journal* **496**: 605–614. DOI: 10.1086/305424 (cit. on p. 25).

Eisenstein, D. J., W. Hu, J. Silk, A. S. Szalay (1998). Can Baryonic Features Produce the Observed 100 h-1 Mpc Clustering? en. *The Astrophysical Journal Letters* **494**:1, L1. DOI: 10.1086/311171 (cit. on p. 26).

Eisenstein, D. J., I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton, et al. (2005). Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *The Astrophysical Journal* **633**: 560–574. DOI: 10.1086/466512 (cit. on p. 26).

Elyiv, A., F. Marulli, G. Pollina, M. Baldi, E. Branchini, et al. (2015). Cosmic voids detection without density measurements. *Monthly Notices of the Royal Astronomical Society* **448**: 642–653. DOI: 10.1093/mnras/stv043 (cit. on p. 44).

Endo, T., H. Tashiro, A. J. Nishizawa (2020). The Alcock Paczynski test with voids in 21cm intensity field. *Monthly Notices of the Royal Astronomical Society*. DOI: 10.1093/mnras/staa2822 (cit. on pp. 64, 127, 162, 163).

Falck, B., K. Koyama, G.-B. Zhao, M. Cautun (2018). Using Voids to Unscreen Modified Gravity. *mnras* **475**:3, 3262–3272. DOI: 10.1093/mnras/stx3288 (cit. on p. 60).

Fang, Y., N. Hamaus, B. Jain, S. Pandey, G. Pollina, et al. (2019). Dark Energy Survey year 1 results: the relationship between mass and light around cosmic voids. *Monthly Notices of the Royal Astronomical Society* **490**: 3573–3587. DOI: 10.1093/mnras/stz2805 (cit. on pp. 54, 62).

Feldman, H. A., N. Kaiser, J. A. Peacock (1994). Power-spectrum analysis of three-dimensional redshift surveys. *The Astrophysical Journal* **426**: 23–37. DOI: 10.1086/174036 (cit. on p. 73).

Fillmore, J. A., P. Goldreich (1984). Self-similar spherical voids in an expanding universe. *The Astrophysical Journal* **281**: 9–12. DOI: 10.1086/162071 (cit. on p. 56).

Frenk, C. S., S. D. M. White (2012). Dark matter and cosmic structure. *Annalen der Physik* **524**: 507–534. DOI: 10.1002/andp.201200212 (cit. on p. 20).

Fry, J. N., R. Giovanelli, M. P. Haynes, A. L. Melott, R. J. Scherrer (1989). Void statistics, scaling, and the origins of large-scale structure. *The Astrophysical Journal* **340**: 11–22. DOI: 10.1086/167372 (cit. on p. 42).

Fujimoto, M. (1983). Dynamics of ellipsoidal voids of matter in an expanding universe. *Publications of the Astronomical Society of Japan* **35**: 159–171. URL: http://adsabs.harvard.edu/abs/1983PASJ...35..159F (cit. on p. 56).

Fukuda, Y., T. Hayakawa, E. Ichihara, K. Inoue, K. Ishihara, et al. (1998). Measurements of the Solar Neutrino Flux from Super-Kamiokande's First 300 Days. *Physical Review Letters* **81**:6, 1158–1162. DOI: 10.1103/PhysRevLett.81.1158 (cit. on p. 21).

Fukugita, M., T. Ichikawa, J. E. Gunn, M. Doi, K. Shimasaku, et al. (1996). The Sloan Digital Sky Survey Photometric System. *The Astronomical Journal* **111**: 1748. DOI: 10.1086/117915 (cit. on p. 68).

Gamow, G. (1948). The Origin of Elements and the Separation of Galaxies. *Physical Review* **74**: 505–506. DOI: 10.1103/PhysRev.74.505.2 (cit. on p. 24).

Gil-Marín, H., J. E. Bautista, R. Paviot, M. Vargas-Magaña, S. de la Torre, et al. (2020). The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: measurement of the BAO and growth rate of structure of the luminous red galaxy sample from the anisotropic power spectrum between redshifts 0.6 and 1.0. *arXiv e-prints* **2007**: arXiv:2007.08994. URL: http://adsabs.harvard.edu/abs/2020arXiv200708994G (cit. on pp. 27, 71, 96, 115, 117, 177).

Goldwirth, D. S., L. N. da Costa, R. van de Weygaert (1995). The two-point correlation function and the size of voids. en. *Monthly Notices of the Royal Astronomical Society* **275**:4, 1185–1194. DOI: 10.1093/mnras/275.4.1185 (cit. on p. 60).

Gottlöber, S., E. L. Łokas, A. Klypin, Y. Hoffman (2003). The structure of voids. en. *Monthly Notices of the Royal Astronomical Society* **344**:3, 715–724. DOI: 10.1046/j.1365-8711.2003.06850.x (cit. on p. 44).

Granett, B. R., M. C. Neyrinck, I. Szapudi (2008). An Imprint of Superstructures on the Microwave Background due to the Integrated Sachs-Wolfe Effect. *The Astrophysical Journal Letters* **683**: L99. DOI: 10.1086/591670 (cit. on p. 63).

Gregory, S. A., L. A. Thompson (1978). The Coma/A1367 supercluster and its environs. *The Astrophysical Journal* **222**: 784–799. DOI: 10.1086/156198 (cit. on pp. 40, 41).

Gunn, J. E., M. Carr, C. Rockosi, M. Sekiguchi, K. Berry, et al. (1998). The Sloan Digital Sky Survey Photometric Camera. *The Astronomical Journal* **116**: 3040–3081. DOI: 10.1086/300645 (cit. on p. 68).

Gunn, J. E., W. A. Siegmund, E. J. Mannery, R. E. Owen, C. L. Hull, et al. (2006). The 2.5 m Telescope of the Sloan Digital Sky Survey. *The Astronomical Journal* **131**: 2332–2359. DOI: 10.1086/500975 (cit. on p. 69).

Guzzo, L., M. Scodeggio, B. Garilli, B. R. Granett, A. Fritz, et al. (2014). The VIMOS Public Extragalactic Redshift Survey (VIPERS). An unprecedented view of galaxies and large-scale structure at $0.5 < z < 1.2$. *aap* **566**: A108. DOI: 10.1051/0004-6361/201321489 (cit. on p. 45).

Hahn, O., C. Porciani, C. M. Carollo, A. Dekel (2007). Properties of dark matter haloes in clusters, filaments, sheets and voids. *Monthly Notices of the Royal Astronomical Society* **375**: 489–499. DOI: 10.1111/j.1365-2966.2006.11318.x (cit. on p. 44).

Hamaus, N., M.-C. Cousinou, A. Pisani, M. Aubert, S. Escoffier, et al. (2017). Multipole analysis of redshift-space distortions around cosmic voids. *Journal of Cosmology and Astroparticle Physics* **2017**:07, 014–014. DOI: 10.1088/1475-7516/2017/07/014 (cit. on pp. 64, 89, 93, 102, 116, 117, 120).

Hamaus, N., A. Pisani, J.-A. Choi, G. Lavaux, B. D. Wandelt, et al. (2020). Precision cosmology with voids in the final BOSS data. *arXiv e-prints* **2007**: arXiv:2007.07895. URL: http://adsabs.harvard.edu/abs/2020arXiv200707895H (cit. on pp. 89, 116, 117, 121).

Hamaus, N., A. Pisani, P. M. Sutter, G. Lavaux, S. Escoffier, et al. (2016). Constraints on Cosmology and Gravity from the Dynamics of Voids. *Phys. Rev. Lett.* **117**:9, 091302. DOI: 10.1103/PhysRevLett.117.091302 (cit. on p. 89).

Hamaus, N., P. M. Sutter, G. Lavaux, B. D. Wandelt (2015). Probing cosmology and gravity with redshift-space distortions around voids. *Journal of Cosmology and Astro-Particle Physics* **2015**: 036. DOI: 10.1088/1475-7516/2015/11/036 (cit. on p. 64).

Hamaus, N., P. M. Sutter, B. D. Wandelt (2014a). Universal Density Profile for Cosmic Voids. en. *Physical Review Letters* **112**:25, 251302. DOI: 10.1103/PhysRevLett.112.251302 (cit. on pp. 57, 156).

Hamaus, N., B. D. Wandelt, P. M. Sutter, G. Lavaux, M. S. Warren (2014b). Cosmology with Void-Galaxy Correlations. en. *Physical Review Letters* **112**:4, 041304. DOI: 10.1103/PhysRevLett.112.041304 (cit. on p. 88).

Hamilton, A. J. S. (1992). Measuring Omega and the real correlation function from the redshift correlation function. *apjl* **385**: L5–L8. DOI: 10.1086/186264 (cit. on p. 93).

Hamilton, A. J. S. (1998). Linear Redshift Distortions: a Review. **231**: 185. DOI: 10.1007/978-94-011-4960-0_17 (cit. on p. 93).

Hanany, S. (1997). The MAXIMA and BOOMERANG experiments. en. URL: https://inspirehep.net/literature/440678 (cit. on p. 25).

Hartlap, J., P. Simon, P. Schneider (2007). Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix. *Astronomy and Astrophysics* **464**: 399–404. DOI: 10.1051/0004-6361:20066170 (cit. on p. 95).

Hausman, M. A., D. W. Olson, B. D. Roth (1983). The evolution of voids in the expanding universe. *The Astrophysical Journal* **270**: 351–359. DOI: 10.1086/161128 (cit. on p. 56).

Hawken, A. J., B. R. Granett, A. Iovino, L. Guzzo, J. A. Peacock, et al. (2017). The VIMOS Public Extragalactic Redshift Survey. Measuring the growth rate of structure around cosmic voids. *\aap* **607**: A54. DOI: 10.1051/0004-6361/201629678 (cit. on pp. 64, 89, 118).

Hawken, A. J., D. Michelett, B. Granett, A. Iovino, L. Guzzo (2016). Measuring the growth rate of structure around cosmic voids. **308**: 571–574. DOI: 10.1017/S1743921316010590 (cit. on pp. 64, 119).

Hawken, A. J., M. Aubert, A. Pisani, M.-C. Cousinou, S. Escoffier, et al. (2020). Constraints on the growth of structure around cosmic voids in eBOSS DR14. en. *Journal of Cosmology and Astroparticle Physics* **2020**:06, 012–012. DOI: 10.1088/1475-7516/2020/06/012 (cit. on pp. 64, 89, 110, 123).

Heitmann, K., H. Finkel, A. Pope, V. Morozov, N. Frontiere, et al. (2019a). The Outer Rim Simulation: A Path to Many-core Supercomputers. *The Astrophysical Journal Supplement Series* **245**: 16. DOI: 10.3847/1538-4365/ab4da1 (cit. on p. 74).

Heitmann, K., T. D. Uram, H. Finkel, N. Frontiere, S. Habib, et al. (2019b). HACC Cosmological Simulations: First Data Release. *The Astrophysical Journal Supplement Series* **244**: 17. DOI: 10.3847/1538-4365/ab3724 (cit. on p. 74).

Higuchi, Y., M. Oguri, T. Hamana (2013). Measuring the mass distribution of voids with stacked weak lensing. *Monthly Notices of the Royal Astronomical Society* **432**: 1021–1031. DOI: 10.1093/mnras/stt521 (cit. on p. 62).

Hinshaw, G., D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, et al. (2013). NINE-YEAR *WILKINSON MICROWAVE ANISOTROPY PROBE* (*WMAP*) OBSERVATIONS: COSMOLOGICAL PARAMETER RESULTS. *The Astrophysical Journal Supplement Series* **208**:2, 19. DOI: 10.1088/0067-0049/208/2/19 (cit. on p. 25).

Hoffman, G. L., E. E. Salpeter, I. Wasserman (1983). Spherical simulations of holes and honeycombs in Friedmann universes. *The Astrophysical Journal* **268**: 527–539. DOI: 10.1086/160976 (cit. on p. 56).

Hoffman, Y., O. Metuki, G. Yepes, S. Gottlöber, J. E. Forero-Romero, et al. (2012). A kinematic classification of the cosmic web. en. *Monthly Notices of the Royal Astronomical Society* **425**:3, 2049–2057. DOI: 10.1111/j.1365-2966.2012.21553.x (cit. on p. 44).

Hou, J., A. G. Sánchez, A. J. Ross, A. Smith, R. Neveux, et al. (2020). The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: BAO and RSD measurements from anisotropic clustering analysis of the Quasar Sample in configuration space between redshift 0.8 and 2.2. *arXiv e-prints* **2007**: arXiv:2007.08998. URL: http://adsabs.harvard.edu/abs/2020arXiv200708998H (cit. on pp. 96, 115, 117, 177).

Howlett, C., A. J. Ross, L. Samushia, W. J. Percival, M. Manera (2015). The clustering of the SDSS main galaxy sample – II. Mock galaxy catalogues and a measurement of the growth of structure from redshift space distortions at z = 0.15. en. *Monthly Notices of the Royal Astronomical Society* **449**:1, 848–866. DOI: 10.1093/mnras/stu2693 (cit. on p. 117).

Hoyle, F., M. S. Vogeley (2004). Voids in the Two-Degree Field Galaxy Redshift Survey. en. *The Astrophysical Journal* **607**:2, 751. DOI: 10.1086/386279 (cit. on pp. 44, 53).

Hu, W., N. Sugiyama (1995). Anisotropies in the cosmic microwave background: an analytic approach. en. *The Astrophysical Journal* **444**: 489. DOI: 10.1086/175624 (cit. on p. 24).

Hu, W., N. Sugiyama, J. Silk (1997). The physics of microwave background anisotropies. en. *Nature* **386**:6620, 37–43. DOI: 10.1038/386037a0 (cit. on p. 25).

Hubble, E. (1929). A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science* **15**: 168–173. DOI: 10.1073/pnas.15.3.168 (cit. on p. 12).

Hui, L., A. Stebbins, S. Burles (1999). A Geometrical Test of the Cosmological Energy Contents Using the Lyα Forest. *The Astrophysical Journal Letters* **511**: L5–L8. DOI: 10.1086/311826 (cit. on p. 126).

Icke, V. (1984). Voids and filaments. *Monthly Notices of the Royal Astronomical Society* **206**: 1P–3P. DOI: 10.1093/mnras/206.1.1P (cit. on pp. 56, 57, 60).

Ikeuchi, S., K. Tomisaka, J. P. Ostriker (1983). The structure and expansion law of a shock wave in an expanding universe. *The Astrophysical Journal* **265**: 583–596. DOI: 10.1086/160703 (cit. on p. 42).

Jamieson, D., M. Loverde (2019). Separate universe void bias. *Physical Review D* **100**: 123528. DOI: 10.1103/PhysRevD.100.123528 (cit. on p. 54).

Jennings, E., C. M. Baugh, S. Pascoli (2012). Testing dark energy using pairs of galaxies in redshift space. en. *Monthly Notices of the Royal Astronomical Society* **420**:2, 1079–1091. DOI: 10.1111/j.1365-2966.2011.20064.x (cit. on p. 126).

Jennings, E., Y. Li, W. Hu (2013). The abundance of voids and the excursion set formalism. *\mnras* **434**:3, 2167–2181. DOI: 10.1093/mnras/stt1169 (cit. on pp. 53, 60, 85).

Jôeveer, M., J. Einasto, E. Tago (1978). Spatial distribution of galaxies and of clusters of galaxies in the southern galactic hemisphere. en. *Monthly Notices of the Royal Astronomical Society* **185**:2, 357–370. DOI: 10.1093/mnras/185.2.357 (cit. on pp. 40, 41).

Kaiser, N. (1987). Clustering in real space and in redshift space. en. *Monthly Notices of the Royal Astronomical Society* **227**:1, 1–21. DOI: 10.1093/mnras/227.1.1 (cit. on pp. 87, 89).

Kaiser, N., G. Squires (1993). Mapping the dark matter with weak gravitational lensing. en. *The Astrophysical Journal* **404**: 441. DOI: 10.1086/172297 (cit. on p. 20).

Kauffmann, G., A. P. Fairall (1991). Voids in the distribution of galaxies: an assessment of their significance and derivation of a void spectrum. en. *Monthly Notices of the Royal Astronomical Society* **248**:2, 313–324. DOI: 10.1093/mnras/248.2.313 (cit. on p. 44).

Kauffmann, G., A. L. Melott (1992). The void spectrum in two-dimensional numerical simulations of gravitational clustering. *The Astrophysical Journal* **393**: 415–430. DOI: 10.1086/171515 (cit. on p. 60).

Kim, Y.-R., R. A. C. Croft (2007). A potentially pure test of cosmic geometry: galaxy clusters and the real space Alcock-Paczyński test. *Monthly Notices of the Royal Astronomical Society* **374**: 535–546. DOI: 10.1111/j.1365-2966.2006.11168.x (cit. on p. 126).

Kirshner, R. P., A. Oemler Jr., P. L. Schechter, S. A. Shectman (1981). A million cubic megaparsec void in Bootes. en. *The Astrophysical Journal* **248**: L57. DOI: 10.1086/183623 (cit. on pp. 41, 59).

Kirshner, R. P., A. Oemler Jr., P. L. Schechter, S. A. Shectman (1987). A survey of the Bootes void. *The Astrophysical Journal* **314**: 493–506. DOI: 10.1086/165080 (cit. on pp. 41, 59).

Kitaura, F.-S., C.-H. Chuang, Y. Liang, C. Zhao, C. Tao, et al. (2016). Signatures of the Primordial Universe from Its Emptiness: Measurement of Baryon Acoustic Oscillations from Minima of the Density Field. *Physical Review Letters* **116**: 171301. DOI: 10.1103/PhysRevLett.116.171301 (cit. on p. 64).

Komatsu, E., K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, et al. (2011). SEVEN-YEAR *WILKINSON MICROWAVE ANISOTROPY PROBE* (*WMAP*) OBSERVATIONS: COSMOLOGICAL INTERPRETATION. *The Astrophysical Journal Supplement Series* **192**:2, 18. DOI: 10.1088/0067-0049/192/2/18 (cit. on pp. 25, 74).

Kovács, A., C. Sánchez, J. García-Bellido, J. Elvin-Poole, N. Hamaus, et al. (2019). More out of less: an excess integrated Sachs-Wolfe signal from supervoids mapped out by the Dark Energy Survey. *Monthly Notices of the Royal Astronomical Society* **484**: 5267–5277. DOI: 10.1093/mnras/stz341 (cit. on p. 63).

Kovács, A., C. Sánchez, J. García-Bellido, S. Nadathur, R. Crittenden, et al. (2017). Imprint of DES superstructures on the cosmic microwave background. *Monthly Notices of the Royal Astronomical Society* **465**: 4166–4179. DOI: 10.1093/mnras/stw2968 (cit. on p. 63).

Krause, E., T.-C. Chang, O. Doré, K. Umetsu (2013). The Weight of Emptiness: The Gravitational Lensing Signal of Stacked Voids. *The Astrophysical Journal Letters* **762**: L20. DOI: 10.1088/2041-8205/762/2/L20 (cit. on p. 62).

Kusenko, A. (2009). Sterile neutrinos. en. *New Journal of Physics* **11**:10, 105007. DOI: 10.1088/1367-2630/11/10/105007 (cit. on p. 21).

Lake, K., R. Pim (1985). Development of voids in the thin-wall approximation. I - General characteristics of spherical vacuum voids. *The Astrophysical Journal* **298**: 439–447. DOI: 10.1086/163629 (cit. on p. 56).

Lambas, D. G., M. Lares, L. Ceccarelli, A. N. Ruiz, D. J. Paz, et al. (2016). The sparkling Universe: the coherent motions of cosmic voids. *Monthly Notices of the Royal Astronomical Society* **455**: L99–L103. DOI: 10.1093/mnrasl/slv151 (cit. on p. 88).

Landy, S. D., A. S. Szalay (1993). Bias and variance of angular correlation functions. *\apj* **412**: 64–71. DOI: 10.1086/172900 (cit. on pp. 92, 102).

Laureijs, R., J. Amiaux, S. Arduini, J. .-.-L. Augueres, J. Brinchmann, et al. (2011). Euclid Definition Study Report. *arXiv e-prints* **1110**: arXiv:1110.3193. URL: http://adsabs.harvard.edu/abs/2011arXiv1110.3193L (cit. on pp. 28, 118).

Lavaux, G., B. D. Wandelt (2010). Precision cosmology with voids: definition, methods, dynamics. en. *Monthly Notices of the Royal Astronomical Society* **403**:3, 1392–1408. DOI: 10.1111/j.1365-2966.2010.16197.x (cit. on p. 44).

Lavaux, G., B. D. Wandelt (2012). PRECISION COSMOGRAPHY WITH STACKED VOIDS. en. *The Astrophysical Journal* **754**:2, 109. DOI: 10.1088/0004-637X/754/2/109 (cit. on pp. 63, 127, 129, 135, 163).

Lee, J., D. Park (2006). Rotation of Cosmic Voids and Void Spin Statistics. *The Astrophysical Journal* **652**: 1–5. DOI: 10.1086/507936 (cit. on p. 61).

Lee, J., D. Park (2009). Constraining the Dark Energy Equation of State with Cosmic Voids. *The Astrophysical Journal Letters* **696**: L10–L12. DOI: 10.1088/0004-637X/696/1/L10 (cit. on p. 61).

Lemaître, G. (1927). Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Annales de la Société Scientifique de Bruxelles* **47**: 49–59. URL: http://adsabs.harvard.edu/abs/1927ASSB...47...49L (cit. on p. 12).

Li, X.-D., C. Park, C. G. Sabiu, J. Kim (2015). Cosmological constraints from the redshift dependence of the Alcock–Paczynski test and volume effect: galaxy two-point correlation function. en. *Monthly Notices of the Royal Astronomical Society* **450**:1, 807–814. DOI: 10.1093/mnras/stv622 (cit. on p. 126).

Linder, E. V. (2003). Exploring the Expansion History of the Universe. *Physical Review Letters* **90**:9, 091301. DOI: 10.1103/PhysRevLett.90.091301 (cit. on pp. 23, 36).

Linder, E. V., R. N. Cahn (2007). Parameterized beyond-Einstein growth. en. *Astroparticle Physics* **28**:4-5, 481–488. DOI: 10.1016/j.astropartphys.2007.09.003 (cit. on p. 36).

Little, B., D. H. Weinberg (1994). Cosmic voids and biased galaxy formation. en. *Monthly Notices of the Royal Astronomical Society* **267**:3, 605–628. DOI: 10.1093/mnras/267.3.605 (cit. on p. 54).

López-Corredoira, M. (2014). ALCOCK-PACZYŃSKI COSMOLOGICAL TEST. en. *The Astrophysical Journal* **781**:2, 96. DOI: 10.1088/0004-637X/781/2/96 (cit. on p. 126).

Lyke, B. W., A. N. Higley, J. N. McLane, D. P. Schurhammer, A. D. Myers, et al. (2020). The Sloan Digital Sky Survey Quasar Catalog: Sixteenth Data Release. *The Astrophysical Journal Supplement Series* **250**: 8. DOI: 10.3847/1538-4365/aba623 (cit. on p. 72).

Madau, P., M. Dickinson (2014). Cosmic Star-Formation History. *Annual Review of Astronomy and Astrophysics* **52**: 415–486. DOI: 10.1146/annurev-astro-081811-125615 (cit. on p. 71).

Mao, Q., A. A. Berlind, R. J. Scherrer, M. C. Neyrinck, R. Scoccimarro, et al. (2017). Cosmic Voids in the SDSS DR12 BOSS Galaxy Sample: the Alcock–Paczyński test. en. *The Astrophysical Journal* **835**:2, 160. DOI: 10.3847/1538-4357/835/2/160 (cit. on pp. 53, 63, 127, 129–131, 135, 151, 161, 162).

Marinoni, C., A. Buzzi (2010). A geometric measure of dark energy with pairs of galaxies. en. *Nature* **468**:7323, 539–541. DOI: 10.1038/nature09577 (cit. on p. 126).

Martel, H., I. Wasserman (1990). Simulation of cosmological voids in Lambda greater than 0 Friedmann models. *The Astrophysical Journal* **348**: 1–25. DOI: 10.1086/168208 (cit. on p. 57).

Massey, R., D. Harvey, J. Liesenborgs, J. Richard, S. Stach, et al. (2018). Dark matter dynamics in Abell 3827: new data consistent with standard cold dark matter. en. *Monthly Notices of the Royal Astronomical Society* **477**:1, 669–677. DOI: 10.1093/mnras/sty630 (cit. on p. 20).

Mather, J. C., E. S. Cheng, D. A. Cottingham, R. E. Eplee Jr., D. J. Fixsen, et al. (1994). Measurement of the cosmic microwave background spectrum by the COBE FIRAS instrument. *The Astrophysical Journal* **420**: 439–444. DOI: 10.1086/173574 (cit. on p. 25).

Matsubara, T., Y. Suto (1996). Cosmological Redshift Distortion of Correlation Functions as a Probe of the Density Parameter and the Cosmological Constant. *The Astrophysical Journal Letters* **470**: L1. DOI: 10.1086/310290 (cit. on p. 126).

McDonald, P., J. Miralda-Escudé (1999). Measuring the Cosmological Geometry from the Lyα Forest along Parallel Lines of Sight. *The Astrophysical Journal* **518**: 24–31. DOI: 10.1086/307264 (cit. on p. 126).

Melchior, P., P. M. Sutter, E. S. Sheldon, E. Krause, B. D. Wandelt (2014). First measurement of gravitational lensing by cosmic voids in SDSS. en. *Monthly Notices of the Royal Astronomical Society* **440**:4, 2922–2927. DOI: 10.1093/mnras/stu456 (cit. on p. 62).

Melott, A. L. (1983). Clustering velocities in the adiabatic picture of galaxy formation. en. *Monthly Notices of the Royal Astronomical Society* **205**:3, 637–641. DOI: 10.1093/mnras/205.3.637 (cit. on p. 42).

Melott, A. L. (1987). Voids and velocities in initially Gaussian models for large-scale structure. en. *Monthly Notices of the Royal Astronomical Society* **228**:4, 1001–1023. DOI: 10.1093/mnras/228.4.1001 (cit. on p. 60).

Micheletti, D., A. Iovino, A. J. Hawken, B. R. Granett, M. Bolzonella, et al. (2014). The VIMOS Public Extragalactic Redshift Survey. Searching for cosmic voids. *Astronomy and Astrophysics* **570**: A106. DOI: 10.1051/0004-6361/201424107 (cit. on pp. 44, 45).

Milgrom, M. (1983). A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis. *The Astrophysical Journal* **270**: 365–370. DOI: 10.1086/161130 (cit. on p. 21).

Müller, V., S. Arbabi-Bidgoli, J. Einasto, D. Tucker (2000). Voids in the Las Campanas Redshift Survey versus cold dark matter models. *Monthly Notices of the Royal Astronomical Society* **318**: 280–288. DOI: 10.1046/j.1365-8711.2000.03775.x (cit. on pp. 44, 53).

Myers, A. D., N. Palanque-Delabrouille, A. Prakash, I. Pâris, C. Yeche, et al. (2015). The SDSS-IV extended Baryon oscillation spectroscopic survey: Quasar target selection. *The Astrophysical Journal Supplement Series* **221**:2, 27 (cit. on p. 72).

Nadathur, S., S. Hotchkiss (2015a). The nature of voids - I. Watershed void finders and their connection with theoretical models. *Monthly Notices of the Royal Astronomical Society* **454**: 2228–2241. DOI: 10.1093/mnras/stv2131 (cit. on p. 49).

Nadathur, S., S. Hotchkiss (2015b). The nature of voids - II. Tracing underdensities with biased galaxies. *Monthly Notices of the Royal Astronomical Society* **454**: 889–901. DOI: 10.1093/mnras/stv1994 (cit. on pp. 54, 85).

Nadathur, S., S. Hotchkiss, J. M. Diego, I. T. Iliev, S. Gottlöber, et al. (2015). Self-similarity and universality of void density profiles in simulation and SDSS data. *Monthly Notices of the Royal Astronomical Society* **449**: 3997–4009. DOI: 10.1093/mnras/stv513 (cit. on p. 57).

Nadathur, S. (2016). Testing cosmology with a catalogue of voids in the BOSS galaxy surveys. *Monthly Notices of the Royal Astronomical Society* **461**: 358–370. DOI: 10.1093/mnras/stw1340 (cit. on p. 154).

Nadathur, S., P. M. Carter, W. J. Percival, H. A. Winther, J. Bautista (2019a). Beyond BAO: improving cosmological constraints from BOSS with measurement of the void-galaxy cross-correlation. *Physical Review D* **100**:2, 023504. DOI: 10.1103/PhysRevD.100.023504 (cit. on pp. 64, 65, 89, 92, 116, 117).

Nadathur, S., P. M. Carter, W. J. Percival, H. A. Winther, J. E. Bautista (2019b). REVOLVER: REal-space VOid Locations from suVEy Reconstruction. *Astrophysics Source Code Library*, ascl:1907.023. URL: http://adsabs.harvard.edu/abs/2019ascl.soft07023N (cit. on pp. 44, 49).

Nadathur, S., R. Crittenden (2016). A Detection of the Integrated Sachs-Wolfe Imprint of Cosmic Superstructures Using a Matched-filter Approach. *apj* **830**:1, L19. DOI: 10.3847/2041-8205/830/1/L19 (cit. on p. 63).

Nadathur, S., S. Hotchkiss (2014). A robust public catalogue of voids and superclusters in the SDSS Data Release 7 galaxy surveys. *Monthly Notices of the Royal Astronomical Society* **440**: 1248–1262. DOI: 10.1093/mnras/stu349 (cit. on p. 53).

Nadathur, S., S. Hotchkiss, R. Crittenden (2017). Tracing the gravitational potential using cosmic voids. *Monthly Notices of the Royal Astronomical Society* **467**: 4067–4079. DOI: 10.1093/mnras/stx336 (cit. on p. 63).

Nadathur, S., W. J. Percival (2019). An accurate linear model for redshift space distortions in the void-galaxy correlation function. *Monthly Notices of the Royal Astronomical Society* **483**:3, 3472–3487. DOI: 10.1093/mnras/sty3372 (cit. on pp. 90, 121, 122).

Nadathur, S., W. J. Percival, F. Beutler, H. A. Winther (2020a). Testing Low-Redshift Cosmic Acceleration with Large-Scale Structure. *Physical Review Letters* **124**: 221301. DOI: 10.1103/PhysRevLett.124.221301 (cit. on p. 66).

Nadathur, S., A. Woodfinden, W. J. Percival, M. Aubert, J. Bautista, et al. (2020b). The completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: geometry and growth from the anisotropic void-galaxy correlation function in the luminous red galaxy sample. *arXiv e-prints* **2008**: arXiv:2008.06060. URL: http://adsabs.harvard.edu/abs/2020arXiv200806060N (cit. on pp. 64, 65, 107, 116–118, 122).

Neveux, R., E. Burtin, A. de Mattia, A. Smith, A. J. Ross, et al. (2020). The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: BAO and RSD measurements from the anisotropic power spectrum of the Quasar sample between redshift 0.8 and 2.2. *arXiv e-prints* **2007**: arXiv:2007.08999. URL: http://adsabs.harvard.edu/abs/2020arXiv200708999N (cit. on pp. 96, 115, 117, 177).

Neyrinck, M. C. (2008). zobov: a parameter-free void-finding algorithm. en. *Monthly Notices of the Royal Astronomical Society* **386**:4, 2101–2109. DOI: 10.1111/j.1365-2966.2008.13180.x (cit. on pp. 44, 46, 48).

Nusser, A. (2005). The Alcock-Paczyński test in redshifted 21-cm maps. *Monthly Notices of the Royal Astronomical Society* **364**: 743–750. DOI: 10.1111/j.1365-2966.2005.09603.x (cit. on p. 126).

Okumura, T. et al. (2016). The Subaru FMOS galaxy redshift survey (FastSound). IV. New constraint on gravity theory from redshift space distortions at $z\sim 1.4$. *Publ. Astron. Soc. Jap.* **68**:3, 38. DOI: 10.1093/pasj/psw029 (cit. on pp. 118, 119, 178).

Ostriker, J. P., P. J. E. Peebles, A. Yahil (1974). The size and mass of galaxies, and the mass of the universe. *The Astrophysical Journal Letters* **193**: L1–L4. DOI: 10.1086/181617 (cit. on p. 19).

Ostriker, J. P., P. J. Steinhardt (1995). The observational case for a low-density Universe with a non-zero cosmological constant. en. *Nature* **377**:6550, 600–602. DOI: 10.1038/377600a0 (cit. on p. 21).

Padilla, N. D., L. Ceccarelli, D. G. Lambas (2005). Spatial and dynamical properties of voids in a Λ cold dark matter universe. en. *Monthly Notices of the Royal Astronomical Society* **363**:3, 977–990. DOI: 10.1111/j.1365-2966.2005.09500.x (cit. on pp. 44, 58, 64, 89).

Pan, D. C., M. S. Vogeley, F. Hoyle, Y.-Y. Choi, C. Park (2012). Cosmic voids in Sloan Digital Sky Survey Data Release 7. *Monthly Notices of the Royal Astronomical Society* **421**: 926–934. DOI: 10.1111/j.1365-2966.2011.20197.x (cit. on p. 64).

Park, D., J. Lee (2007). Void Ellipticity Distribution as a Probe of Cosmology. *Physical Review Letters* **98**: 081301. DOI: 10.1103/PhysRevLett.98.081301 (cit. on p. 61).

Paz, D., M. Lares, L. Ceccarelli, N. Padilla, D. G. Lambas (2013). Clues on void evolution-II. Measuring density and velocity profiles on SDSS galaxy redshift space distortions. *Monthly Notices of the Royal Astronomical Society* **436**: 3480–3491. DOI: 10.1093/mnras/stt1836 (cit. on pp. 58, 64, 89).

Peebles, P. J. E. (1974). The Nature of the Distribution of Galaxies. *Astronomy and Astrophysics* **32**: 197. URL: http://adsabs.harvard.edu/abs/1974A%26A....32..197P (cit. on p. 40).

Peebles, P. J. E. (1980). *The large-scale structure of the universe*. Princeton University Press (cit. on pp. 29, 35).

Peebles, P. J. E. (1982). The peculiar velocity around a hole in the galaxy distribution. *The Astrophysical Journal* **257**: 438–441. DOI: 10.1086/160001 (cit. on p. 56).

Peebles, P. J. E., M. G. Hauser (1974). Statistical Analysis of Catalogs of Extragalactic Objects. III. The Shane-Wirtanen and Zwicky Catalogs. *The Astrophysical Journal Supplement Series* **28**: 19. DOI: 10.1086/190308 (cit. on p. 92).

Penzias, A. A., R. W. Wilson (1965). A Measurement of Excess Antenna Temperature at 4080 Mc/s. *The Astrophysical Journal* **142**: 419–421. DOI: 10.1086/148307 (cit. on p. 24).

Percival, W. J., A. J. Ross, A. G. Sánchez, L. Samushia, A. Burden, et al. (2014). The clustering of Galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: including covariance matrix errors. *Monthly Notices of the Royal Astronomical Society* **439**: 2531–2541. DOI: 10.1093/mnras/stu112 (cit. on p. 96).

Perico, E. L. D., R. Voivodic, M. Lima, D. F. Mota (2019). Cosmic voids in modified gravity scenarios. *Astronomy and Astrophysics* **632**: A52. DOI: 10.1051/0004-6361/201935949 (cit. on p. 60).

Perlmutter, S., G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, et al. (1999). Measurements of Ω and Λ from 42 High-Redshift Supernovae. en. *The Astrophysical Journal* **517**:2, 565. DOI: 10.1086/307221 (cit. on p. 22).

Pezzotta, A., S. de la Torre, J. Bel, B. R. Granett, L. Guzzo, et al. (2017). The VIMOS Public Extragalactic Redshift Survey (VIPERS). *\aap* **604**: A33. DOI: 10.1051/0004-6361/201630295 (cit. on pp. 118, 119, 178).

Phillipps, S. (1994). A possible geometric measurement of the cosmological constant. en. *Monthly Notices of the Royal Astronomical Society* **269**:4, 1077–1081. DOI: 10.1093/mnras/269.4.1077 (cit. on p. 126).

Pisani, A., G. Lavaux, P. M. Sutter, B. D. Wandelt (2014). Real-space density profile reconstruction of stacked voids. *Monthly Notices of the Royal Astronomical Society* **443**: 3238–3250. DOI: 10.1093/mnras/stu1399 (cit. on p. 121).

Pisani, A., P. M. Sutter, N. Hamaus, E. Alizadeh, R. Biswas, et al. (2015). Counting voids to probe dark energy. *Physical Review D* **92**: 083531. DOI: 10.1103/PhysRevD.92.083531 (cit. on pp. 60, 65).

Platen, E., R. Van De Weygaert, B. J. T. Jones (2007). A cosmic watershed: the WVF void detection technique. en. *Monthly Notices of the Royal Astronomical Society* **380**:2, 551–570. DOI: 10.1111/j.1365-2966.2007.12125.x (cit. on p. 49).

Platen, E., R. van de Weygaert, B. J. T. Jones, G. Vegter, M. A. A. Calvo (2011). Structural analysis of the SDSS Cosmic Web - I. Non-linear density field reconstructions. *Monthly Notices of the Royal Astronomical Society* **416**: 2494–2526. DOI: 10.1111/j.1365-2966.2011.18905.x (cit. on p. 44).

Pollina, G., M. Baldi, F. Marulli, L. Moscardini (2016). Cosmic voids in coupled dark energy cosmologies: the impact of halo bias. *Monthly Notices of the Royal Astronomical Society* **455**: 3075–3085. DOI: 10.1093/mnras/stv2503 (cit. on p. 54).

Pollina, G., N. Hamaus, K. Dolag, J. Weller, M. Baldi, et al. (2017). On the linearity of tracer bias around voids. *Monthly Notices of the Royal Astronomical Society* **469**: 787–799. DOI: 10.1093/mnras/stx785 (cit. on p. 54).

Pontzen, A., A. Slosar, N. Roth, H. V. Peiris (2016). Inverted initial conditions: Exploring the growth of cosmic structure and voids. *Physical Review D* **93**:10, 103519. DOI: 10.1103/PhysRevD.93.103519 (cit. on p. 156).

Prakash, A., T. C. Licquia, J. A. Newman, A. J. Ross, A. D. Myers, et al. (2016). The SDSS-IV extended baryon oscillation spectroscopic survey: luminous red galaxy target selection. *The Astrophysical Journal Supplement Series* **224**:2, 34 (cit. on p. 71).

Press, W. H., P. Schechter (1974). Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *The Astrophysical Journal* **187**: 425–438. DOI: 10.1086/152650 (cit. on p. 60).

Raghunathan, S., S. Nadathur, B. D. Sherwin, N. Whitehorn (2020). The Gravitational Lensing Signatures of BOSS Voids in the Cosmic Microwave Background. *The Astrophysical Journal* **890**: 168. DOI: 10.3847/1538-4357/ab6f05 (cit. on p. 63).

Raichoor, A., J. Comparat, T. Delubac, J.-P. Kneib, C. Yeche, et al. (2017). The SDSS-IV extended Baryon Oscillation Spectroscopic Survey: final emission line galaxy target selection. *\mnras* **471**: 3955–3973. DOI: 10.1093/mnras/stx1790 (cit. on p. 71).

Raichoor, A., A. de Mattia, A. J. Ross, C. Zhao, S. Alam, et al. (2020). The completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Large-scale Structure Catalogues and Measurement of the isotropic BAO between redshift 0.6 and 1.1 for the Emission Line Galaxy Sample. *arXiv e-prints* **2007**: arXiv:2007.09007. URL: http://adsabs.harvard.edu/abs/2020arXiv200709007R (cit. on pp. 71, 72).

Regos, E., M. J. Geller (1991). The evolution of void-filled cosmological structures. *The Astrophysical Journal* **377**: 14–28. DOI: 10.1086/170332 (cit. on p. 57).

Reid, B., S. Ho, N. Padmanabhan, W. J. Percival, J. Tinker, et al. (2016). SDSS-III Baryon Oscillation Spectroscopic Survey Data Release 12: galaxy target selection and large-scale structure catalogues. *\mnras* **455**:2, 1553–1573. DOI: 10.1093/mnras/stv2382 (cit. on p. 70).

Reid, M. J., D. W. Pesce, A. G. Riess (2019). An Improved Distance to NGC 4258 and Its Implications for the Hubble Constant. *The Astrophysical Journal* **886**:2, L27. DOI: 10.3847/2041-8213/ab552d (cit. on p. 27).

Ricciardelli, E., V. Quilis, J. Varela (2014). On the universality of void density profiles. en. *Monthly Notices of the Royal Astronomical Society* **440**:1, 601–609. DOI: 10.1093/mnras/stu307 (cit. on p. 57).

Riess, A. G., A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, et al. (1998). Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *The Astronomical Journal* **116**: 1009–1038. DOI: 10.1086/300499 (cit. on p. 22).

Ronconi, T., S. Contarini, F. Marulli, M. Baldi, L. Moscardini (2019). Cosmic voids uncovered - first-order statistics of depressions in the biased density field. *Monthly Notices of the Royal Astronomical Society* **488**: 5075–5084. DOI: 10.1093/mnras/stz2115 (cit. on p. 55).

Ronconi, T., F. Marulli (2017). Cosmological exploitation of cosmic void statistics. New numerical tools in the CosmoBolognaLib to extract cosmological constraints from the void size function. *\aap* **607**: A24. DOI: 10.1051/0004-6361/201730852 (cit. on pp. 53, 60).

Rood, H. J. (1988). Voids. *Annual Review of Astronomy and Astrophysics* **26**: 245–294. DOI: 10.1146/annurev.aa.26.090188.001333 (cit. on p. 39).

Ross, A. J., J. Bautista, R. Tojeiro, S. Alam, S. Bailey, et al. (2020). The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Large-scale Structure Catalogs for Cosmological Analysis. *Monthly Notices of the Royal Astronomical Society*. DOI: 10.1093/mnras/staa2416 (cit. on pp. 71, 72, 123).

Rubin, V. C., W. K. Ford Jr. (1970). Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *The Astrophysical Journal* **159**: 379. DOI: 10.1086/150317 (cit. on p. 19).

Ryden, B. S. (1994). Self-similar Expansion of Axisymmetric Voids. *The Astrophysical Journal* **423**: 534. DOI: 10.1086/173832 (cit. on pp. 57, 60).

Ryden, B. S. (1995). Measuring Q 0 from the Distortion of Voids in Redshift Space. en. *The Astrophysical Journal* **452**: 25. DOI: 10.1086/176277 (cit. on pp. 63, 126, 127).

Sachs, R. K., A. M. Wolfe (1967). Perturbations of a Cosmological Model and Angular Variations of the Microwave Background. *The Astrophysical Journal* **147**: 73. DOI: 10.1086/148982 (cit. on p. 63).

Sami, M. (2007). "Models of dark energy". en. *The Invisible Universe: Dark Matter and Dark Energy*. Ed. by L. Papantonopoulos. Lecture Notes in Physics. Berlin, Heidelberg: Springer, pp. 219–256. DOI: 10.1007/978-3-540-71013-4_8 (cit. on p. 23).

Sánchez, C., J. Clampitt, A. Kovacs, B. Jain, J. García-Bellido, et al. (2017). Cosmic voids and void lensing in the Dark Energy Survey Science Verification data. *Monthly Notices of the Royal Astronomical Society* **465**: 746–759. DOI: 10.1093/mnras/stw2745 (cit. on pp. 43, 44, 54, 62).

Schaap, W. E. (2007). "DTFE: the Delaunay Tessellation Field Estimator". PhD thesis. URL: http://adsabs.harvard.edu/abs/2007PhDT.......433S (cit. on p. 46).

Schuster, N., N. Hamaus, A. Pisani, C. Carbone, C. D. Kreisch, et al. (2019). The bias of cosmic voids in the presence of massive neutrinos. *arXiv e-prints*, arXiv:1905.00436 (cit. on p. 54).

Scolnic, D. M., D. O. Jones, A. Rest, Y. C. Pan, R. Chornock, et al. (2018). The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample. *The Astrophysical Journal* **859**:2, 101. DOI: 10.3847/1538-4357/aab9bb (cit. on p. 22).

Shandarin, S., H. A. Feldman, K. Heitmann, S. Habib (2006). Shapes and sizes of voids in the Lambda cold dark matter universe: excursion set approach. *Monthly Notices of the Royal Astronomical Society* **367**: 1629–1640. DOI: 10.1111/j.1365-2966.2006.10062.x (cit. on pp. 44, 60).

Sheth, R. K., R. Van De Weygaert (2004). A hierarchy of voids: much ado about nothing. en. *Monthly Notices of the Royal Astronomical Society* **350**:2, 517–538. DOI: 10.1111/j.1365-2966.2004.07661.x (cit. on pp. 44, 53, 58–60).

Shoji, M., J. Lee (2012). Voids in Redshift Space. *arXiv e-prints* **1203**: arXiv:1203.0869. URL: http://adsabs.harvard.edu/abs/2012arXiv1203.0869S (cit. on p. 66).

Simpson, F., J. A. Peacock (2010). Difficulties distinguishing dark energy from modified gravity via redshift distortions. *Physical Review D* **81**:4, 043512. DOI: 10.1103/PhysRevD.81.043512 (cit. on p. 127).

Slipher, V. M. (1913). The radial velocity of the Andromeda Nebula. *Lowell Observatory Bulletin* **2**: 56–57. URL: http://adsabs.harvard.edu/abs/1913LowOB...2...56S (cit. on pp. 11, 39).

Slipher, V. M. (1917). Nebulae. *Proceedings of the American Philosophical Society* **56**: 403–409. URL: http://adsabs.harvard.edu/abs/1917PAPhS..56..403S (cit. on pp. 11, 12, 39).

Smee, S. A., J. E. Gunn, A. Uomoto, N. Roe, D. Schlegel, et al. (2013). The Multi-object, Fiber-fed Spectrographs for the Sloan Digital Sky Survey and the Baryon Oscillation Spectroscopic Survey. \aj **146**: 32. DOI: 10.1088/0004-6256/146/2/32 (cit. on pp. 68–70).

Smith, A., E. Burtin, J. Hou, R. Neveux, A. J. Ross, et al. (2020). The Completed SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: N-body Mock Challenge for the Quasar Sample. *arXiv e-prints* **2007**: arXiv:2007.09003. URL: http://adsabs.harvard.edu/abs/2020arXiv200709003S (cit. on pp. 75, 103).

Smoot, G. F., C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, et al. (1992). Structure in the COBE differential microwave radiometer first year maps. en. *Astrophys.J.Lett.* **396**: L1–L5. DOI: 10.1086/186504 (cit. on p. 25).

Sołtan, A. M. (2019). ISW in ΛCDM or something else? *Monthly Notices of the Royal Astronomical Society* **488**: 2732–2742. DOI: 10.1093/mnras/stz1913 (cit. on p. 63).

Song, Y.-S., W. J. Percival (2009). Reconstructing the history of structure formation using redshift distortions. \jcap **2009**:10, 004–004. DOI: 10.1088/1475-7516/2009/10/004 (cit. on p. 115).

Sotiriou, T. P., V. Faraoni (2008). f(R) Theories Of Gravity. en. *Rev.Mod.Phys.* **82**: 451–497. DOI: 10.1103/RevModPhys.82.451 (cit. on p. 24).

Spolyar, D., M. Sahlén, J. Silk (2013). Topology and Dark Energy: Testing Gravity in Voids. en. *Physical Review Letters* **111**:24, 241103. DOI: 10.1103/PhysRevLett.111.241103 (cit. on p. 60).

Stark, C. W., A. Font-Ribera, M. White, K.-G. Lee (2015). Finding high-redshift voids using Lyman α forest tomography. *Monthly Notices of the Royal Astronomical Society* **453**: 4311–4323. DOI: 10.1093/mnras/stv1868 (cit. on p. 44).

Sutter, P. M., G. Lavaux, N. Hamaus, B. D. Wandelt, D. H. Weinberg, et al. (2014a). Sparse sampling, galaxy bias, and voids. *Monthly Notices of the Royal Astronomical Society* **442**: 462–471. DOI: 10.1093/mnras/stu893 (cit. on pp. 53, 54, 85).

Sutter, P. M., G. Lavaux, B. D. Wandelt, D. H. Weinberg (2012a). A First Application of the Alcock-Paczynski Test to Stacked Cosmic Voids. \apj **761**:2, 187. DOI: 10.1088/0004-637X/761/2/187 (cit. on pp. 58, 63, 127, 129, 135).

Sutter, P. M., G. Lavaux, B. D. Wandelt, D. H. Weinberg (2012b). A Public Void Catalog from the SDSS DR7 Galaxy Redshift Surveys Based on the Watershed Transform. *The Astrophysical Journal* **761**: 44. DOI: 10.1088/0004-637X/761/1/44 (cit. on p. 53).

Sutter, P. M., G. Lavaux, B. D. Wandelt, D. H. Weinberg, M. S. Warren, et al. (2014b). Voids in the SDSS DR9: observations, simulations, and the impact of the survey mask. *Monthly Notices of the Royal Astronomical Society* **442**: 3127–3137. DOI: 10.1093/mnras/stu1094 (cit. on p. 58).

Sutter, P. M., A. Pisani, B. D. Wandelt, D. H. Weinberg (2014c). A measurement of the Alcock–Paczyński effect using cosmic voids in the SDSS. en. *Monthly*

Notices of the Royal Astronomical Society **443**:4, 2983–2990. DOI: 10.1093/mnras/stu1392 (cit. on pp. 63, 127, 129, 130, 135, 151, 162).

Sutter, P., G. Lavaux, N. Hamaus, A. Pisani, B. Wandelt, et al. (2015). VIDE: The Void IDentification and Examination toolkit. en. *Astronomy and Computing* **9**: 1–9. DOI: 10.1016/j.ascom.2014.10.002 (cit. on pp. 44, 49).

Tamone, A., A. Raichoor, C. Zhao, A. de Mattia, C. Gorgoni, et al. (2020). The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Growth rate of structure measurement from anisotropic clustering analysis in configuration space between redshift 0.6 and 1.1 for the Emission Line Galaxy sample. *arXiv e-prints* **2007**: arXiv:2007.09009. URL: http://adsabs.harvard.edu/abs/2020arXiv200709009T (cit. on pp. 96, 115, 117, 177).

Taylor, A., B. Joachimi, T. Kitching (2013). Putting the Precision in Precision Cosmology: How accurate should your data covariance matrix be? *\mnras* **432**: 1928. DOI: 10.1093/mnras/stt270 (cit. on p. 95).

Thompson, L. A., S. A. Gregory (2011). An Historical View: The Discovery of Voids in the Galaxy Distribution. *arXiv e-prints* **1109**: arXiv:1109.1268. URL: http://adsabs.harvard.edu/abs/2011arXiv1109.1268T (cit. on p. 41).

Tifft, W. G., S. A. Gregory (1976). Direct observations of the large-scale distribution of galaxies. *The Astrophysical Journal* **205**: 696–708. DOI: 10.1086/154325 (cit. on p. 40).

Tifft, W. G., S. A. Gregory (1978). Observations of the Large Scale Distribution of Galaxies. **79**: 267. URL: http://adsabs.harvard.edu/abs/1978IAUS...79..267T (cit. on p. 41).

Tully, R. B., D. Pomarède, R. Graziani, H. M. Courtois, Y. Hoffman, et al. (2019). Cosmicflows-3: Cosmography of the Local Void. *The Astrophysical Journal* **880**: 24. DOI: 10.3847/1538-4357/ab2597 (cit. on p. 44).

Tyson, J. A., F. Valdes, R. A. Wenk (1990). Detection of systematic gravitational lens galaxy image alignments - Mapping dark matter in galaxy clusters. *The Astrophysical Journal Letters* **349**: L1–L4. DOI: 10.1086/185636 (cit. on p. 20).

Van de Weygaert, R., E. van Kampen (1993). Voids in gravitational instability scenarios – I. Global density and velocity fields in an Einstein–de Sitter universe. en. *Monthly Notices of the Royal Astronomical Society* **263**:2, 481–526. DOI: 10.1093/mnras/263.2.481 (cit. on pp. 53, 57).

Vargas-Magaña, M., J. E. Bautista, J.-C. Hamilton, N. G. Busca, É. Aubourg, et al. (2013). An optimized correlation function estimator for galaxy surveys. *\aap* **554**: A131. DOI: 10.1051/0004-6361/201220790 (cit. on p. 92).

Vaucouleurs, G. d. (1971). The Large-Scale Distribution of Galaxies and Clusters of Galaxies. en. *Publications of the Astronomical Society of the Pacific* **83**: 113. DOI: 10.1086/129088 (cit. on p. 39).

Verza, G., A. Pisani, C. Carbone, N. Hamaus, L. Guzzo (2019). The void size function in dynamical dark energy cosmologies. en. *Journal of Cosmology and Astroparticle Physics* **2019**:12, 040–040. DOI: 10.1088/1475-7516/2019/12/040 (cit. on p. 60).

Vielzeuf, P., A. Kovács, U. Demirbozan, P. Fosalba, E. Baxter, et al. (2019). Dark Energy Survey Year 1 Results: the lensing imprint of cosmic voids on the Cosmic Microwave Background. *arXiv e-prints* **1911**: arXiv:1911.02951. URL: http://adsabs.harvard.edu/abs/2019arXiv191102951V (cit. on p. 63).

Voivodic, R., M. Lima, C. Llinares, D. F. Mota (2017). Modeling void abundance in modified gravity. en. *Physical Review D* **95**:2, 024018. DOI: 10.1103/PhysRevD.95.024018 (cit. on p. 60).

Weinberg, S. (2015). Gravitation and Cosmology. en. URL: https://inspirehep.net/literature/1410180 (cit. on p. 13).

Wong, K. C., S. H. Suyu, G. C.-F. Chen, C. E. Rusu, M. Millon, et al. (2020). H0LiCOW – XIII. A 2.4 per cent measurement of H0 from lensed quasars: 5.3σ tension between early- and late-Universe probes. en. *Monthly Notices of the Royal Astronomical Society* **498**:1, 1420–1439. DOI: 10.1093/mnras/stz3094 (cit. on p. 28).

Wright, E. L., P. R. M. Eisenhardt, A. K. Mainzer, M. E. Ressler, R. M. Cutri, et al. (2010). The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. *\aj* **140**: 1868–1881. DOI: 10.1088/0004-6256/140/6/1868 (cit. on p. 70).

Yoo, J., Y. Watanabe (2012). Theoretical models of dark energy. *International Journal of Modern Physics D* **21**:12, 1230002. DOI: 10.1142/S0218271812300029 (cit. on p. 23).

York, D. G., J. Adelman, J. E. Anderson Jr., S. F. Anderson, J. Annis, et al. (2000). The Sloan Digital Sky Survey: Technical Summary. *\aj* **120**: 1579–1587. DOI: 10.1086/301513 (cit. on p. 68).

Zel'dovich, Y. B. (1970). Gravitational instability: An approximate theory for large density perturbations. *Astronomy and Astrophysics* **5**: 84–89. URL: http://adsabs.harvard.edu/abs/1970A%26A.....5...84Z (cit. on p. 30).

Zeldovich, Y. B. (1976). Structure of the Universe. *Comments on Astrophysics* **6**: 157. URL: http://adsabs.harvard.edu/abs/1976ComAp...6..157Z (cit. on p. 30).

Zeldovich, Y. B., J. Einasto, S. F. Shandarin (1982). Giant voids in the Universe. en. *Nature* **300**:5891, 407–413. DOI: 10.1038/300407a0 (cit. on p. 41).

Zhao, C., C.-H. Chuang, J. Bautista, A. de Mattia, A. Raichoor, et al. (2020). The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: one thousand multi-tracer mock catalogues with redshift evolution and systematics for galaxies and quasars of the final data release. *arXiv e-prints* **2007**: arXiv:2007.08997. URL: http://adsabs.harvard.edu/abs/2020arXiv200708997Z (cit. on p. 74).

Zhao, C., C. Tao, Y. Liang, F.-S. Kitaura, C.-H. Chuang (2016). DIVE in the cosmic web: voids with Delaunay triangulation from discrete matter tracer distributions. *Monthly Notices of the Royal Astronomical Society* **459**: 2670–2680. DOI: 10.1093/mnras/stw660 (cit. on p. 44).

Zwicky, F. (1933). Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta* **6**: 110–127. URL: http://adsabs.harvard.edu/abs/1933AcHPh...6..110Z (cit. on p. 19).

Zwicky, F. (1967). Compact and Dispersed Cosmic Matter, Part I. *Advances in Astronomy and Astrophysics* **5**: 267. URL: http://adsabs.harvard.edu/abs/1967AdA%26A...5..267Z (cit. on p. 39).