Alma Mater Studiorum — Università di Bologna

SCUOLA DI SCIENZE
Corso di Laurea Magistrale in Astrofisica e Cosmologia
Dipartimento di Fisica e Astronomia

Modelling the size distribution of Cosmic Voids

Tesi di Laurea Magistrale

Candidato:
Tommaso Ronconi

Relatore:
Lauro Moscardini

Correlatori:
Marco Baldi
Federico Marulli

Sessione I
Anno Accademico 2015-2016

Contents

Abstract

The currently accepted flat-ΛCDM Standard Cosmological Model accounts for the existence of two major energy components, Cold Dark Matter and Dark Energy, and for the flatness of the metric. Despite the theoretical framework of the Standard Cosmological Model is well outlined and generally accepted by the scientific community, there are still an amount of unresolved issues, which span from the definition of the initial conditions, to the physical nature of the major components of the Universe. The appearance of today's large scale structure depends on the initial conditions and on the metric and energy properties of the Universe. Thus, from the observable characteristics of these structures, it should be possible to impose constraints on the cosmological model's defining parameters.

In this context, Cosmic Voids are a major component of the large-scale distribution of matter in the Universe. They emerge out of the density troughs on the primordial Gaussian field of density fluctuations and are between the largest observable structures in the Universe (Guzzo and The Vipers Team, 2013; Nadathur, 2016).

With their relatively simple structure and shape, voids are useful probes of a variety of cosmic parameters and have indeed been shown to possess great potential for constraining Dark Energy and testing theories of gravity (Krause et al., 2013; Sutter et al., 2014b; Pollina et al., 2016). In spite of their popularity as cosmological tools, a gap of knowledge between Cosmic Void observations, simulations and theory still persists. Voids are for their nature only mildly non-linear and tend to become more spherical as they evolve (Icke, 1984), which suggests that their evolution should be easier to reconstruct than that of positive perturbations. Therefore, the spherical evolution approximation is particularly suited for describing the expansion of cosmic voids (Blumenthal et al., 1992). This approximation was used to build a theoretical model for the prediction of the void statistical distribution as a function of size (Sheth and van de Weygaert, 2004, hereafter SvdW), with the same approach used to derive the halo mass function. Since the Dark Matter halo distribution provides a large amount of cosmological information, the negative of this information should also be found in the Cosmic Void distribution. However, the SvdW model of void size function has been proven unsuccesful in reproducing the results obtained from cosmological simulations: previous works could only fit the observed distributions with a SVdW size function in relaxed conditions (i.e. letting the lower threshold δ_v vary to better represent the data), since voids found in simulations are usually under-predicted by models (Colberg et al., 2005; Pisani et al., 2015; Nadathur and Hotchkiss, 2015). This undermines the possibility of using voids as cosmological probes.

The goal of our thesis work is to cover the gap between theoretical predictions and measured distributions of cosmic voids. To this end we develop an algorithm to identify voids in simulated tracer distributions,

consistently with theory. We reconsider the accepted SvdW model, by inspecting the possibilities offered by a recently proposed refinement (the Vdn model, Jennings et al., 2013). Comparing the void catalogues obtained applying our algorithm to a set of simulations which differ in resolution, total size and cosmology, we validate the Vdn model, finding out that it is reliable over a large range of radii, at all the redshifts considered and for all the cosmological models inspected.

We have then extended our discussion to more realistic cases, searching for a size function model for voids identified in a distribution of biased tracers. We find that, naively applying the same procedure used for the unbiased tracers to a halo mock distribution does not provide successfull results, suggesting that the Vdn model requires to be reconsidered when dealing with biased samples. Thus, we test two alternative extensions of the model and find that two scaling relations exist: both the Dark Matter void radii and the underlying Dark Matter density contrast scale with the halo-defined void radii. We use these findings to develop a semi-analytical model which gives promising results.

Sommario

Il modello cosmologico attualmente accettato, chiamato flat-ΛCDM, prevede che l'Universo sia composto principalmente da due componenti energetiche principali, Materia Oscura Fredda ed Energia Oscura, e che abbia una metrica piatta. Nonostante la struttura teorica del Modello Cosmologico Standard sia ben delineata e generalmente accettata dalla comunità scientifica, sono ancora presenti svariati problemi non risolti. Tali problemi spaziano dalla definizione delle condizioni iniziali, alla natura fisica delle componenti principali dell'Universo. L'aspetto delle strutture su grande scala dipende dalle condizioni iniziali, dalla metrica e dalle proprietà energetiche dell'Universo. Perciò, dalle caratteristiche di queste strutture dovrebbe essere possibile imporre dei limiti ai parametri che definiscono il modello cosmologico.

In questo contesto, i vuoti cosmici sono una componente rilevante della distribuzione su grande scala delle materia. Si formano dalle depressioni del campo Gaussiano di fluttuazioni in densità e sono tra le più grandi strutture osservabili nell'Universo (Guzzo and The Vipers Team, 2013; Nadathur, 2016).

Grazie alla loro forma relativamente semplice, i vuoti sono utili sonde di una varietà di parametri cosmologici ed è stato infatti dimostrato che hanno un grande potenziale nel porre limiti sull'Energia Oscura e per testare le teorie di gravitazione (Krause et al., 2013; Sutter et al., 2014b; Pollina et al., 2016). A dispetto della loro popolarità come sonde cosmologiche, le osservazioni, i risultati delle simulazioni e i modelli teorici di vuoti non sono tuttora in totale accordo.

I vuoti sono per loro natura solo lievemente non lineari e tendono a diventare sferici mentre evolvono (Icke, 1984). Questo suggerisce che la loro evoluzione dovrebbe essere più semplice da modellare rispetto alle perturbazioni di densità positive. Per tale motivo, l'approssimazione di evoluzione sferica è particolarmente adatta a descrivere l'espansione dei Vuoti Cosmici (Blumenthal et al., 1992). Questa approssimazione è stata infatti utilizzata per costruire un modello teorico in grado di predire la distribuzione statistica dei vuoti in funzione delle loro dimensioni (Sheth and van de Weygaert, 2004, da quì in avanti SvdW), con lo stesso approccio utilizzato per derivare la funzione di massa degli aloni. Dato che la distribuzione degli aloni di Materia Oscura fornisce una grande quantità di informazione cosmologica, il negativo di questa informazione dovrebbe trovarsi nella distribuzione di vuoti cosmici. Ciononostante, il modello di funzione delle dimensioni di SvdW si è rivelato di scarso successo nel riprodurre i risultati ottenuti da simulazioni cosmologiche: nei lavori precedenti è stato soltanto possibile adattare la curva teorica alle distribuzioni osservate, lasciando libere le condizioni sui parametri (ovvero lasciando che la soglia di densità inferiore che entra nel modello, δ_v, variasse per meglio rappresentare i dati), dato che i vuoti trovati nelle simulazioni sarebbero altrimenti sottostimati dai modelli (Colberg et al., 2005; Pisani et al., 2015; Nadathur and Hotchkiss, 2015). Questo

mina la possibilità di utilizzare i vuoti come sonde cosmologiche.

L'obiettivo primario di questo lavoro di tesi è di colmare il divario tra le predizioni teoriche e le distribuzioni misurate di vuoti cosmici. A tal fine, abbiamo sviluppato un algoritmo in grado di identificare i vuoti nelle distribuzioni simulate di traccianti, in modo da renderlo consistente con la teoria. Abbiamo riconsiderato il modello attualmente accettato di SvdW, indagando sulle possibilità offerte da un suo recente miglioramento (il cosiddetto modello Vdn, Jennings et al., 2013). Comparando i cataloghi di vuoti ottenuti dall'applicazione del nostro algoritmo con una raccolta di simulazioni che differiscono in risoluzione, dimensione totale e cosmologia, abbiamo validato il modello Vdn, scoprendo che è affidabile su un largo intervallo di raggi, a tutti i redshift considerati e per tutte le cosmologie prese in esame.

Abbiamo quindi esteso la nostra discussione a casi più realistici, cercando una funzione di dimensioni in grado di fornire predizioni affidabili sui vuoti identificati nelle distribuzioni di traccianti affetti da bias. Abbiamo dimostrato che, applicare semplicemente la stessa procedura utilizzata nel caso di traccianti non affetti da bias ad una distribuzione simulata di aloni, non fornisce buoni risultati. Questo suggerisce che il modello Vdn richieda di essere rielaborato per poter trattare con campioni affetti da bias. Per questo, abbiamo vagliato due possibili estensioni alternative del modello, trovando che esistono due relazioni di scala: sia il raggio dei vuoti trovati nella distribuzione di traccianti non affetti da bias, che il contrasto in densità della materia oscura contenuta nei vuoti tracciati dagli aloni sono entrambi dipendenti dal raggio del vuoto negli aloni associato. Abbiamo infine utiliziamo queste scoperte per sviluppare un modello semi-analitico che ha dato risultati promettenti.

Introduction

One fundamental step in our understanding of the Universe was made possible by Einstein's General Theory of Relativity, since the deriving theory of gravity provides the mathematical structure of modern cosmological models. Our current understanding of cosmology is based on a set of assumptions none of which are seriously challenged by current data. At the very base of all these assumptions stands the *Cosmological Principle*, that is actually the sum of two assumptions, the first is isotropy, that is the property of presenting the same characteristics independently of the line of sight chosen by the observer. With the observation of the Universe in different directions it is possible to prove this property which, at the present time, was always confirmed. The latter assumption of the Cosmological Principle is homogeneity, that requires that we are not in a special place in the Universe. This statement is not as simple to test. When it is taken singularly, it is called *Copernican Principle*.

Furthermore, it is also implicitly assumed that photons are conserved. This is encoded in the *Etherington relation* which states that in a cosmological theory based on the Einstein's theory of gravity, distance measures are unique and it is then possible to link angular diameter distances to luminosity distances. This assumption is verified only in the case of photons traveling along unique, null geodesics.

These are the pillars on which standard cosmology is built. Thanks to these assumptions a wide range of models have been proposed since Einstein's *Cosmological Considerations* (Einstein, 1917), which together with the early work of Friedmann (1922) and Hubble's discovery of the expansion of the Universe (Hubble, 1929), depicted an accepted scenario. The exact definition of the Standard Cosmological Model improves with time following the progress of experiments in measuring the cosmological parameters and along with the attainment of a better understanding of physics. Nevertheless, most of cosmologists agree on a fundamental picture, the so-called *Big Bang* scenario, which describes the Universe as a system evolving from a highly compressed state existing around 10^{10} years ago. Besides the observed isotropic expansion, which may suggest that in earlier epochs the Universe was in a more compact configuration, this hypothesis is supported by two fundamental arguments. The first is the need of a compelling explanation for the production of the observed element abundances which, in the Big Bang model, is naturally solved in the primordial nucleosynthesis model, which predicts a distribution of elements pleasantly close to that actually measured. The other crucial argument rises from the thermal evolution implied by the Big Bang scenario. At some point a primordial fluid of several components decouples from radiation and from this moment on the mean free path of photons becomes infinite. A resulting black body fossil

radiation should be observed isotropically distributed in the present Universe. This radiation was actually observed in the microwave wavelength, at the black body temperature of approximatelly $3\,K$, by Penzias and Wilson (1965), resulting in the ultimate success of the Big Bang model. A wide range of information on both the primordial Universe and its geometry is encoded in this *Cosmic Microwave Background* (CMB) and a great scientific effort has been dedicated to its study in the last decades. Furthermore, the observed thermal anysotropies on the CMB spatial distribution are extremely small, which is another relevant proof in favor of the cosmological principle.

There is a set of fundamental ingredients which are necessary to define a cosmological model, first of all the Einstein's Field Equations, relating the geometry of the Universe with its matter and energy content. Their solution led to the *Friedmann-Robertson-Walker* metric which encodes the Cosmological Principle

Throughout the 20th century it became widely accepted that our Universe is not only made of components visible thanks to their interaction with photons, but also of an actually dominant amount of exotic and mysterious components which are still not completely understood. It is proved with observations of a wide range of phenomena (see e.g. Bertone et al., 2005) that matter is overwhelmed by a non-barionic fraction which now apparently interacts only through gravity with the rest of the Universe, the so-called *Cold Dark Matter* (CDM) (Zwicky, 1933, 1937). Moreover, in the late 1990s, a tension between the model predictions and the measured galaxy correlation on large scale was discovered (Maddox et al., 1990). Thanks to high redshift Supernovae distance measurements, the accelerated rate of expansion of the Universe was finally observed (Perlmutter et al., 1999; Riess et al., 1998; Schmidt et al., 1998). Several theories try to explain this unexpected behaviour both introducing the effect of some kind of *Dark Energy* or by a modified theory of gravity.

The currently accepted cosmological model is called *flat-*Λ*CDM*. It is defined by six main parameters and on the assumption of flatness. The energy content of the Universe is dominated by a *cosmological constant*, Λ, and by CDM. The current density distribution is the result of initial Gaussian adiabatic fluctuations seeded in an inflational period of fierce accelerated expansion.

Despite its notable success in describing many cosmological data sets, the Standard Cosmological Model leaves many questions unanswered. They span from the comprehension of the physics of inflation to the possible presence of relativistic species at the decoupling epoch, beyond the known photons and neutrinos. Moreover it is still not certain what the mass of neutrinos might be, nor whether the primordial helium abundance is consisten with the Big Bang nucleosynthesis predictions and if the initial fluctuations were actually adiabatic. Last but not least, the nature of dark matter and dark energy is still poorly understood.

The structures observed in the present Universe are formed as a result of the gravitational collapse of the initial perturbations in the density field. On the large scale their occurence drammatically depends not only on the action of gravity itself but on the background evolution too, namely on the same characteristics of the Universe that outline a cosmological model. Almost all viable existing theories for structure formation are hierarchical: the matter distribution evolves through a sequence of ever larger structures. Hierarchical scenarios of structure formation are

successful in explaining the evolution of gravitationally bound virialized haloes and, in particular, the *excursion set approach* provides a useful framework to evaluate their formation histories. This approach was originally proposed by Bond et al. (1991) as a re-formulation of the earlier Press and Schechter (1974) work on the prediction of the dark matter halo statistical distribution. It provides an analytic approximation for the distribution of halo masses, merger rates and formation times, and it can be extended to provide estimates of the mass distribution in randomly chosen positions in space. One of the key assumptions of this approach is inherited from the original work of Press and Schechter (1974): the initial perturbations evolve into virialized objects with a spherical collapse and their distribution can be predicted by a *halo mass function*. If one considers the large scale structure (LSS) of the Universe, both in the redshift surveys and in the numerical simulations, a common feature is the presence of filaments and sheets of matter. They represent a distinctly non-virialized distribution of matter for which gravitational contraction of initially unspherical density peaks has only been accomplished along one or two dimensions. Several modifications of the halo mass function have been proposed (Sheth et al., 2001; Sheth and Tormen, 2002) to account for ellipsoidal collapse and thus to better model the actual matter distribution in the Universe.

Cosmic voids occupy a significant fraction of the Universe volume. They are large underdense regions where matter is evacuated and then squeezed in between their boundaries, as a result of the repulsing effect of gravity. Figure 1 illustates how cosmic voids appear as a major component in the galaxy distribution. They have been identified in all the wide angle surveys ever obtained (see e.g. Colless et al., 2003a; Tegmark et al., 2004; Micheletti et al., 2014). While galaxy clusters encapsulate most of the mass in the Universe, voids are a dominant component of the total volume over the megaparsec-scale. For their nature, voids are only mildly nonlinear objects, since while evolving they encounter a crucial physical limit: matter density cannot be less than zero. This results in the minimum possible density contrast achievable by a region, $\delta = -1$. Furthermore, conversely to what happens to overdensities, the evolution of voids removes the eventual unsphericity of the initial density perturbation (Icke, 1984). These properties suggest that the employment of a simplified spherical expansion model may be more accurate in describing the void formation and evolution than it is in the description of overdensities (Blumenthal et al., 1992).

Cosmic voids have gathered in the last decade a large popularity as cosmological tools. With typical sizes over the tens of megaparsec they are between the largest observable structures in the Universe. For this reason they are particularly suited in bringing information about several hot topics. They are indeed extreme objects, meaning that they generate from the shortest wavelengths of the matter perturbation power spectrum. Moreover, their underdense nature makes them suitable for a wide range of dark energy measurements.

The excursion set model has been applied to the study of underdensities, developing a theoretical model to predict the size distribution of cosmic voids (Sheth and van de Weygaert, 2004). It is based on the assumption that cosmic voids are nearly spherical regions which have gone through shell-crossing, i.e. where the inner more underdense matter shells inside voids have reached the outer as a result of the differential outward acceleration they experience. Nevertheless, this model lacks of

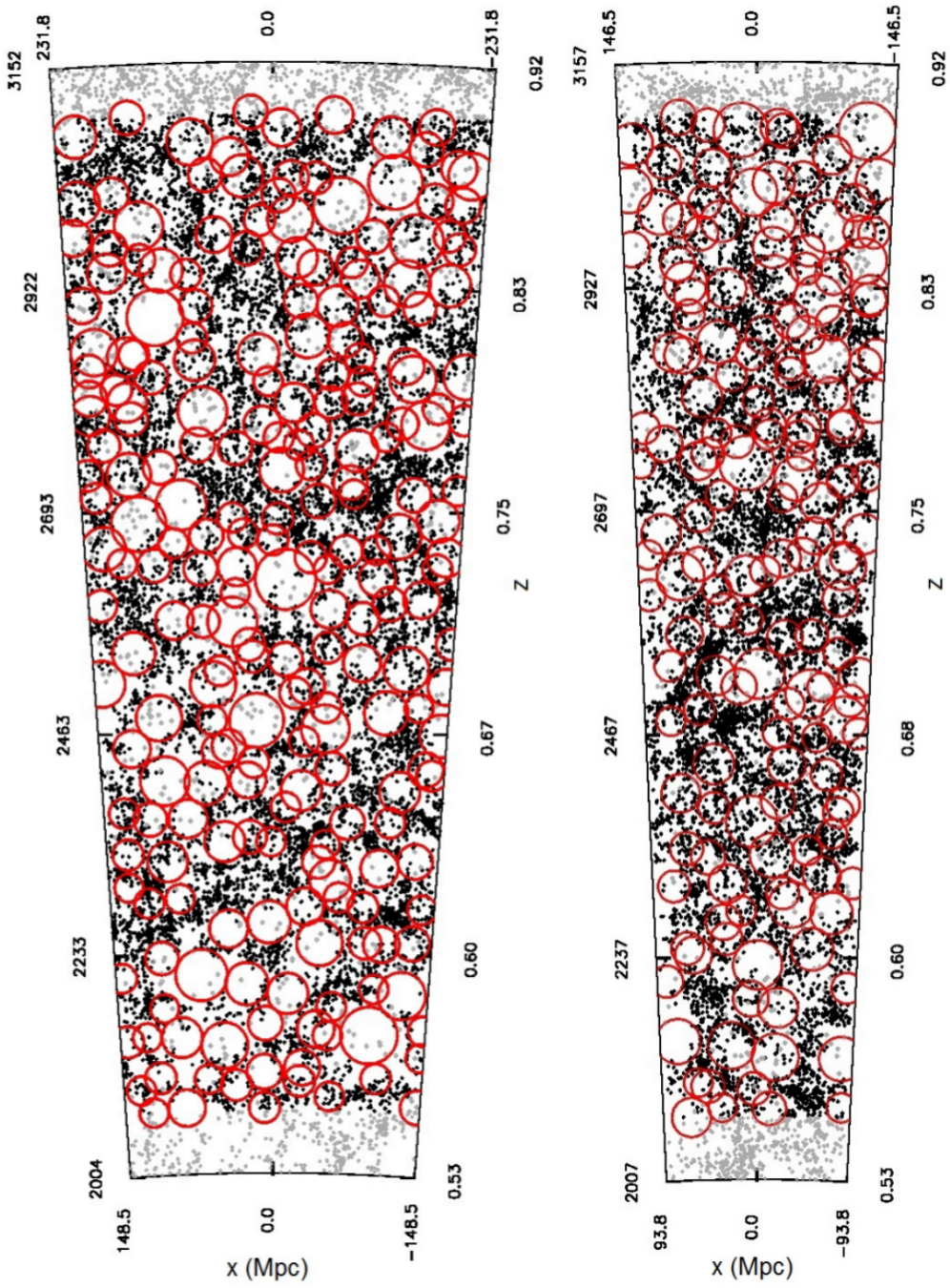

Figure 1: Two samples of galaxies obtained from the VIMOS Public Extragalactic Redshift Survey (VIPERS) at redshift $0.5 < z < 0.9$. *In both the shown slices spherical cosmic voids are obtained as maximal spheres embedding a fixed level of density and are shown in red. The grey points are galaxies detected as isolated and galaxies outside the sample redshift range. The scales show comoving distance in Mpc and the corresponding redshift. Figure from Micheletti et al. (2014).*

a complete validation, since it has been proven to systematically underpredict the number of voids actually measured both in surveys and simulations. Jennings et al. (2013) have proposed a modification of the model and demonstrated that it may better predict the actual size distribution of cosmic voids.

The aim of this thesis work is to obtain a reliable method for the analysis of the distribution of cosmic voids with an eye to the possible future employments. We search for a valid theoretical model capable of predicting the void size distribution in simulated but realistic catalogues of mass tracers.

Our work is organized as follows: In Chapter 1 we give a general introduction to the cosmological framework. In Chapter 2 the theory of non-linear evolution is outlined, focusing on the spherical expansion of underdensities which leads to the formation of cosmic voids. In Chapter 3 we present the main algorithms and computational tools we used in our work and finally we describe the algorithm we have built to find spherical non-overlapped voids. Chapter 4 is dedicated to the validation of our fiducial model, trying out its reliability for void abundances in a set of N-body simulations with different conditions. While in the previous Chapter we make use of unbiased tracers in Chapter 5 we extend our discussion to more realistics cases, searching for a void size function capable of making predictions for voids identified in the distributions of biased tracers. In Chapter 6 we finally draw out our conclusions and give some hints on the possible prosecution of our work.

Chapter 1

The cosmological framework

The growth of the primordial quantum fluctuations in the gravitational potential leads to the observed Large Scale Structure (LSS) of the Universe. As a result, studying the structures which populate the Universe should give information about the Universe as a whole, its energy content and the processes which dictated its evolution from the primordial stages to what is observed today. Since on sufficiently large scales the predominant interaction is gravity, a cosmological model should be described in terms of Albert Einstein's theory of General Relativity (GR, Einstein, 1915).

In the context of GR, gravity is encoded in the geometry of space-time, which is described by a *metric tensor*, $g_{\mu\nu}$. The minimum separation between two events in space-time, (t, x_1, x_2, x_3) and (t', x_1', x_2', x_2'), is given by

$$\mathrm{d}s^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu = g_{00}\,\mathrm{d}t^2 + 2\,g_{0i}\,\mathrm{d}x^i\,\mathrm{d}t + g_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j \ , \tag{1.1}$$

where $g_{00}\,\mathrm{d}t^2$ is the time term, $g_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j$ the spatial term and $2\,g_{0i}\,\mathrm{d}x^i\,\mathrm{d}t$ the mixed term.

The space-time metric is connected with the energy-momentum tensor, $T_{\mu\nu}$, which decribes the matter and energy content of the Universe, by the *Einstein's Field Equation of GR*:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu} \ , \tag{1.2}$$

where G is the gravitational constant and c is the light speed. $R_{\mu\nu}$ and R are the Ricci Tensor and Scalar, respectively, defined from the Riemann Tensor $R^\sigma_{\lambda\mu\nu}$:

$$R_{\mu\nu} \equiv R^\sigma_{\mu\sigma\nu} \ ,$$

$$R \equiv R^\nu_\nu = g^{\mu\nu}R_{\mu\nu} \ .$$

The left-hand side of equation (1.2) is the *Einstein's Tensor*, $G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$, which returns information about the space-time metric.

Equation (1.2) states that the matter and energy content of the Universe determine its geometry. In the general Big-Bang model the different energy components are considered as perfect fluids, with an energy-momentum tensor

$$T_{\mu\nu} = -p\,g_{\mu\nu} + (p + \rho c^2)u_\mu u_\nu \ , \tag{1.3}$$

where p is the pressure, ρc^2 is the energy density and u_λ are the components of the 4-velocity.

All the models developed in modern cosmology are based on the previous equations and on a fundamental assumption, the *Cosmological Principle*(CP), which states that on sufficiently large scales the Universe is both *isotropic* and *homogeneous*. As a result of the growth of the initial perturbations, with time passing this scale grows. Nowadays the CP is respected on scales over some hundreds of megaparsecs. The CP basically means that there are neither preferential directions nor positions in the Universe.

The metric tensor can be expressed in a simple form which encodes the CP, as we will discuss in the next Section.

1.1 The Friedmann-Robertson-Walker metric

Under the hypotesis of homogeneity and isotropy stated in the CP, it is possible to define a universal time so that, at each moment, the spatial 3D metric $\mathrm{d}l^2 = g_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j$ (where $i,j = 1,2,3$), is the same in each position in space. To do that, the mixed terms in eq. (1.1) have to be null, meaning that the g_{0i} components of the metric tensor have to be equal to zero. This leads to the general form

$$\mathrm{d}s^2 = c^2\,\mathrm{d}t^2 - g_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j = (c\,\mathrm{d}t)^2 - \mathrm{d}l^2\,. \tag{1.4}$$

Any point in space-time is labelled with a set of spatial coordinates x_i ($i = 1,2,3$), called *comoving coordinates*, which are the coordinates at rest with the Universe' expansion. In the same way, it is possible to define the *proper time*, t, as the time measured by a clock at rest with the Universe' expansion. To determine g_{ij} we have to find a 3-space metric which follows the requirements of homogeneity and isotropy of the CP.

The 3-dimensional Rienmann tensor R_{ijkl} determines the curvature of a 3-space while the scalar product of its contraction $R_{ij} \equiv R^l_{ilj}$ (the Ricci curvature tensor), with the 3-space metric, g_{ij}, gives the scalar curvature R. With the CP, R_{ijkl} do not depend on the derivatives of the metric g_{ij}, while the scalar R has to be constant. For the simmetry properties of the deriving form of the Rienmann tensor, it is possible to define the most general metric of the form (1.4) embedding the CP, which is the maximally symmetric *Friedmann-Robertson-Walker metric*:

$$\mathrm{d}s^2 = (c\;\mathrm{d}t)^2 - a^2(t)\left[\frac{\mathrm{d}r^2}{1-\kappa r^2} + r^2(\sin^2\theta\,\mathrm{d}\phi^2 + \mathrm{d}\theta^2)\right], \tag{1.5}$$

where κ is the *curvature parameter* and $a(t)$ is the *expansion* (or *scale*) *factor*, which is a function of time and determines the overall scale of the spatial part of the metric.

Let us consider a free particle at rest in the origin of the comoving coordinates system at some time. Because of the absence of a preferential direction, no acceleration can be induced to a particle which moves along a geodesic. Thanks to the required homogeneity, each world line $x^i = const$ is a geodesic.

In a universe where the metric (1.5) is valid, the spatial hypersurface curvatures can be either positive, zero or negative, depending on the value of κ. This parameter can be either *(i)* $\kappa = +1$ (*closed universe*) which means the universe has a positive Gaussian curvature $1/r^2$ and the geometry is hyperspherical (pseudo-Euclidean), *(ii)* $\kappa = 0$ (*flat universe*), the space has an infinite curvature radius and the usual Cartesian geometry is valid (Euclidean), and *(iii)* $\kappa = -1$ *(open universe)*, the universe has a negative Gaussian curvature, leading to a hyperbolic geometry (pseudo-Euclidean). For the flat $\kappa = 0$ space the homogeneity and isotropy are obvious properties, as we have a familiar Euclidean 3D space.

The $\kappa = +1$ space is closed, meaning that its volume is finite, as its geometry is analogous to that of a sphere in a 3-dimensional space, and its set of boundaries is empty. Introducing a new coordinate defined as

$$\chi = F(r) := \int \frac{\mathrm{d}r}{\sqrt{1-\kappa r^2}} = \begin{cases} \sin^{-1} r & (\text{for } \kappa = +1) \\ r & (\text{for } \kappa = 0) \\ \sinh^{-1} r & (\text{for } \kappa = -1) \end{cases} , \tag{1.6}$$

it is possible to re-write the spatial part of the metric (1.5) in terms of (χ, θ, ϕ):

$$\mathrm{d}l^2 = a^2 \left[\mathrm{d}\chi^2 + f^2(\chi)(\mathrm{d}\theta^2 + \sin^2\theta \, \mathrm{d}\phi^2) \right] , \tag{1.7}$$

where

$$f(\chi) = \begin{cases} \sin\chi & (\text{for } \kappa = +1) \\ \chi & (\text{for } \kappa = 0) \\ \sinh\chi & (\text{for } \kappa = -1) \end{cases} . \tag{1.8}$$

For $\kappa = +1$, equation (1.7) represents a 3-dimensional sphere of radius a, embedded in a 4-dimensional pseudo-Euclidean space:

$$x_0^2 + x_1^2 + x_2^2 + x_3^2 = a^2 ,$$

where (x_0, x_1, x_2, x_3) are the Cartesian coordinates. It is also possible to define a set of angular coordinates on the sphere using the relations

$$\begin{cases} x_0 = a \, \sin\chi \\ x_1 = a \, \cos\chi \sin\theta \sin\phi \\ x_2 = a \, \cos\chi \sin\theta \cos\phi \\ x_3 = a \, \cos\chi \cos\theta \end{cases} , \tag{1.9}$$

so that the line element on the sphere

$$\mathrm{d}L^2 = \mathrm{d}x_0^2 + \mathrm{d}x_1^2 + \mathrm{d}x_2^2 + \mathrm{d}x_3^2$$

becomes

$$\mathrm{d}L^2_{sphere} \equiv a^2 \left[\mathrm{d}\chi^2 + \sin^2\chi(\mathrm{d}\theta^2 + \sin^2\theta \, \mathrm{d}\phi^2) \right] ,$$

which is identical to the (1.7) with $\kappa = +1$. The space of the $\kappa = 1$ metric is completely covered, at a fixed time, by the range of angles

$$0 \leq \chi \leq \pi \quad , \quad 0 \leq \theta \leq \pi \quad , \quad 0 \leq \phi \leq 2\pi ,$$

and, as already said, it has a finite volume:

$$V = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \mathrm{d}\theta \int_0^{\pi} \mathrm{d}\chi \sqrt{g} = a^3 \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \sin\theta \, \mathrm{d}\theta \int_0^{\pi} \sin^2\chi \, \mathrm{d}\chi = 2\pi^2 a^3 \ .$$

The surface of this volume can be parameterized in terms of χ,

$$S = 4\pi a^2 \sin^2\chi \ ,$$

and varies as χ changes in time. Firstly it increases, it reaches a maximum value at $\chi = \pi/2$ and then it decreases. In such a space, $S(\chi)$ is systematically larger than in a Euclidean space, the sum of the internal angles of a triangle is larger than π.

Analogues equations for the $\kappa = -1$ space can be obtained by substituting all the trigonometric functions with the respective hyperbolic functions. In this case the space is open and infinite and its properties are more like that of a Euclidean space. The sum of the angles of a triangle is less than π and $S(\chi)$ is systematically smaller than in the Euclidean space.

1.1.1 The Hubble Law and the reddening of distant objects

With a FRW metric, the *proper distance*, d_{pr}, between two events is

$$d_{pr} = \int_0^r \frac{a(t)\,\mathrm{d}r'}{\sqrt{1-\kappa r'^2}} = a(t)F(r) \ , \tag{1.10}$$

and depends on time through the scale parameter $a(t)$. It is also possible to define a *comoving distance*

$$d_c := d_{pr}(t_0) = a(t_0)F(r) = a_0 F(r) \ , \tag{1.11}$$

which is the value of the proper distance at the present time, $t = t_0$. The two measures are related by the equation

$$d_{pr} = \frac{a(t)}{a_0} d_c \ . \tag{1.12}$$

To see how fast two points drift apart, as a consequence of the Universe expansion, let us do the derivative of equation (1.10) with respect to time:

$$v = \frac{\mathrm{d}\, d_{pr}}{\mathrm{d}\, t} = \frac{\mathrm{d}}{\mathrm{d}\, t}\big[a(t)F(r)\big] = \frac{\mathrm{d}\, a(t)}{\mathrm{d}\, t}F(r) + a(t)\frac{\mathrm{d}\, F(r)}{\mathrm{d}\, t} = \dot{a}(t)F(r) + a(t)\dot{F}(r) \ ,$$

where $\dot{F} = 0$, so that it becomes

$$v = \frac{\dot{a}(t)}{a(t)} d_{pr} = H(t) d_{pr} \ , \tag{1.13}$$

which is the well known *Hubble Law*, where we have defined the *Hubble Parameter* $H(t) \equiv \dot{a}/a$. This parameter measures the isotropic expansion velocity and has the same value for the whole Universe at a given cosmic time. The global motion of objects in the Universe with respect to each other is called *Hubble Flow*. It has to be noticed that $H(t)$ is measured in s^{-1} and, by taking the oversimplified assumption

that for all its history the Universe has expanded with the same rate, its inverse is a rough estimate of the time passed from the beginning of the expansion.

Since on large scales, objects are affected by this expansion, their motion will result in a reddening of their observed spectrum. Formally, this reddening is due to the expansion of the electromagnetic wavelength with the space in which it is embedded. Let us consider once again the FRW metric in polar coordinates, eq. (1.5), with $\mathrm{d}\phi = \mathrm{d}\theta = 0$. Massless particles have $\mathrm{d}s^2 = 0$ by definition. Photons move along null geodesic with the expansion of the Universe. If a photon is emitted from a point with comoving coordinates $(r, 0, 0)$, at some (earlier) time t_e, and it is observed at time t_0, integrating the metric along the path we have

$$\int_{t_e}^{t_0} \frac{c\,\mathrm{d}t}{a(t)} = \int_0^r \frac{\mathrm{d}r}{\sqrt{1-\kappa r^2}} =: F(r) \ . \tag{1.14}$$

When a second photon emitted from the source at $t'_e = t_e + \delta t_e$ reaches the observer at $t'_0 = t_0 + \delta t_0$, given that $F(r)$ does not change because comoving coordinates have been assumed, eq. (1.14) can be written

$$\int_{t'_e}^{t'_0} \frac{c\,\mathrm{d}t}{a(t)} = F(r) \ . \tag{1.15}$$

If δt_e and, therefore, δt_0 are small, the equivalence between the two equations (1.14) and (1.15) implies that $\delta t_0/a_0 = \delta t'_0/a(t'_0) = \delta t/a(t)$. This affects the frequencies of the emitted and received light which depend on δt_0 and δt_e, respectively. In particular $\nu_e a(t_e) = \nu_0 a_0$ or, equivalently,

$$\frac{a(t)}{\lambda_e} = \frac{a_0}{\lambda_0}$$

from which

$$1 + z \equiv \frac{a_0}{a(t)} \ . \tag{1.16}$$

In the last equation, we have defined a new variable which is more directly observable than the expansion parameter a. This variable is called *redshift* and is usually implemented in equations to account for the expansion of the Universe. It is defineded as

$$z \equiv \frac{\lambda_0 - \lambda_e}{\lambda_e} \ .$$

Using the same line of thought for a particle instead of a photon, it is possible to demonstrate that the Hubble law is naturally predicted by the FRW metric.

1.1.2 Friedmann Equations

Without making any assumption on the geometry of the Universe, we apply the FRW metric (1.5) in the field equations (1.2). Out of the 16 equations of the system only two are independent. These are the evolution equations for $a(t)$ and describe the dynamic evolution of the Universe:

$$\ddot{a} = -\frac{4\pi}{3} G \left(\rho + \frac{3p}{c^2} \right) a \ , \tag{1.17}$$

$$\dot{a}^2 + \kappa c^2 = \frac{8\pi}{3} G\rho a^2 \ , \tag{1.18}$$

where a dot indicates a derivation with respect to the cosmic time t. Equation (1.17) and (1.18) are the *First* and the *Second Friedmann Equations*, respectively.

Each of these two equations can be recovered from the other one by applying the *adiabatic condition*,

$$\mathrm{d}\mathcal{U} = -p \ \mathrm{d}V \ ,$$

where $\mathcal{U}$ is the internal energy of the Universe. As long as the Universe is considered as a closed system, it expands and evolves without losing energy. So this latter equation means that the Universe expands adiabatically. It can be rewritten in the following equivalent forms:

$$\mathrm{d}(\rho c^2 a^3) = -p \, \mathrm{d}a^3 \tag{1.19}$$

$$\dot{\rho} a^3 = \frac{\mathrm{d}}{\mathrm{d}\, t}\left[a^3(\rho c^2 + p)\right] \tag{1.20}$$

$$\dot{\rho} + 3\left(\rho + \frac{p}{c^2}\right)\frac{\dot{a}}{a} = 0 \tag{1.21}$$

The density, ρ, in these equations has to be considered as the sum of all the densities of the components of the Universe.

1.1.3 Friedmann Models

The Friedmann Equations (1.17) and (1.18) define a set of cosmological models which have been studied and improved for almost a century since their first proposal (Friedmann, 1922). In universes where eq. (1.17) is valid, the evolution cannot be static ($\ddot{a} = 0$) unless

$$\rho = -3\frac{p}{c^2} \ , \tag{1.22}$$

which implies that either density or pressure must be negative. If this is not the case, then the Universe will eventually collapse under its own self-gravity. Since Einstein was firmly convinced that the Universe must be static, he modified the eq. (1.2) by introducing the *cosmological constant* Λ:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R - \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \ . \tag{1.23}$$

With an appropriate choice of Λ, one can obtain a static, even though unstable, cosmological model.

When Hubble finaly discovered the expansion of the Universe, Einstein regretted the introduction of this constant but, at the end of the 20th century, Λ-models made a come back as a result of the discovery of the accelerated expansion of the Universe (Perlmutter et al., 1999; Riess et al., 1998; Schmidt et al., 1998). The cosmological constant acts as a "repulsive force" which counteracts the attractive pull of gravity. The scientific community now believes that the introduction of Λ in eq. (1.17) is a valid method to account for the observed cosmic acceleration. This acceleration is still not well understood, as it does not exist a fully consistent and natural explanation for its origin, and it is generally labelled as the resulting effect of some kind of *Dark Energy*.

To solve eq.s (1.17) and (1.18), that depend on three parameters, $a(t)$, ρ and p, it is necessary to introduce an equation of state for the fluid composing the Universe. The mean free path of a particle is smaller than the typical scale of macro-physics phenomena which involve structure formation and evolution in the Universe. It is then acceptable to adopt the perfect fluid approximation when studying cosmology. The equation of state of such a fluid can be expressed by the general form

$$p = w\rho c^2 \ , \tag{1.24}$$

where w is defined so that the sound speed is

$$c_s = \left(\frac{\partial\, p}{\partial\, \rho}\right)_S^{1/2} = c\sqrt{w} \ .$$

This operative definition also sets the Zel'dovich range $0 \leq w < 1$ in which w is physically acceptable[1]. Different components have different values of w. The "ordinary" components can be divided into two families, relativistic and non-relativistic. The equation of state of a non-degenerate ultrarelativistic fluid in thermal equilibrium has $w = 1/3$. Even though a fluid with a similar equation of state is usually referred to as a *radiative fluid*, it can be used to describe relativistic particles of any form (e.g. photons, neutrinos). On the other hand, $w = 0$ is used to represent any material whose contribution to pressure can be neglected (i.e. *dust*). This is a good approximation to the behaviour of any form of non-relativistic fluid, such as ordinary matter. For what concerns Λ, by analogy to the other fluids and following eq. (1.23), it is defined to behave as a perfect fluid with $w = -1$.

Therefore the total energy-momentum tensor of eq. (1.2) is given by the sum of all the components with different state equation:

$$T_{\mu\nu} \equiv \sum_i T^{(i)}_{\mu\nu} \ .$$

Combining eq.s (1.19) and (1.24), the following dependence, which describes how the densities of the different components vary through cosmic time, can be derived:

$$\rho_w \propto a^{-3(1+w)} \propto (1+z)^{-3(1+w)} \ . \tag{1.25}$$

As a result of eq. (1.25) we can assert that, in different cosmic epochs, different components have prevailed over the others.

Eq. (1.18) can be re-written to derive the curvature of the assumed FRW metric

$$\frac{\kappa}{a^2} = \frac{1}{c^2}\left(\frac{\dot{a}}{a}\right)^2\left(\frac{\rho}{\rho_c} - 1\right) \ , \tag{1.26}$$

where

$$\rho_c(t) := \frac{3}{8\pi G}\left(\frac{\dot{a}}{a}\right)^2 = \frac{3\,H^2(t)}{8\pi G} \tag{1.27}$$

is called the *critical density* and expresses the density a Friedmann Universe must have to be flat ($\kappa = 0$). It can be used to define the dimensionless *density parameter*

$$\Omega(t) = \frac{\rho(t)}{\rho_c(t)} \ , \tag{1.28}$$

[1] Nevertheless, physically significant cases with $w < 0$ exist (e.g. during inflation).

where the density ρ can be either that of a single component (Ω_w) or of the sum of all of them (Ω_T). In a flat universe $\Omega_T = 1$, while in open or closed universes it is < 1 or > 1, respectively.

From the condition of adiabatic expansion one can derive the evolution of the density parameter of each component, keeping in mind that the equation of state parameter w has to vary accordingly:

$$\Omega_w(z) = \Omega_{0\,w}(1+z)^{3(1+w)} \ . \tag{1.29}$$

Equation (1.29) also justifies the need to constrain the values of cosmological parameters at present time. All of our models are obtained as approximations around the point in space-time where more information is available: the present time. As long as an evolving model is assumed, the time is relevant, while for the Copernican principle the position in space is not.

It is useful to re-write the Second Friedmann Equation in terms of Ω, H and z, which are more representative parameters of the observable universe. Eq. (1.16) and the definitions in eq.s (1.28) and (1.13) can be combined with eq. (1.18), becoming

$$H^2(z) = H_0^2(1+z)^2\left[1 - \sum_i \Omega_{0\,w_i} + \sum_i \Omega_{0\,w_i}(1+z)^{1+3w_i}\right] \ , \tag{1.30}$$

where $H(z) = \dot{a}/a$ is the Hubble parameter at a generic time t, $\sum_i \Omega_{0\,w_i} = \Omega_T$ is the sum of all the i-th components, labelled with their value of w_i, and where the subscript 0 indicates the measure of the parameter at $z = 0$. The quantity $1 - \Omega_T$ is then a measure of curvature.

Let us now consider eq. (1.17) instead, and re-write it by applying the eq. (1.2). It becomes

$$\ddot{a} = -\frac{4\pi}{3} G\rho(1+3w)\, a \ . \tag{1.31}$$

This last equation implies that, if w fits in the Zel'dovich interval, $\ddot{a}$ has always been less than zero and so the graph of $a(t)$ is concave. While the existence of the Hubble Law implies that $H(z)$ is positive, eq. (1.30) states that $a(t)$ grows monotonically. All these considerations allow us to say that in a Friedmann Universe, at a certain finite time in the past ($z_{BB} > 0$) $a(z_{BB})$ must have been equal to zero. At this point a singularity both in density and in the Hubble parameter may exist in all the models of the Universe filled with fluids having $-\frac{1}{3} \leq w \leq 1$. The scale factor of the Universe vanishes, while the density and the expansion speed diverge, leading to the BB scenario.

Note that the expansion of the universe emerging from the BB model is a result of the initial conditions describing a homogeneus and isotropic Friedmann Universe, and it is not due to pressure, which here acts to decelerate the expansion (Coles and Lucchin, 2002).

1.1.4 Flat VS Curved Universes

It has already been demonstrated that at different cosmic epochs different components prevail over the others. It is then possible to roughly approximate our Universe as if it was made only by one component as long as it is studied in periods

far from equivalences. The general solution for a mono-component fluid in a flat universe ($\Omega_0 = 1$) is called *Einstein-de Sitter Model* (*EdS*). In this model the eq. (1.30) reduces to

$$H(z) = H_0(1+z)^{\frac{3(1+w)}{2}} \ . \tag{1.32}$$

Some useful dependencies derived by applying the EdS model are reported in Table 1.1 in the general case and in two particular cases, a dust universe and a perfect radiative fluid.

General	$w = 0$	$w = 1/3$
$a(t) = a_0\left(\frac{t}{t_0}\right)^{\frac{2}{3(1+w)}}$	$a(t) \propto t^{2/3}$	$a(t) \propto t^{1/2}$
$t = t_0(1+z)^{\frac{2}{3(1+w)}}$	$t \propto (1+z)^{-3/2}$	$t \propto (1+z)^{-2}$
$H(t) = \frac{2}{3(1+w)}t^{-1}$	$H(t) = \frac{2}{3t}$	$H(t) = \frac{1}{2t}$
$t_0 = \frac{2}{3(1+w)}\frac{1}{H_0}$	$t_0 = \frac{2}{3H_0}$	$t_0 = \frac{1}{2H_0}$
$\rho = \frac{1}{6\pi G(1+w)^2}\frac{1}{t^2}$	$\rho = \frac{1}{6\pi G}\frac{1}{t^2}$	$\rho = \frac{3}{32\pi G}\frac{1}{t^2}$

Table 1.1: Dependencies of the EdS model in three different cases: general case, dust universe ($w = 0$), radiative fluid ($w = 1/3$).

In the general case we have $\Omega_0 \neq 1$ but, at high redshift, the term $(1-\Omega_0)$ inside parentheses in eq. (1.30) is negligible with respect to the other term, which grows faster with increasing redshift. If we define the critical value z^* so that

$$|1-\Omega_0| = \Omega_0(1+z^*)^{1+3w} \ , \tag{1.33}$$

inside the interval $z > z^*$, eq. (1.30) will reduce to

$$H(z) = H_0\sqrt{\Omega_0}(1+z)^{\frac{3(1+w)}{2}} \ ,$$

which differs from eq. (1.32) by only a factor $\sqrt{\Omega_0}$. A curved universe will then behave as an EdS universe until $z \gg z^*$, meaning that it is acceptable to consider the young universe as flat. For the range of values of Ω_0 suggested by present data, this approximation is valid up to $z \approx 10$.

1.2 The Standard Cosmological Model

As it has already been said, mono-component models are acceptable approximations but, in order to have a more realistic representation of the Universe, all of its components must be taken into account. The *flat-*Λ*CDM* cosmological model takes its name from the assumed flat geometry and its major components at $z = 0$, namely the cosmological constant and a dust component called Cold Dark Matter (CDM). It is supported by a set of observable cosmological parameters which are

constrained by constantly improving experiments. Our Universe is composed by a radiative fluid together with different forms of dark matter (DM) and energy (DE).

The *radiative fluid* component is made up by all the particles that are relativistic at a given time. At $z = 0$ they are basically photons and possibly neutrinos, but their contribution is negligible. The density parameter of photons can be estimated by measuring the energy density of the CMB, which gives us $\Omega_{0\,r} \approx 10^{-5}$.

The *matter component* is divided in baryonic (i.e. the "ordinary" matter) and non-baryonic (i.e. DM). The first interacts with photons and is then observable, while the latter only interacts via gravity, making its direct observation impossible with standard methods based on the emission of light. Even though we cannot see it directly there is strong evidence of its existence, e.g. by the gravitational anomalies found in galaxy clusters. In a brief publication Ostriker (1999) reported the importance of "an extraordinarily prescient paper" by Fritz Zwicky published in October of 1937 (Zwicky, 1937) in which the bases of our current believes on matter were founded. The ideas of Zwicky were based on a set of observations of the Coma cluster started 4 years before (Zwicky, 1933) and considerations on the virial theorem. He had already noticed the existence of a discrepancy in the mass of the cluster estimated using luminosities and the measured velocity dispersion of the galaxies which make up the cluster. He found that the mass-to-light ratio of the system was two orders of magnitude greater than that of the solar neighborhood and concluded that, if his results were to be confirmed both by more data on velocity dispersions and by other means[2], "dark matter is present in a much greater density than luminous matter".

Since Zwicky's discoveries and later works, from the 1970s a lot has been done to understand the nature of this elusive component of our Universe. The distribution of DM has been studied from the scale of galaxies, through the study of the rotation curves of disk galaxies and of the velocity dispersion of elliptical galaxies, to the larger cosmological scales, in the CMB's spectrum. All these data have confirmed the existence of a ruling component in the matter distribution which is nothing like what we are used to. A large zoology of possible explanations and candidates for DM has been proposed so far. Even though the majority of the scientific community accepts that it should be made of *weakly interacting massive particles* (*WIMPs*) within the Standard Cosmological Model, other explanations involve both modifications of the theory of gravitation or exotic particles. For a (almost) complete review of what is currently known about DM one may refer to Bertone et al. (2005) and Kuhlen et al. (2012).

From the early 1990s, a tension in the model of flat matter-dominated Universe came out when measuring galaxy correlations on large scales (Maddox et al., 1990), leading to a revived interest for the cosmological constant Λ, initially proposed by Einstein at the dawn of modern cosmology. This culminated with the discovery of the accelerated expansion at the end of the decade, which led to the need of a modification in the cosmological models through the introduction of the DE component. Despite the refinements under which the standard ΛCDM model has gone during all the 21st century and the remarkable success it achieved, its theoretical roots

[2]He also suggested some other experiments which could have been done to confirm his theories, such as gravitational lensing studies. In his paper of 1933 he also used the term "Cold Dark Matter" for the very first time (*dunkle [kalte] matter*).

are yet poorly defined (Baldi, 2012b, and references therein). The mostly accepted explanation (i.e. the cosmological constant) is not an explanation, and supposes the existence of a physical mechanism not yet discovered capable of causing the observed acceleration, but a wide range of alternatives has been proposed.

In a Friedmann Universe with FRW metric (1.5) Einstein's field equations are written as in eq. (1.23). If we consider $T_{\mu\nu} = 0$, eq. (1.23) describes the empty space, and so $\Lambda \neq 0$ is a measure of the *energy of vacuum* which acts as a repulsive force, counteracting that of gravity. Eq. (1.23) can be re-written to resemble eq. (1.2) if we enclose the energy of vacuum in the energy-momentum tensor eq. (1.3) is then re-written as

$$\tilde{T} = -p\, g_{\mu\nu} + (p + \rho c^2) u_\mu u_\nu + \Lambda g_{\mu\nu} = -\tilde{p}\, g_{\mu\nu} + (\tilde{p} + \tilde{\rho} c^2) u_\mu u_\nu \ , \tag{1.34}$$

where $\tilde{p} = p + p_\Lambda$ and $\tilde{\rho} = \rho + \rho_\Lambda$ with

$$p_\Lambda := -\frac{\Lambda c^4}{8\pi G} \quad \text{and} \quad \rho_\Lambda := \frac{\Lambda c^2}{8\pi G} \ . \tag{1.35}$$

In the flat-ΛCDM model, Λ is a true constant and the equation of state then leads to $w = -1$. From eq. (1.29) it is possible to see that, while matter ($w = 0$) and radiation ($w = 1/3$) have densities that evolve with redshift, respectively as $\propto (1+z)^3$ and as $(1+z)^4$, and since the energy density of DE is constant throughout the whole expansion history of the Universe (eq. (1.35)), Ω_Λ ($w = -1$) only depends by the critical density evolution, thus by the evolution of the scale factor.

Let us now define a decelleration parameter,

$$q(t) := -\frac{a(t)\ddot{a}(t)}{\dot{a}^2(t)} \equiv -\frac{1}{H^2(t)} \frac{\ddot{a}(t)}{a(t)} \ , \tag{1.36}$$

which expresses a dimensionless measure of the second derivative of the scale factor $a(t)$.

In a mono-component matter Universe with $w = 0$ and $p = 0$, the First Friedmann Equation with definition (1.29) leads to $q = \Omega_M/2$, which implies that $q > 0$ independently of time. In such a Universe the expansion is decelerated, that is not compatible with present observations. Introducing the cosmological constant in the First Friedmann Equation one obtains a multi-component model with

$$\ddot{a} = -aH^2 \frac{\Omega_M}{2} + aH^2 \Omega_\Lambda \ ,$$

where $\Omega_\Lambda \equiv \frac{\Lambda c^2}{3H^2}$. Using the definition (1.36), it becomes:

$$q = \frac{\Omega_M}{2} - \Omega_\Lambda \ .$$

If the expansion is accelerated, it must be $q < 0$ and so the condition

$$\Omega_\Lambda > \frac{\Omega_M}{2}$$

has to be satisfied. With the currently accepted values of the two parameters $\Omega_M \approx 0.3$ and $\Omega_\Lambda \approx 0.7$, it results that $q \approx -0.55$.

Our Universe is approximately flat ($\Omega_M + \Omega_\Lambda \approx 1$), as also confirmed by recent measurements, but the evolution of the scale parameter cannot be that described previously. In the earlier moments of its evolution, each universe evolves as EdS ($\ddot{a} < 0$) but, because of the observed later time acceleration ($\ddot{a} > 0$), the graph has to present a flex where $\ddot{a}(z_f) = 0$. This change in the behaviour of the model is found to occur at $z_f \approx 0.68$.

Using eq. (1.29), if we impose $\Omega_M(z_{eq}) = \Omega_\Lambda(z_{eq})$ we can obtain the *redshift of equivalence*, $z_{eq} \approx 0.33$. Since this result implies that the Λ-component's contribution became relevant only at recent times, it is refered to as the *Coincidence Problem*.

1.3 Jeans Theory and Expanding Universes

The anisotropies in the CMB are a fossil of the density perturbations in the cosmic fluid at the time of recombination. From the amplitude of the fluctuations in the temperature power spectrum it is possible to derive the amplitude

$$\frac{\delta T}{\overline{T}} \approx 10^{-5}$$

for primordial perturbations, where $\overline{T}$ is the mean black body temperature of the CMB. However, if one considers what is observed in the present Universe in terms of galaxies and clusters at megaparsec scales, and in terms of voids, filaments and sheets on scales over 10 Mpc, the Universe appears to be quite inhomogeneous and shows the charachteristics of a highly non-linear evolution. In the largest surveys it is however possible to recover the homogeneity on scales over hundreds of Mpc, and fluctuations arrange around amplitudes of approximately 10^2.

The non-linear structure of the Universe is a consequence of *gravitational instability*, which makes the small initial perturbations grow. Matter is attracted by overdense regions and evacuated from underdense ones, thus amplifying already existing inhomogeneities. This process is analytically described by the Jeans theory which makes use of the Newtonian theory of gravity to find solutions which aim at the description of how fast an initial perturbation has to grow to reproduce the inhomogeneities observed today. The results derived in this theory are applicable only to non-relativistic matter on scales not exceeding the Hubble horizon (i.e. the sphere causally connected with an observer). The *Hubble horizon* is defined as

$$R_H := a(t) \int_{t_{BB}}^{t} \frac{c\,\mathrm{d}t'}{a(t')} \,, \tag{1.37}$$

where time $t_{BB} = 0$ is the instant when the expansion started. On scales larger than the horizon, gravity is the only interaction which plays a role, but it has to be treated by means of a fully relativistic theory. On the other hand, at sub-horizon scales, the Jeans theory provides a reliable theory of perturbations in an expanding Universe. We will first present the general arguments of this theory in the simplyfied case of a static Universe, even though it is known that this is not the case, and then we extend the solutions to an expanding Universe.

1.3.1 Instability in a Static Universe

We assume a homogeneous and isotropic background with constant, time-independent matter density, $\rho(\mathbf{x},t) = const$, in a static non-expanding Universe[3]. Let us consider the following system of hydrodinamic equations in adiabatic conditions:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \qquad \text{mass conservation} \tag{1.38a}$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} = -\frac{1}{\rho}\nabla\rho - \nabla\Phi \qquad \text{Euler equation} \tag{1.38b}$$

$$\nabla^2\Phi = 4\pi G\rho \qquad \text{Poisson equation} \tag{1.38c}$$

$$p = p(\rho, S) = p(\rho) \qquad \text{equation of state} \tag{1.38d}$$

$$\frac{\mathrm{d}\, S}{\mathrm{d}\, t} = 0 \qquad \text{adiabatic condition} \tag{1.38e}$$

where $\mathbf{v}$ is the velocity vector of a fluid element, Φ the gravitational potential, p the pressure and S the entropy. Eq.s (1.38) are solved for a background (B) solution

$$\rho_B = const \qquad p_B = const$$
$$\Phi_B = const \qquad \mathbf{v} = 0\ .$$

Introducing a small perturbation δA, where A is one of the five variables of the system, then substituting this term in the system (1.38), one obtains the hydrodinamic system for a fluctuation in the density field. If we define the adimensional density fluctuation

$$\delta(\mathbf{x},t) := \frac{\delta\rho(\mathbf{x},t)}{\rho_B}\ , \tag{1.40}$$

and require it to be small, $\delta \ll 1$, the system (1.38) can be linearized. The solving equation for density is a differential equation of $\delta(\mathbf{x},t)$ which, in the Fourier space, has the form

$$\ddot{\delta}_{\mathbf{k}} + (k^2c_s^2 - 4\pi G\rho_B)\delta_{\mathbf{k}} = 0\ , \tag{1.41}$$

where $k = |\mathbf{k}|$, $\delta_{\mathbf{k}} = \delta_k(t)$ is the Fourier transform of $\delta(\mathbf{x},t)$, $c_s^2 := (\partial p/\partial\rho)_S$ is the sound speed and a dot denotes a derivative with respect to time t. Eq. (1.41) has two independent solutions $\delta_{\mathbf{k}} \propto \exp(\pm i\,\omega(k)t)$, where

$$\omega(k) = \sqrt{k^2c_s^2 - 4\pi G\rho_B}\ . \tag{1.42}$$

The behaviour of these adiabatic perturbations depends crucially on the sign of the expression under the square root in (1.42). Defining the *Jeans length*,

$$\lambda_J := \frac{2\pi}{k} = c_s\left(\frac{\pi}{G\rho_B}\right)^{1/2}\ , \tag{1.43}$$

as the scale corresponding to $\omega(k_J) = 0$, the perturbation propagates as a sound wave when $\lambda < \lambda_J$ with a phase velocity $c_{ph} = \omega/k$. This velocity tends to the sound speed in the limit $k \gg k_J$ or, equivalently, for small scales $\lambda \ll \lambda_J$.

[3] Even though the assumption of matter at rest causes an inconsistency in the Poisson equation, as it would require a null gravitational field, it will be automatically solved by the introduction of an expansion factor.

On the other hand, for very large scales, gravity dominates and the fluctuation grows rapidly as an exponential. This represents the phenomenon of the *Jeans instability*: small density fluctuations evolve according to their scale and, if they can grow, they do it exponentially.

1.3.2 Instability in an Expanding Universe

In an expanding homogeneous and isotropic Universe, the background density is a function of time, $\rho_B = \rho_B(t)$, and the velocities are not null but obey the Hubble law, $\mathbf{v}_B = H(t)\mathbf{x}$. Substituting these expressions into (1.38a) one obtains the equation

$$\dot{\rho}_B + 3H(t)\rho_B = 0 \ , \tag{1.44}$$

which states that the total mass of non-relativistic matter is conserved. Introducing a fluctuation in equations (1.38) and accounting for a double component velocity, $\mathbf{u} := \dot{\mathbf{x}} = H(t)\mathbf{x}+\mathbf{v}$, where the first term accounts for the Hubble flow and the second describes the peculiar velocity of the fluid, it is possible to recover the solution for every Fourier mode $\delta = \delta_{\mathbf{k}}(t)\exp(i\mathbf{kx})$:

$$\ddot{\delta}_k + 2H(t)\dot{\delta}_k + (k^2c_s^2 - 4\pi G\rho_B)\delta_k = 0 \ . \tag{1.45}$$

The last equation is the *dispersion relation*, where the term $2H(t)\dot{\delta}_k$ represents the Hubble friction, while $k^2c_s^2\delta_k$ accounts for the characteristic velocity field of the fluid, both these terms oppose to the growth of fluctuations.

The eq. (1.45) is a second-order differential equation for $\delta(\mathbf{x}, t)$ with growing and decaying mode solutions:

$$\delta(\mathbf{x}, t) = A(\mathbf{x})D_+(t) + B(\mathbf{x})D_-(t) \ , \tag{1.46}$$

where A and B are functions of the comoving coordinate and D_+ and D_- are functions of the time coordinate. A growing solution of eq. (1.45) has scale bigger than the Jeans scale (expressed in physical units)

$$\lambda_J = \frac{2\pi a}{k_J} = c_s\sqrt{\frac{\pi}{G\rho_B}} \ ,$$

and depends on the cosmological model through $H(t)$.

We can rewrite the eq. (1.45) for δ as an equation for the growth factor

$$\ddot{D} + 2H(z)\dot{D} - \frac{3}{2}\Omega_{m,0}H_0^2(1+z)^3D = 0 \ ,$$

where the time dependence has been substituted with a dependency on redshift.

If we apply the dependencies in Table 1.1 for an EdS Universe with $\Omega_m = 1$, the growing and decaying solutions are

$$D_+(t) \propto t^{2/3} \propto a(t)$$

and

$$D_-(t) \propto t^{-1} \propto a^{-3/2} \ ,$$

respectively.

A more complicated algebraic expression for $D_+(t)$ in the case of $\Omega_m < 1$, $\Omega_\Lambda = 0$ is given in Peebles (1980).

For a flat Universe with a cosmological constant, the solution to the differential equation can be written in integral form

$$D_+(z) = \frac{H(z)}{H_0} \int_z^\infty \frac{\mathrm{d}z'(1+z')}{H^3(z')} \left[\int_0^\infty \frac{\mathrm{d}z'(1+z')}{H^3(z')} \right]^{-1} ,$$

where the factor in brackets makes the normalization $D_+ = 1$ at $z = 0$.

For other equations of state of DE, or for more components (curvature, radiation), no general integral expression exists, and the differential equation must be solved directly.

To a good approximation, $\mathrm{d}\ln D_+ / \mathrm{d}\ln a \approx \Omega_m^{0.55}$, implying $D_+ \propto a$ while $\Omega_m \approx 1$ and a slowing of growth as Ω_m falls below unity.

If we assume that density fluctuations appear at some very early time, then at a much later time we care only about the growing mode $D_+(t)$, which we can just call $D(t)$. Thus

$$\delta(\mathbf{x}, t) = \delta(\mathbf{x}, t_i) \frac{D(t)}{D(t_i)} . \tag{1.47}$$

As long as $\delta \ll 1$, density fluctuations simply "grow in place" in comoving coordinates.

1.3.3 Primordial Density Fluctuations

The theory developed in the previous Section applies to a single Fourier mode of the density field $\delta(\mathbf{x}, t) = D_+(t)\delta(\mathbf{x})$. In a realistic representation of the growth of structures a superposition of waves has to be considered. We model the density field as a superposition of Fourier modes with different amplitudes. This approach requires a statistical description of the initial perturbations and implies that any comparison between theory and observations will also have to be statistical.

The spatial Fourier transform of $\delta(\mathbf{x})$ is

$$\delta(\mathbf{k}) = \frac{1}{(2\pi)^3} \int \mathrm{d}^3\mathbf{x}^{-i\mathbf{k}\cdot\mathbf{x}} \delta\mathbf{x} . \tag{1.48}$$

We can define the *power spectrum* of the density field as the variance of the amplitudes at a given value of $\mathbf{k}$:

$$\langle \delta(\mathbf{k}) \delta^*(\mathbf{k}') \rangle = (2\pi)^3 P(k) \delta_D^{(3)}(\mathbf{k} - \mathbf{k}') , \tag{1.49}$$

where $\delta_D^{(3)}$ is the 3-dimensional Dirac delta function. In definition (1.49) the CP is expressed by $P(\mathbf{k}) = P(k)$. The power spectrum is related via a Fourier transform to its analogous quantity in the real space, the *correlation function*, $\xi(r)$, of $\delta(\mathbf{x})$:

$$\xi(r) = \frac{1}{(2\pi)^3} \int \mathrm{d}^3\mathbf{k}\, P(k) e^{i\mathbf{k}\cdot\mathbf{x}} , \tag{1.50}$$

where $\xi(r)$ is defined by

$$\langle \delta(\mathbf{x}) \delta(\mathbf{x}') \rangle = \xi(|\mathbf{x} - \mathbf{x}'|) = \xi(\mathbf{r}) = \xi(r) , \tag{1.51}$$

with r representing the distance.

The shape of the initial fluctuation spectrum is assumed to be imprinted in the Universe at some arbitrarily early time. Since the accepted inflation scenario states that the seeds of today's structures are generated by stochastic quantum fluctuations in a scalar field (i.e. the *inflaton*), the initial $P(k)$ should produce a power-law form

$$P(k) = Ak^n \ , \tag{1.52}$$

where the spectral index is generally assumed to be close to unity, $n = 1$ (Zeldovich, 1972). While the shape of the power spectrum may be fixed by these considerations, the amplitude is not and the gap should be plugged with observations. Nowadays the most reliable measure of A is obtained from the temperature fluctuations in the Cosmic Microwave Background (CMB).

In the real space the assumption of stochastical generation of fluctuations implies their amplitudes have a Gaussian distribution. Although the mean value of the perturbation $\delta(\mathbf{x}) \equiv \delta$ across the statistical ensemble is identically zero by definition, its mean square value, i.e. its *variance* σ^2, is not. It is possible to show that

$$\sigma^2 = \langle|\delta(\mathbf{x})^2|\rangle = \sum_{\mathbf{k}} \langle|\delta_{\mathbf{k}}|^2\rangle = \frac{1}{V_u}\sum_k \delta_k^2 \ , \tag{1.53}$$

where the average is taken over an ensemble of realisations of volume V_u. The quantity δ_k is related to the power spectrum by $\delta_k^2 = P(k)$. If we take the limit $V_u \to \infty$ and assume that the density field fulfils the CP, we find

$$\sigma^2 \to \frac{1}{2\pi^2}\int_0^\infty P(k)k^2 \, \mathrm{d}k \ . \tag{1.54}$$

In eq. (1.54) the variance σ^2 is defined for each point in space. The measure of σ^2 requires the reconstruction of the density field, which is far from trivial. In fact, especially in today's Universe, the observable mass tracers are not distributed smoothly across space.

Instead of using a punctual variance, it is convenient to describe the fluctuation field as a function of some resolution scale R. Let $\langle M\rangle$ be the mean mass found inside a spherical volume of radius R, then we can define a density fluctuation from discrete tracers as

$$\delta_M = \frac{M - \langle M\rangle}{\langle M\rangle} \ . \tag{1.55}$$

With this definition, and with that of eq. (1.53) we obtain the *mass variance*

$$\sigma_M^2 = \langle\delta_M^2\rangle = \frac{\langle(M - \langle M\rangle)^2\rangle}{\langle M\rangle^2} \ , \tag{1.56}$$

which is the convolution of the punctual variance with a window function W, since

$$\delta_M(\mathbf{x}) = \delta(\mathbf{x}) \otimes W(\mathbf{x}, R) \ . \tag{1.57}$$

From the last two equations, it is possible to render the dependency on the power spectrum in the definition of mass variance, passing in the limit expressed by eq. (1.54),

$$\sigma_M^2 = \frac{1}{(2\pi)^3}\int \mathrm{d}^3\mathbf{k}\langle P(k)\rangle \hat{W}^2(\mathbf{k}, R) \ , \tag{1.58}$$

where $\hat{W}$ is the Fourier-trasform of the window function and which, as it must be, is a function of R and therefore of M.

The meaning of the window function is in the implications of such an expression. Since the higher values of k tend to be averaged out within the window volume, σ_M^2 is dominated by perturbation components with wavelength $\lambda \approx k^{-1} > R$. The Zel'dovich spectrum (1.52) is a growing function of k, so waves with much larger λ contribute only a small amount.

1.3.4 Non-Linear Evolution

When the density contrast of the perturbations approaches unity ($\delta \approx 1$), the linear theory approximation cannot be applied anymore, since its basic assumption fades. It is then necessary to abandon the small-perturbations approximation and develop a non-linear theory for the evolution of anisotropies.

The formation of the cosmic web, consisting of haloes, voids, filaments and walls, is a natural outcome of the gravitational instability of these anisotropies. This scenario is favoured by an overwhelming body of observational and theoretical evidences.

After the linear regime breaks down, when δ becomes comparable to unity, a *quasi-linear* (or *weakly non-linear*) regime sets in. During this regime, even though the density contrast may remain less than unity, the perturbations distribution function starts to evolve into a non-Gaussian shape. The power spectrum changes by virtue of a complicated cross-talk between different wave-modes.

For this kind of problems, analytical methods for special cases are available, but a general solution is gained using numerical N-body experiments, which allow the anlysis of strongly non-linear regimes. Since a physical limit for the underdensity level is set at $\delta = -1$, voids always evolve as quasi-linear structures. We will present in Chapter 2 the theoretical basis for the statistical prediction of the void distribution, through an analytical solution of their non-linear evolution. We will then present in Chapter 3 how do N-body simulations work and which are the methods used in this work to verify that the predictions of the suggested model are correct.

Moreover, further into the non linear regime, bound structures form. The baryonic content in these objects eventually becomes important and the simple treatment presented in the previous Sections is no longer correct. The spatial distribution of galaxies may be very different from the distribution of the DM, even on large scales, as a result of hydrodynamical effects, star formation, heating and cooling of gas. It is both possible to develop semi-analytical models based on empirically derived relations to address this problem, or simply to assume that the point-like distribution of haloes, galaxies or clusters, $n(\mathbf{r}) = \sum_i \delta_D^{(3)}(\mathbf{r} - \mathbf{r}_i)$, bears a simple functional relationship to the underlying $\delta(\mathbf{r})$. It is generally believed that relative fluctuations in the object number counts and matter density fluctuations are proportional to each other, at least within sufficiently large volumes. This is expressed by a linear biasing prescription:

$$\frac{\delta n(\mathbf{r})}{\langle n \rangle} = b \frac{\delta \rho(\mathbf{r})}{\langle \rho \rangle} \ , \tag{1.59}$$

where b is what is usually called the *(linear) bias parameter*. With such an ansatz, the linear theory can be used on large scales to relate galaxy clustering statistics to those of the underlying density fluctuations, e.g. $\sigma^2_{gxy} = b^2\sigma^2_M$. We will get a deeper insight into these issues in Section 2.4.

Chapter 2

Cosmic Voids in the large scale structure

The cosmic web, consisting of haloes, voids, filaments and walls, is well predicted by the CDM model (e.g. Bond et al., 1998) and confirmed by large galaxy surveys (e.g. de Lapparent et al., 1986; Colless et al., 2003b; Alam et al., 2015; Guzzo and The Vipers Team, 2013). Voids emerge out of the density troughs in the primordial Gaussian field of density fluctuations. They have been shown to have a great potential for constraining DE and testing theories of gravity via several measurements. These measurements include: distance measurements via the Alcock-Paczyǹski Test (Ryden, 1995; Lavaux and Wandelt, 2012; Sutter et al., 2014b), weak gravitational lensing of voids (Krause et al., 2013; Melchior et al., 2014; Clampitt and Jain, 2015; Gruen et al., 2016; Sánchez et al., 2016), the signal of the Integrated Sachs-Wolfe (ISW) effect associated with voids (Sachs and Wolfe, 1967; Granett et al., 2008; Nadathur et al., 2012; Ilić et al., 2013; Flender et al., 2013; Cai et al., 2014; Planck Collaboration et al., 2014, 2015b; Aiola et al., 2015), void ellipticity as a probe for the DE equation of state (Lavaux and Wandelt, 2010; Sutter et al., 2015a; Bos et al., 2012; Li et al., 2012; Pisani et al., 2015), void abundances and profiles for testing theories of gravity and cosmology (Li et al., 2012; Massara et al., 2015; Barreira et al., 2015; Zivick et al., 2015; Cai et al., 2015; Lam et al., 2015; Clampitt et al., 2013; Li et al., 2012), coupled DE (Pollina et al., 2016), the nature of DM (Yang et al., 2015), baryon acoustic oscillations in void clustering (Kitaura et al., 2016), and redshift-space distortions in voids (Hamaus et al., 2015).

Despite their popularity and great potential as a cosmologica tool, a gap of knowledge between the evolution of individual voids through simulations and observations versus theory still persists. In particular, even though a statistical description of the void number density as a function of the size has been developed by Sheth and van de Weygaert (2004), it has been proven with simulations to lack of a validation (Nadathur and Hotchkiss, 2015; Sutter et al., 2012; Pisani et al., 2015; Colberg et al., 2005).

Blumenthal et al. (1992) studied analytically the evolution of (isolated) spherically symmetric voids. Using N-body simulations (Section 3.2) they showed that adjacent voids collide, producing thin walls and filaments as the matter between them is squeezed. The peculiar velocities perpendicular to the void walls are mainly tan-

gential motions which cause the evacuation of the separating thin wall, eventually resulting in the merger of neighbouring voids.

The time-scale on which the internal substructure of a void is erased is approximately the same as that when the void itself approches non-linearity. As this re-arrangement of structure progresses to ever larger scales, the same basic process of collision, evacuation and merging repeats. N-body simulations with more generic cosmological circumstances yelded similar results (van de Weygaert and van Kampen, 1993).

In this Chapter we present the theoretical foundations that leads to the development of a statistical model to predict voids distributions. Their study is approached with a scheme similar to that of overdensities: starting from the spherical evolution of perturbations in the primordial density field (Section 2.1), we present the *excursion sets formalism* in Section 2.2 which leads to the desired statistical distribution of voids in terms of their size (Section 2.3). We present both the "classical" Sheth and van de Weygaert model (Sheth and van de Weygaert, 2004, hereafter SvdW) and its modification to account for the peculiar characteristics of void evolution developed by Jennings et al. (2013) (hereafter JLH). Our work will provide a validation of the latter model, demostrating its applicability to the matter distribution predicted by classical cosmological models.

We will also present an insight on a possible extension of the JLH model to situations where the density field is not traced by the DM distribution but by a halo distribution in Section 2.4. This will provide a step forward the exploitation of the void size function in real situations, such as the galaxy distribution provided by large cosmological surveys.

2.1 Spherical evolution

As we have seen in Section 1.3, it is possible to define the evolution of a density perturbation in a linear regime, until the density contrast within it is $\delta(\mathbf{x}, t) \ll 1$. Nevertheless, structures in the Universe are found with a wide range of degrees of non-linearity.

An analytical solution for the non-linear evolution of an *isolated spherically symmetric* perturbation exists. This case is the simplest non-trivial model for the decoupling of an anysoptropy in the density field, from the general Hubble expansion (see e.g. Peebles, 1980).

Let us consider a spherical top-hat perturbation modelled as a set of concentric shells, perfectly uniform and without substructures. The radial distribution of the density field inside the shell does not affect its evolution, which only depends on the total energy of the perturbation.

At some initial time t_i, the density contrast field in the Universe, $\delta(\mathbf{x}, t_i)$, divides its volume into several *underdense*, $\delta(\mathbf{x}, t) < 0$, and *overdense*, $\delta(\mathbf{x}, t) > 0$, regions. It is reasonable to expect that a sufficiently high overdensity will eventually decouple from the Hubble flow and collapse. The net effect of gravity on an underdensity is instead repulsive, causing the region to expand faster than the Hubble flow, which, nevertheless, enhances this outward flow. The final result of this behaviour is the evacuation of matter from the region.

Thanks to the assumption of the CP it is reasonable to hypotize that each region in the Universe can be treated as an independent Friedmann Universe, until it evolves adiabatically. If the only relevant interaction is gravity, in our Newtonian dissertation a top-hat perturbation satisfies the last hypothesis until the total enclosed mass is conserved, i.e. until all the shells can be studied separately.

The initial density distribution of our model is

$$\rho(r, t_i) = \rho_B(t_i) + \delta\rho(r_i, t_i) = \rho_B(t_i)[1 + \delta_i(r)] \ , \tag{2.1}$$

where $\delta_i(r) = \delta_i(r, t_i)$ is the initial density contrast which has to be a function of r. In our description it is convenient to use the proper radial coordinate $r = a(t)|\mathbf{x}|$, where $\mathbf{x}$ is the comoving coordinate.

The motion of a thin shell of particles located at a distance r is governed by the equation

$$\frac{\mathrm{d}^2 r}{\mathrm{d}\, t^2} = -\frac{GM}{r^2} = -\frac{4\pi G}{3}\rho_B\big(1 + \Delta\big)\, r \ , \tag{2.2}$$

where

$$M = \frac{4\pi}{3} r_i^3 \rho_B\big(1 + \Delta_i\big) \ , \tag{2.3}$$

$$\Delta_i = \frac{3}{r_i^3}\int_0^{r_i} \delta_i(r) r^2 \,\mathrm{d}r \ , \tag{2.4}$$

with r_i the initial radius of the shell with enclosed mass M and Δ_i is the average value of δ_i within r_i at time t_i.

Eq.(2.2) is analogous to the First Friedmann Equation (1.17) for a monocomponent Universe,

$$\frac{\mathrm{d}^2 a}{\mathrm{d}\, t^2} = -\frac{4\pi G}{3}\rho_B a \ , \tag{2.5}$$

from the similarity of these two eq.s we can see that the perturbation evolves like a Universe of a different mean density, but with the same initial expansion rate.

From eq. (2.2), the first integral of motion is

$$\frac{1}{2}\left(\frac{\mathrm{d}\, r}{\mathrm{d}\, t}\right)^2 - \frac{GM}{r} = \frac{1}{2}\left(\frac{\mathrm{d}\, r}{\mathrm{d}\, t}\right)^2 - \frac{4\pi G}{3}\rho_B(1 + \Delta) r^2 = E \ , \tag{2.6}$$

where E is a constant of integration, namely, the total energy of the perturbation. The sign of E determines the fate of the given mass shell, whether it will eventually decouple from the Hubble flow and collapse or expande forever. If $E < 0$ it follows from eq.(2.6) that, as r increases $\dot{r}$ will eventually become zero and later negative, implying a contraction and collapse. On the other hand, if $E > 0$, then $\dot{r}^2$ will never become zero and the shell will expand forever.

The evolution of a spherical over- or underdensity is *entirely* and *solely* determined by the effective over- or underdensity within the radius r of the shell and by its peculiar velocity v_p (Sheth and van de Weygaert, 2004).

Unsurprisingly, eq.(2.6) is analogous to the Second Friedmann Equation (1.18)

$$\frac{1}{2}\left(\frac{\mathrm{d}\, a}{\mathrm{d}\, t}\right)^2 - \frac{4\pi G}{3}\rho_B a^2 = -\frac{K}{2} \ , \tag{2.7}$$

which describes the expansion rate of the Universe. The constant K, as we have seen in Section 1.1, is the curvature parameter which is $K = -1$ for an open Universe, $K = 0$ for a flat Universe and $K = +1$ for a closed Universe. Note that eq. (2.7) is obtained from the First Friedmann Equation (1.17) applying the adiabatic condition (1.20).

Given these similarities, we can procede as it is done to derive the parametric solutions of the Friedmann equations for curved universes (see e.g. Peebles, 1980), therefore the parametric solutions for a mass shell are

$$r = A(\cosh\theta - 1) \ , \qquad t = B(\sinh\theta - \theta) \ , \qquad E > 0 \ ; \tag{2.8a}$$

$$r = A(1 - \cos\theta) \ , \qquad t = B(\theta - \sin\theta) \ , \qquad E < 0 \ ; \tag{2.8b}$$

with A and B two constants related by

$$A^3 = GMB^2 \ . \tag{2.9}$$

The parameter θ is the *development angle*, which parametrizes all the physical relevant quantities relating to the mass shell.

The homogeneous background model is described by similar equations. Choosing the expansion parameter $a(t) = r_B(t)$ so that it encloses the same mass M as in eq.s (2.8), we have

$$r_B = A_B(\cosh\eta - 1) \ , \qquad t = B_B(\sinh\eta - \eta) \ , \qquad K < 0 \ ; \tag{2.10a}$$

$$r_B = A_B(1 - \cos\eta) \ , \qquad t = B_B(\eta - \sin\eta) \ , \qquad K > 0 \ ; \tag{2.10b}$$

with A_B and B_B two constants related by

$$A_B^3 = GMB_B^2 \ . \tag{2.11}$$

The solutions for the EdS case (or for the zero energy case) are that listed in Table 1.1.

We can then calculate the mean density inside a shell by applying $\rho = 3M/(4\pi r^3)$ and substituting with the dependencies in system (2.8) and in eq. (2.11), we obtain

$$\rho(r,t) = \begin{cases} \dfrac{3}{4\pi G t^2}\dfrac{(\sinh\zeta - \zeta)^2}{(\cosh\zeta - 1)^3} & \text{open} \\ \dfrac{3}{4\pi G t^2}\dfrac{2}{9} & \text{critical} \\ \dfrac{3}{4\pi G t^2}\dfrac{(\zeta - \sin\zeta)^2}{(1 - \cos\zeta)^3} & \text{closed} \end{cases} \ , \tag{2.12}$$

which are valid for both the shell (in which case "open" means $E > 0$) and the background Universe (where "open" means $K > 0$). In the most general form, the density contrast of a perturbation in a monocomponent Universe can be obtained from

$$1 + \Delta(r,t) = \frac{\rho(r,t)}{\rho_B(r,t)} = \frac{f(\theta)}{f(\eta)} \ , \tag{2.13}$$

where θ and η are the development angles of the perturbation and of the background Universe, respectively. The *cosmic density function* $f(\zeta)$ is defined as

$$f(\zeta) = \begin{cases} \dfrac{(\sinh\zeta - \zeta)^2}{(\cosh\zeta - 1)^3} & \text{open} \\ 2/9 & \text{critical} \\ \dfrac{(\zeta - \sin\zeta)^2}{(1 - \cos\zeta)^3} & \text{closed} \end{cases} . \tag{2.14}$$

Moreover, the velocity of expansion or contraction of a spherical shell can be written in terms of θ and η. Let us consider the peculiar velocity of a shell

$$v_p(r,t) = v(r,t) - H(t)r(t) , \tag{2.15}$$

where $v(r,t)$ is the total velocity of the shell and $H(t)$ is the Hubble parameter of the background Universe. We may define a general Hubble parameter H_s for an unspecified shell,

$$H_s(t) = \frac{\dot{r}}{r} = \frac{1}{r}\frac{\mathrm{d}\,r}{\mathrm{d}\,t} = \frac{1}{t}\,g(\zeta) , \tag{2.16}$$

where we have defined the *cosmic velocity function* as

$$g(\zeta) = \begin{cases} \dfrac{\sinh\zeta(\sinh\zeta - \zeta)}{(\cosh\zeta - 1)^2} & \text{open} \\ 2/3 & \text{critical} \\ \dfrac{\sin\zeta(\zeta - \sin\zeta)}{(1 - \cos\zeta)^2} & \text{closed} \end{cases} . \tag{2.17}$$

The peculiar velocity of eq. (2.15) can then be rewritten as

$$v_p(r,t) = H(t)r(t)\left[\frac{g(\theta)}{g(\eta)} - 1\right] . \tag{2.18}$$

With eq. (2.13) and eq. (2.18) we obtained explicit expressions for the evolution of a spherical perturbation in FRW backgrounds with no cosmological constant. In the next sections we will focus to the case of a spherical perturbation evolving in an EdS background Universe.

2.1.1 Overdensities

Consider the evolution of an initially overdense shell in a EdS Universe. Such a shell will initially expand slower than the Hubble flow, until it gradually halts reaching a maximum radius r_m. After this instant, t_{ta}, it turns around and starts collapsing. Since at turn-around the velocity of the shell is $v(r_m, t_{ta}) = 0$, for the closed case eq. (2.17) gives a velocity

$$v(r,t) = H_s(r,t)\,r(t) \propto \frac{\sin\theta}{1 - \cos\theta} . \tag{2.19}$$

From the relation in eq. (2.19) we obtain that at turn-around $\theta_{ta} = \pi$, the resulting density is then

$$1 + \Delta(r_m, t_{ta}) = (3\pi/4)^2 \approx 5.6 \ , \tag{2.20}$$

which is highly non-linear and implies the shell has expanded by a factor of $5.6^{1/3} \approx 1.8$ from what it was initially. The last relation suggests that the collapsing region is nearly 6 times denser than the background Universe, at turn-around. This corresponds to a density contrast of $\Delta \approx 4.6$ which is already in the non-linear regime.

It is interesting to calculate the corresponding value of Δ in linear regime. For a spherical overdensity with initial peculiar velocity $v_p(t_i) = 0$ in an EdS Universe, the evolution of the density contrast is given by both the growing ($D_+ \propto t^{2/3}$) and decaying ($D_- \propto t^{-1}$) mode

$$\delta = \delta_+(t_i)\left(\frac{t}{t_i}\right)^{2/3} + \delta_-(t_i)\left(\frac{t}{t_i}\right)^{-1} \ , \tag{2.21}$$

where $\delta_\pm(t_i)$ are the growing (+) and decaying (−) fraction of the initial perturbation δ_i. The combination in eq. (2.21) is necessary to satisfy the correct boundary condition on the velocity: $V_i = 0$ requires that $\delta_+(t_i) = \frac{3}{5}\delta_i$, where δ_i is the initial density contrast (Coles and Lucchin, 2002).

It is reasonable to assume that the deaying mode will become negligible after a short time. Then the remaining perturbation will then be just $\delta \approx \delta_+(t_i)$. It then follows that, in linear approximation, the *turn-around density contrast* is

$$\delta_{ta} = \frac{3}{5}\left(\frac{3\pi}{4}\right)^{2/3} \approx 1.062 \ . \tag{2.22}$$

Full collapse is instead associated with $\theta = 2\pi$ and the linearly extrapolated initial overdensity reaches the threshold *collapse density contrast*:

$$\delta_c = \frac{3}{5}\left(\frac{3\pi}{2}\right)^{2/3} \approx 1.686 \ . \tag{2.23}$$

We will see in the next Sections the crucial importance of these two threshold values. Nevertheless, even though a strictly gravitational description fixes the comoving radius of the overdensity to zero ($r(2\pi) = 0$), in reality the matter in the collapsing object will eventually virialize. In fact, the more the radius shrinks, the less the approximation of independently evolving shells is correct. Their interaction and energy exchange during the collapse drives the overdensity towards an equilibrium configuration. Therefore, it is usual to assume that the final size of a collapsed spherical object is finite and equal to its virial radius.

This model, except for very special cases, is not to be trusted once the overdensity has stopped expanding and started the collapse. In fact the mass distribution is strongly unstable against the development of non-radial motion, amplifing the initial asymmetries (Lin et al., 1965; Peebles, 1980).

2.1.2 Underdensities

The evolution of underdense spherical regions is different from that of their overdense counterparts. The net outward directed peculiar acceleration is directly proportional to the meam density contrast $\Delta(r, t)$ of the void. Since they are more

underdense, the inner shells "feel" a stronger outward acceleration than the outer shells.

In the case of an EdS model, $\Omega = 1$, from eq. (2.13) the density deficit evolves as

$$1 + \Delta(r,t) = \frac{f_{\rm open}(\theta)}{f_{\rm EdS}(\eta)} = \frac{9}{2}\frac{(\sinh\theta - \theta)^2}{(\cosh\theta - 1)^3} \ . \tag{2.24}$$

It is interesting to compare this result to the corresponding linear initial density deficit $\Delta^L(r,t)$

$$\Delta^L(r,t) = \Delta^L(z) = \frac{\Delta_i}{1+z} = \frac{1}{1+z}\left(\frac{3}{5}\Delta_{+,i} + \frac{2}{5}\Delta_{-,i}\right) \approx \frac{3}{5}\frac{\Delta_{+,i}}{1+z} \ , \tag{2.25}$$

where we keep only the growing solution $\Delta_{+,i}$, since the decaying solution $\Delta_{-,i}$ gradually fades.

The constants in eq. (2.8) are

$$A = \frac{r_i}{2\Delta_i} \ , \tag{2.26}$$

$$B = \frac{3}{4}\frac{t_i}{\Delta_i^{3/2}} \ , \tag{2.27}$$

where r_i and t_i are the initial coordinates of the perturbation (Blumenthal et al., 1992). Eq. (2.25), with eq.s (2.26) and (2.27) and considering the solutions from Table 1.1, gives

$$\Delta_i^L(\theta) = -\left(\frac{3}{4}\right)^{2/3}\frac{3}{5}\left(\sinh\theta - \theta\right)^{2/3} \ . \tag{2.28}$$

A direct consequence of the differential outward expansion within and around the void and of the radial decrease of the expansion rate, is the accumulation of shells near the boundary of the void. The density deficit $\Delta(r)$ of the void decreases as a function of radius r, down to a minimum at the centre. Shells that were initially close to the centre will ultimately catch up with the shells further outside, until they eventually pass them. This marks the event of *shell crossing*.

For an ideal top-hat underdensity, this process also results in the formation of an infinitely dense ridge in the outermost region of the void. From this moment on the evolution of the void may be described by a self-similar outward moving shell (Suto et al., 1984). Even though this extreme condition is verified only by ideally steep underdensities, it is a reasonable approximation.

In spherically symmetric negative perturbations, shell crossing occurs as the fluctuation becomes non-linear. The density grows on the edges where $\delta\rho/\rho = 1$ is reached before shell crossing. Nevertheless shell crossing defines a characteristic time that defines the formation of the void at a non-linear level. The solutions in eq. (2.8) represent a family of trajectories labeled by r_i and parametrized by θ. We can find out when and where shell crossing first occurs by differentiating the parametrized solutions w.r.t. r and θ, and requiring that $\mathrm{d}r$ and $\mathrm{d}t$ vanish:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \mathrm{d}r/r \\ \mathrm{d}\theta \end{bmatrix} = 0 \ , \tag{2.29}$$

where A_{ij} are functions of Δ_i and θ (Jennings et al., 2013). The homogeneous system of linear equations (2.29) has non-zero solutions when $\det A = 0$, thus the shell crossing condition is

$$\frac{\sinh\theta_{sc}(\sinh\theta_{sc} - \theta_{sc})}{(\theta_{sc} - 1)^2} = \frac{8}{9} \ . \tag{2.30}$$

Therefore shell crossing first happens at a development angle $\theta_{sc} \approx 3.49$. Moreover note that, from eq. (2.16) and from the relation (2.30), at shell crossing the void has a precisely determined excess Hubble expansion rate: $H_{sc} = (4/3)H(t_{sc})$.

Substituting θ_{sc}, we can now obtain all the relevant values to deal with voids. First of all, the non-linear (NL) density contrast is given by eq. (2.13)

$$1 + \delta_v^{NL}(r, t) \approx 0.21 \ . \tag{2.31}$$

It follows that $|\delta_{sc}| \approx 0.8 \lesssim 1$, which means that a void is only nearly non-linear. This should not surprise, since there is a physical limit, $(\delta\rho/\rho)_{lim} = -1$, to the underdensity degree that a region can reach. An underdensity which evolved towards shell crossing, has expanded by a factor $(1 + \delta_{sc}^{NL})^{-1/3} \approx 1.72$ w.r.t. its initial size.

Last but not least, from eq. (2.28) we obtain the linearly extrapolated underdensity of a void

$$\delta_v \approx -2.71 \ , \tag{2.32}$$

which is the underdense peer of the threshold found in eq. (2.23).

It has been demonstrated (Icke, 1984) that voids in the matter distribution are likely to be roughly spherical, differently from collapsing objects, which are more likely to evolve into filamentary or sheet-like structures. We have already said that in the collapse the asphericities of a perturbation are magnified. Since the expansion of a void can be adressed as the time reversal of the collapse of an overdensity, it follows that in the underdense case initial asphericities are cancelled.

2.2 Excursion set formalism

With the hypothesis that, at initial time t_i, the density fluctuations are small, $\delta \ll 0$ and are distributed as a Gaussian field, their value at some later time $t > t_i$ is completely determined by the power spectrum of the fluctuations $P(k, t)$.

The linear growth $\delta(\mathbf{x}, t) = \delta_i D(t)/D(t_0)$, with $D(t)$ the growth factor found in Section 1.3.2, is valid until the collapse, for overdensities, or until the shell crossing, for underdensities. The linear thresholds found for the spherical model in the previous sections are weakly dependent on the cosmological model, via the density parameters Ω_m and Ω_Λ.

The statistical description of the distribution of structures in the Universe has been developed, in the first place, to predict the distribution of collapsed objects (Press and Schechter, 1974; Sheth and Tormen, 2002; Sheth et al., 2001; Lacey and Cole, 1993), but was also extended to underdense regions (i.e. voids, Sheth and van de Weygaert, 2004; Jennings et al., 2013). These approaches are based on the *excursion set formalism* (Bond et al., 1991; Zentner, 2007).

Instead of considering a density fluctuation field δ evolving in time towards some critical static threshold, δ_c or δ_v, we can move the time dependency to the

threshold value, multipling all the expressions by $D(t_0)/D(t)$. We then have a linear fluctuation field $\delta(\mathbf{x}) \equiv \delta(\mathbf{x}, t_0)$ rescaled to time t_0 and a critical threshold $\delta_{c|v}(t) = \delta_{c|v} D(t_0)/D(t)$ that changes with time t.

The infinitesimal mass element in $\mathbf{x}$ belongs to a halo of mass M at time t if the linear fluctuation $\delta_f(\mathbf{x}, R)$, centred in $\mathbf{x}$ and filtered over a sphere of radius $R \propto M^{1/3}$, is higher than the required threshold:

$$\mathbf{x} \in M \iff \delta_f(\mathbf{x}, R) \geq \delta_{c|v}(t) \ . \tag{2.33}$$

Note that, since mass and radius depend on each other, the same arguments are valid both for haloes and voids.

This dissertation is possible because the filtered density field $\delta_f(\mathbf{x}, R)$ is the convolution of the punctual density field $\delta(\mathbf{x})$ with a window function $W(\mathbf{x}, R)$. A Fourier decomposition of $\delta_f(\mathbf{x}, R)$ gives

$$\delta_f(\mathbf{x}, R) = \frac{1}{2\pi^2} \int_0^\infty \mathrm{d}k k^2 \hat{\delta}(k) \hat{W}(kR) \approx \frac{1}{2\pi^2} \int_0^{k_f} \mathrm{d}k k^2 \hat{\delta}(k) \equiv \delta_f(\mathbf{x}, k_f) \ , \tag{2.34}$$

where $k_f \propto 1/R$ is the wave number associated to the filter scale R, an the hat over a variable indicates its Fourier-trasformed value.

The value of $\delta_f(\mathbf{x}, R)$ is the sum of all the fluctuations with $k \lesssim k_f$. All the waves with $k \gtrsim k_f$ are removed by the filter. We also introduce the filtered mass variance,

$$\sigma_R^2 = \frac{1}{2\pi^2} \int_0^\infty \mathrm{d}k k^2 P(k) \hat{W}^2(kR) \approx \frac{1}{2\pi^2} \int_0^{k_f} \mathrm{d}k k^2 P(k) \equiv S(k_f) \ , \tag{2.35}$$

which is a monotonic growing function of k_f in all the relevant cases. Each fluctuation δ_f centred at some point $\mathbf{x}$ in real space and filtered on a scale corresponding to a wave number k_f, traces a path in the two-dimensional space $(S(k_f), \delta_f(\mathbf{x}, k_f))$.

Since $S(k_f \to 0) \to 0$ and $S(k_f \to \infty) \to \infty$, all the trajectories start from $(S, \delta_f) = (0, 0)$, corresponding to a null fluctuation filtered on an infinitely large scale. They will then move away from the origin. The movement of a fluctuation in the space (S, δ_f) depends on the matter distribution around $\mathbf{x}$. Since the distribution of over- or underdensities is Gaussian, each infinitesimal step is stocastic.

With a three-dimensional top-hat filter on k, the contributions of all the scales k to δ_f are uncorrelated. The trajectories are then described by a *Brownian motion* with diffusion equation

$$\frac{\partial Q}{\partial S} = \frac{1}{2} \frac{\partial^2 Q}{\partial {\delta_f}^2} \ , \tag{2.36}$$

where $Q(k_f, S)$ is the probability distribution function of the stocastic variable δ_f for all the trajectories with a given $\sigma^2(k_f) = S$.

The general solution of eq. (2.36) for a free Brownian motion is a Gaussian distribution:

$$Q(\delta_f, S) = \frac{1}{\sqrt{2\pi} S^{1/2}} \exp\left(-\frac{\delta_f^2}{2S}\right) \ , \tag{2.37}$$

with an expected value δ_f and variance S. For a given value of S, eq. (2.37) counts the number density of trajectories within the interval $[\delta_f, \delta_f + \mathrm{d}\delta]$.

The fundamental idea of the *excursion set model* is that, at any given time t, all the trajectories starting from the origin, which pierce the barrier $\delta_{c|v}(t)$ for the first time at some S coordinate, belong to haloes of mass $M(S)$ (or, for underdensities, to voids of size $R(S)$). This conversion between mass, size and variance is possible thanks to eq. (2.35) and thanks to the dependencies $k_f \propto 1/R \propto M^{-1/3}$.

We want to find the smallest value of S where a trajectory crosses the barrier $\delta_{c|v}$. This method allows us to count the number of structures which are defined by a crossing of some barrier $\delta_{c|v}(t)$.

Let us consider some critical scale k_0, corresponding to a variance S, at a given time t_0. There are three possible outcomes for the trajectories:

1. trajectories which have pierced the barrier at $k_f < k_0$ and have not pierced it back:
$$|\delta_f(k_0)| \geq |\delta_{c|v}(t_0)| \ ;$$
2. trajectories which have pierced the barrier at $k_f < k_0$ but have also pierced it back at k'_f, with $k_f < k'_f < k_0$:
$$|\delta_f(k_0)| < |\delta_{c|v}(t_0)| \ \wedge \ \exists\, k_f < k_0 \text{ such that } |\delta_f(k_f)| > |\delta_{c|v}(t_0)| \ ;$$
3. trajectories which never pierced the barrier:
$$|\delta_f(k_f)| < |\delta_{c|v}(t)| \ \forall\, k_f \leq k_0 \ .$$

We want to count type 3 trajectories. In this case eq. (2.37) gives all the trajectories with $|\delta_f(k_0)| < |\delta_{c|v}|(t_0)$. We have then to subtract from $Q(\delta_f, S)$ all the trajectories which, at some smaller k_f, have already pierced the barrier, i.e. type 2 trajectories.

Since we have chosen a top-hat filter in the Fourier-space, and given that all the stocastic Fourier-modes are independent of each other, a trajectory has the same probability to move up or down in the δ_f coordinate, at each infinitesimal step. As a consequence, for each type 2 trajectory a virtual equally probable trajectory exists. It moves from a shifted origin $(S, \delta_f) = (0, 2\delta_{c|v})$ and pierces the barrier in the same point $(S, \delta_{c|v(t)})$ of the real one. The virtual trajectory satisfies the same diffusion eq. (2.36) and its solving probability distribution is equivalent to the probability distribution of the real type 2 trajectory:

$$Q_1(\delta_f, S|\delta_{c|v}(t)) = \frac{1}{\sqrt{2\pi} S^{1/2}} \exp\left[-\frac{(\delta_f - 2\delta_{c|v}(t))^2}{2S}\right] , \tag{2.38}$$

where $\delta_f - 2\delta_{c|v}(t)$ is the Gaussian distribution expected value. The probability of type 3 trajectories can then be expressed as

$$\begin{aligned} Q_2(\delta_f, S|\delta_{c|v}(t))\, \mathrm{d}\delta_f &= [Q(\delta_f, S) - Q_1(\delta_f, S|\delta_{c|v}(t))]\, \mathrm{d}\delta_f \\ &= \frac{1}{\sqrt{2\pi} S^{1/2}} \left\{ \exp\left(-\frac{\delta_f^2}{2S}\right) - \exp\left[-\frac{(\delta_f - 2\delta_{c|v}(t))^2}{2S}\right] \right\} . \end{aligned} \tag{2.39}$$

Eq. (2.39) is the exact solution of the diffusion eq. for stocastic trajectories with an absorbing barrier (Chandrasekhar, 1943).

Of all the possible trajectories, the type 3 fraction is given by the cumulative probability:

$$P_2(S, \delta_{c|v}(t)) = \int_{-\infty}^{\delta_{c|v}(t)} Q_2(\delta_f, S|\delta_{c|v}(t))\, \mathrm{d}\delta_f \ . \tag{2.40}$$

Therefore, the complementary probability $\overline{P_2} = 1 - P_2$ gives all the trajectories that, at a certain time t have already pierced the barrier.

In the overdense case, $\overline{P_2}$ is the number fraction of fluid elements that at a certain time t belong to haloes of mass $M(< S)$. This is the definition of cumulative mass function:

$$P(< S, t) = \overline{P_2}(S, \delta_c(t)) \ . \tag{2.41}$$

2.2.1 Press-Schechter halo mass function and exstension

The Press-Schechter (PS) halo mass function was firstly proposed by Press and Schechter (1974) but its reformulation using the excursion set formalism was made later by Bond et al. (1991). It measures the probability that a certain fluid element in $\mathbf{x}$ at time t belongs to a halo of mass M. It can be obtained differentiating, w.r.t. S, the probability distribution of trajectories which pierced the barrier for the first time at t in some S coordinate:

$$p(S, \delta_c(t)) = -\frac{\partial}{\partial S} \int_{-\infty}^{\delta_c(t)} Q_2(\delta_f, S|\delta_c(t))\, \mathrm{d}\delta_f = -\int_{-\infty}^{\delta_c(t)} \frac{\partial\, Q_2}{\partial\, S}\, \mathrm{d}\delta_f \ , \tag{2.42}$$

where we have moved the derivative inside the integral. We can then substitute the derivation w.r.t. S with a derivation w.r.t. δ_f by applying eq. (2.36)

$$p(S, \delta_c(t)) = -\frac{1}{2} \int_{-\infty}^{\delta_c(t)} \frac{\partial^2 Q_2}{\partial\, {\delta_f}^2}\, \mathrm{d}\delta_f = -\frac{1}{2} \frac{\partial\, Q_2}{\partial\, \delta_f}\bigg|_{-\infty}^{\delta_c(t)} \ . \tag{2.43}$$

Differentiating eq. (2.39) we obtain

$$\frac{\partial\, Q_2}{\partial\, \delta_f} = \frac{1}{\sqrt{2\pi S}} \left\{ -\frac{\delta_f}{S} \exp\left(-\frac{\delta_f^2}{2S}\right) + \frac{\delta_f - 2\delta_c(t)}{S} \exp\left[-\frac{(\delta_f - 2\delta_{c|v}(t))^2}{2S}\right] \right\} \ . \tag{2.44}$$

Since this function is zero in the limit to $-\infty$, in eq. (2.43) only the term in $\delta_c(t)$ survives:

$$\begin{aligned} p(S, \delta_c(t)) = \frac{\mathrm{d}\, f(S)}{\mathrm{d}\, S} &= -\frac{1}{2} \frac{\partial\, Q_2}{\partial\, \delta_f}\bigg|^{\delta_c(t)} = \frac{1}{\sqrt{2\pi S}} \frac{2\delta_c(t)}{S} \exp\left[-\frac{\delta_c^2(t)}{2S}\right] \\ &= \sqrt{\frac{2}{\pi}} \frac{\delta_c(t)}{S^{3/2}} \exp\left[-\frac{\delta_c^2(t)}{2S}\right] \ . \end{aligned} \tag{2.45}$$

This last equation gives the mass fraction in haloes with variance S. To obtain the mass fraction in haloes of mass M we have to change variable with the probability conservation law:

$$p(x)\, \mathrm{d}x = p(y)\, \mathrm{d}y \ \Rightarrow \ p(x) = p(y) \left| \frac{\mathrm{d}\, y}{\mathrm{d}\, x} \right| \ , \tag{2.46}$$

so that

$$\frac{\mathrm{d}\, f(M)}{\mathrm{d}\, M} = \frac{\mathrm{d}\, f(S)}{\mathrm{d}\, S}\left|\frac{\mathrm{d}\, S}{\mathrm{d}\, M}\right| .$$

In our case we use

$$\left|\frac{\mathrm{d}\, S}{\mathrm{d}\, M}\right| = \frac{S}{M}\left|\frac{\mathrm{d}\, \ln S}{\mathrm{d}\, \ln M}\right| = \frac{2S}{M}\left|\frac{\mathrm{d}\, \ln \sigma}{\mathrm{d}\, \ln M}\right| ,$$

to obtain the mass fraction in haloes of mass M:

$$\frac{\mathrm{d}\, f(M)}{\mathrm{d}\, M} = \left(\frac{2}{\pi}\right)^{1/2} \frac{\delta_c(t)}{M\sigma(k_f)}\left|\frac{\mathrm{d}\, \ln \sigma}{\mathrm{d}\, \ln M}\right| \exp\left[-\frac{\delta_c^2(t)}{2\sigma^2(k_f)}\right] . \tag{2.47}$$

Multiplying the last expression by a factor ρ/M, where ρ is the matter density of the Universe, we obtain the PS mass function:

$$\frac{\mathrm{d}\, n(M)}{\mathrm{d}\, M} = \frac{\mathrm{d}\, f(M)}{\mathrm{d}\, M}\frac{\rho}{M} = \left(\frac{2}{\pi}\right)^{1/2} \frac{\rho_0}{M^2}\frac{\delta_c(t)}{\sigma(k_f)}\left|\frac{\mathrm{d}\, \ln \sigma}{\mathrm{d}\, \ln M}\right| \exp\left[-\frac{\delta_c^2(t)}{2\sigma^2(k_f)}\right] , \tag{2.48}$$

where we implicitly assume that the number of haloes inside a given volume is conserved when we move from the linear description of the excursion set model to the abundance of the real collapsed non-linear objects,

$$\frac{\mathrm{d}\, n(M)}{\mathrm{d}\, M} = \frac{\mathrm{d}\, n_L(M_L)}{\mathrm{d}\, M_L} .$$

Since we made the hypothesis that, while it collapses, the total mass in each halo is conserved, $M = M_L$, this assumption should be correct.

The excursion set model not only gives the number density of haloes at some epoch but can also be used to study the merger history of haloes and to derive other characteristics (Lacey and Cole, 1993, 1994): these results are often described as *Press Schechter extended model*.

Each trajectory in the (S, δ_c) space describes the merger history of a particle at some coordinate $\mathbf{x}$. The generic δ_f is replaced with $\delta_c(t) \propto \delta_c(t_0)D(t)/D(0)$, so that each different value of the δ_c coordinate defines a different collapse time. Hierarchical clustering describes the growth of haloes as the result of a series of mergers. In an excursion set description, this process corresponds to follow a trajectory from large values of S and δ_c back towards the origin.

While we go back in time, the threshold decreases. Type 2 trajectories imply a horizontal jump in the (S, δ_c) space, towards the first crossing on the left. In a hierarchical clustering description, these jumps are interpreted as the result of merger events.

To describe a halo with mass M_1 at time t_1, and mass M_2 at time t_2, with $t_2 > t_1$, we select all the trajectories which pass through both points $(S_1, \delta_c(t_1))$ and $(S_2, \delta_c(t_2))$. This request is obtained by requiring that the trajectories start in $(S_2, \delta_c(t_2))$ instead of the origin $(0, 0)$. The probability distribution is then obtained substituting $S \to (S_1 - S_2)$ and $\delta_c(t) \to (\delta_c(t_1) - \delta_c(t_2))$ in eq. (2.45):

$$p(S_1, \delta_c(t_1)|S_2, \delta_c(t_2)) = \frac{(\delta_c(t_1) - \delta_c(t_2))}{\sqrt{2\pi}(S_1 - S_2)^{3/2}} \exp\left[-\frac{(\delta_c(t_1) - \delta_c(t_2))^2}{2(S_1 - S_2)}\right] . \tag{2.49}$$

As we have done to obtain the PS mass function, we multiply by a factor $\mathrm{d}S_1/\,\mathrm{d}M_1$ and require number conservation. This leads to the *Conditional Mass Function*

$$n(M_1,t_1|M_2,t_2)=\left(\frac{2}{\pi}\right)^{1/2}\frac{\rho}{M_1^2}\left|\frac{\mathrm{d}\,\ln\sigma_1}{\mathrm{d}\,\ln M_1}\right|\frac{\sigma_1^2(\delta_c(t_1)-\delta_c(t_2))}{(\sigma_1^2-\sigma_2^2)^{3/2}}\exp\left[-\frac{(\delta_c(t_1)-\delta_c(t_2))^2}{2(\sigma_1^2-\sigma_2^2)}\right], \tag{2.50}$$

which gives the number of haloes of mass M_1 embedded in bigger haloes of mass M_2.

2.3 Void size function

The spherical top-hat model provides a complete description of the non-linear evolution of a spherically symmetric density perturbation. One of the main features of this model is that the evolution does not depends on the initial size or enclosed mass of the region, but only on the amplitude of the initial top-hat density contrast.

We have described how the statistical distribution of collapsed DM haloes is analytically predicted, i.e. the PS mass function and its extension to the study of hierarchical clustering.

The counterpart of this description for underdensities has been firstly proposed by SvdW. Even though their model has proven to be inadequate to describe the real distribution of voids, its basic assumptions are reasonable.

The first difference between a halo based description of the statistics of the LSS and the voidness oriented one is in the variable used to characterize structures. While in the first case we used mass M, in the latter we prefer the size R. Nevertheless, since $M\propto R^3$, the two approaches are equivalent.

SvdW argue that, if a description of the cosmic spatial structure has to be based on the idealization of spherical symmetry, low-density regions are the objects of choice. A succesful representation of the void size statistics has to take into account two main processes. First, a void emerges as a mature well-defined entity through a bottom-up assembly of smaller underdensities. Secondly, the void may interact with its surroundings, participating to the continuing process of hierarchical structure formation.

To take into account these requests in the description, let us first consider the evolution of an isolated void. The most basic and universal properties of evolving spherical voids are their *expansion* (van de Weygaert and van Kampen, 1993), in contrast with their overdense peers which tend to collapse, and the *evacuation* of mass. Moreover they have a *super-Hubble velocity field* which causes a *suppression in structure growth*. Other properties which have already been presented in Section 2.1.1 are the formation of a *boundary ridge* and, last but most importantly, *shell crossing*.

In all our description we will define voids as underdense, spherical, non-overlapped regions, which have gone through shell crossing. We will then search for all the trajectories in the usual (S,δ_f) space which have pierced the negative barrier $\delta_v=-2.71$ (eq. 2.32). To obtain the PS mass function we counted all the trajectories that pierced the barrier δ_c and considered them as belonging to haloes of mass

$M(S)$, where S is the minimum variance at which the barrier is pierced. Ignoring all the other crossing of the barrier allows us to avoid double counting haloes (*cloud-in-cloud* process).

For their extension and natural tendency to embed other structures, other processes are to be taken into account. To address this nature, the void size function has to depends on two barriers instead of only one, the shell crossing level δ_v and the collapse density δ_c. Here follows the main processes that have to be considered:

- *Cloud-in-cloud*, it is the same as for haloes: When a trajectory pierces the δ_c barrier more than once, only the crossing with the smallest value of δ_c has to be considered.

- *Cloud-in-void*: when a trajectory first crosses the δ_v barrier and then crosses the δ_c one, it means that a clump may form in a region that is destined to become a void.

- *Void-in-void*: the formation of a large void by the merger of smaller voids. These is the underdense counterpart of the PS extended formalism presented at the end of Section 2.2.1. These trajectories have to be counted only once, and represent voids of size $R(S)$, where S is the smallest variance where the trajectory crosses δ_v.

- *Void-in-cloud*: this is the main difference between haloes and voids. If a random walk first crosses δ_c and then crosses δ_v on a smaller scale, then the smaller void is contained within a larger collapsed region. Since the larger region has collapsed, the smaller void within it no longer exists, so it should not be counted.

In summary, what distinguishes voids from collapsed objects is that, whereas it may be possible to have a cluster within a void, it does not make physical sense to have a void within a cluster. The excursion set approach allows one to account for this.

From eq. (2.45) the differential fraction of random walks which have firstly crossed a barrier of constant height δ_v in $S \equiv \sigma^2$ is:

$$f_{\ln\sigma} = \frac{\mathrm{d}\, f(S)}{\mathrm{d}\, \ln S} = S\frac{\mathrm{d}\, f(S)}{\mathrm{d}\, S} = \sqrt{\frac{2}{\pi}}\frac{\delta_v}{\sigma}\exp\left[-\frac{\delta_v^2}{2\sigma^2}\right] , \tag{2.51}$$

which is related to the corresponding differential probability distribution by $p(S, \delta_f) = \sigma^2 f_{\ln\sigma}$. Instead of focusing on the enclosed mass, we define the variance σ^2 in terms of the size of the considered region:

$$\sigma^2(R) \equiv S(R) = \frac{1}{2\pi}\int \mathrm{d}k k^2 P(k)|W(k,R)|^2 . \tag{2.52}$$

The distribution $f_{\ln\sigma}$ then gives the excursion set approach's approximation for the fraction of volume that is occupied by voids.

In order to find the probability that a certain region of scale $\sigma^2(R) = S$ will eventually become a void, we have to subtract to the first crossing distribution of

the barrier δ_v ($p(S, \delta_v)$), the subset of trajectories that had crossed δ_c before ever reaching δ_v,

$$p(S, \delta_v, \delta_c) = p(S, \delta_v) - \int_0^S p(S, \delta_v | s, \delta_c)\, \mathrm{d}s \ , \tag{2.53}$$

where $p(S, \delta_v | s, \delta_c) = p(S - s, \delta_c - \delta_v)$ is obtained from eq. (2.49) with $S_1 = S$, $S_2 = s$, $\delta_c(t_1) = \delta_v$ and $\delta_c(t_2) = \delta_c$. While searching for the first crossing of the barrier prevents double countings caused by the void-in-void process, subtracting the second term on the r.h.s. of eq. (2.53) avoids counting errors due to the void-in-cloud process. It has been demonstrated by SvdW that eq. (2.53) leads to

$$f_{\ln\sigma}(\sigma) = 2\sum_{j=1}^{\infty} j\pi x^2 \sin(j\pi\mathcal{D}) \exp\left[-\frac{(j\pi x)^2}{2}\right] \ , \tag{2.54}$$

where

$$\mathcal{D} = \frac{|\delta_v|}{\delta_c + |\delta_v|} \quad \text{and} \quad x = \frac{\mathcal{D}}{|\delta_v|}\sigma \ .$$

In their derivation of this law, SvdW also demostrate that all the random walks cross δ_c at some point, indicating that all mass is associated with gravitationally bound haloes. In contrast, only a fraction $f_{void} = \delta_c/(\delta_c - \delta_v)$ of all random walks cross δ_v without first having crossed δ_c, i.e. only a fraction f_{void} of the mass is associated with voids.

Since the infinite series in eq. (2.54) is cumbersome to work with, it is useful to have a sufficiently accurate closed form expression. A reasonable approximation is given by

$$f_{\ln\sigma}(\sigma) \approx \begin{cases} \sqrt{\dfrac{2}{\pi}}\dfrac{|\delta_v|}{\sigma}\exp\left(-\dfrac{\delta_v^2}{2\sigma^2}\right) & x \leq 0.276 \\ 2\displaystyle\sum_{j=1}^{4} j\pi x^2 \sin(j\pi\mathcal{D}) \exp\left[-\dfrac{(j\pi x)^2}{2}\right] & x > 0.276 \end{cases} \ , \tag{2.55}$$

which was proposed by Jennings et al. (2013) and is accurate at the 0.2% level or better across the domain of validity.

2.3.1 Sheth and van de Weygaert model

By the same reasoning as applied to haloes, the SvdW formula for the abundance of voids in linear theory is given by

$$\frac{\mathrm{d}\, n_L}{\mathrm{d}\, \ln r_L} = \frac{\rho}{M} f_{\ln\sigma}(\sigma) \frac{\mathrm{d}\, \ln\sigma^{-1}}{\mathrm{d}\, \ln r_L} = \frac{f_{\ln\sigma}(\sigma)}{V(r_L)} \frac{\mathrm{d}\, \ln\sigma^{-1}}{\mathrm{d}\, \ln r_L} \ , \tag{2.56}$$

with the subscript L indicating values derived in linear theory and where we express the number density in terms of the linear theory radius of the void r_L.

In the spherical evolution model, the actual void expands from its linear radius. At some epoch of shell crossing $\rho_v \approx 0.21\rho_B$ (eq. 2.31). We have already seen that this means that the radius expands by a factor $(r/r_L) \approx 1.72$. In the SvdW

model the idea relies on equating a number density derived from linear theory to the number density of non-linear voids:

$$\frac{\mathrm{d}\, n}{\mathrm{d}\, \ln r} = \frac{\mathrm{d}\, n_L}{\mathrm{d}\, \ln r_L}\bigg|_{r_L = r/1.72} . \tag{2.57}$$

This results in a shift of the linear model towards larger scales, without any change in amplitude (Figure 2.1). The simple shift on the scale axis means that the total

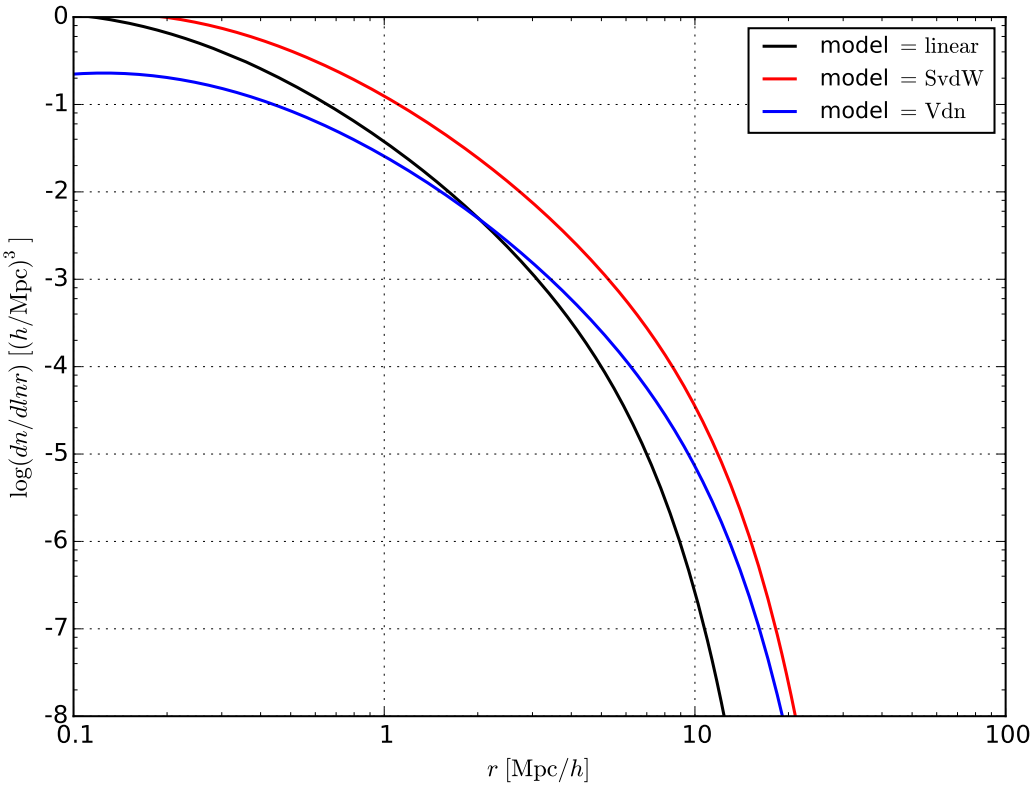

Figure 2.1: Void size function for three different models. The black line represents the linear theory result (eq. (2.56)). The red line is the SvdW model (eq. (2.57)). The blue line is the Vdn model (eq. (2.60)).

comoving number of voids is conserved while they grow ($n = n_L$). This assumption may not be valid, expecially for large voids, given their natural tendency to expand and overlap.

One last remark about the dependency on two thresholds. While the lower threshold is set by the event of shell crossing, there is more freedom in the choice of the overdensity threshold, since we are not sure if a void embedded in an overdensity (void-in-cloud process) should be considered definitively crushed out of existence when the overdensity turns around or after it virializes. To rule out this uncertainty we let the upper threshold vary between the turnaround threshold and the collapse one: $1.06 \leq \delta_c \leq 1.69$. Looking at Figure 2.1, it is possible to appreciate that this uncertainty only affects the smaller voids. We will see in the next Chapters that these scales are only marginally sampled by simulations.

2.3.2 Volume conserving model (Vdn)

A more physical approach to address the evolution of voids through an excursion set approach is to give up on the assumption of isolated spherical expansion, which leads to the total number conservation hypothesis. This request was acceptable for haloes, since they tend to evolve towards a localized geometry. Unfortunately for voids this is not the case. As a result of their natural tendency to expand, overlap and merge, the total number may not be conserved, with the larger voids more affected than the smaller ones.

In particular, the cumulative volume fraction in voids larger than some value R is

$$\mathcal{F}(R) = \int_R^\infty \frac{\mathrm{d}r}{r} V(r) \frac{\mathrm{d}\, n}{\mathrm{d}\, \ln r} \ . \tag{2.58}$$

In the SvdW this integral has values greater than unity for radii of interest (Figure 2.2), which is unphysical.

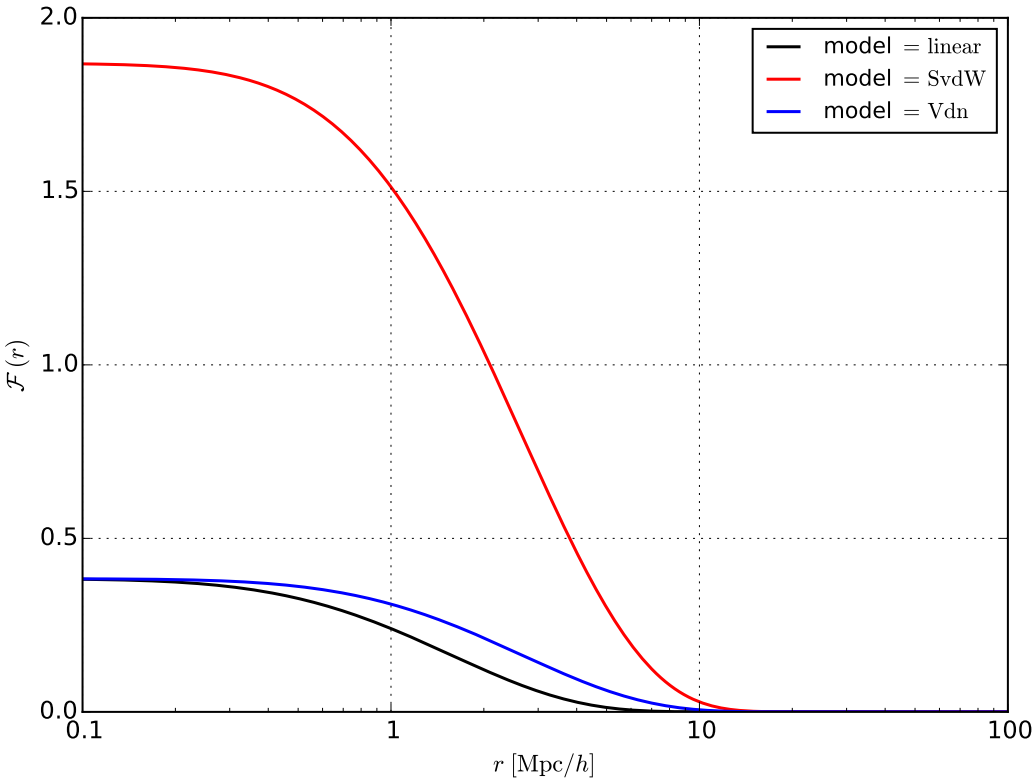

Figure 2.2: Cumulative void size function. The different curves give the fraction of volume occupied by voids, as it is predicted by different models: the black line illustrates the linear theory prediction, the red line the SvdW model and the blue line the Vdn model.

Jennings et al. (2013) suggest a simple fix to this unphysicality. They argue that the volume fraction and shape of the abundance function, rather then the total number of voids, should be fixed during the expansion. If we use eq. (2.58) to define the volume fraction in linear theory, it is conserved if the non-linear abundance satisfies

$$V(r)\, \mathrm{d}n \equiv V(r_L)\, \mathrm{d}n_L\big|_{r_L(r)} \ . \tag{2.59}$$

The abundance of voids varies while they expand from r_L to r as a result of the fusion between neighbouring smaller voids into larger ones. The non-linear theory abundance then becomes

$$\frac{\mathrm{d}\, n}{\mathrm{d}\, \ln r} = \frac{V(r_L)}{V(r)} \frac{\mathrm{d}\, n_L}{\mathrm{d}\, \ln r_L} \frac{\mathrm{d}\, \ln r_L}{\mathrm{d}\, \ln r} \,, \tag{2.60}$$

which gives the so-called *Vdn Model*. Its abundance prediction is shown in Figure 2.1, while its the cumulative volume fraction is shown in Figure 2.2. Note that the volume fraction does not reach unity, as requested.

If the mapping $r(r_L)$ is that resulting from linear theory (i.e. $r = 1.72 r_L$) then the $\mathrm{d}\ln r_L / \mathrm{d}\ln r$ term in eq. (2.60) is unity. As a consequence, the impact of moving from linear to non-linear theory is a shift in both the scale r and the amplitude of $\mathrm{d}n / \mathrm{d}\ln r$. The shape of the distribution does not vary. The ratio $V(r_L)/V(r)$ in eq. (2.60) shows that the Vdn model predicts a factor of $1.72^3 \approx 5$ less voids than the SvdW model at all radii.

The first goal of this work is to validate the Vdn model, thus demonstrating the efficiency of the spherical model for underdensities.

2.4 Halo bias

All our previous dissertation was based on the evolution of density fluctuations in the density field. The most relevant part of mass in the Universe is DM, baryons only have a minor role. This is particularly true when we consider the growth of structures. Since baryons decouple from radiation later than DM, it is the development of gravitational potential wells and peaks, resulting from the aggregation or evacuation of DM, that drives the later evolution of baryons.

Nevertheless, DM is (still) not directly observable since it only interacts weakly with ordinary matter. What we do observe with galaxy surveys is the baryon component's radiation. Baryons naturally gather up into the potential wells generated by haloes, which result from the gravitational aggregation of DM. This means that their emission traces the position of DM.

If we want voids to become efficient tools to investigate cosmological models, we have to first develop a theory capable of compare theoretical results with observations.

In the voids statistical distribution context, the first step to reach such an objective is to develop a model of void size function able to describe the size distribution of voids when the density field is traced by DM haloes.

In the theory developed in the previous Sections we considered voids which form in the DM density field, as a result of spherical expansion of primordial underdensities. Both the halo and the galaxy population, though, trace the underlying DM density field in a biased way. In particular, the halo distribution trace a biased number density field n_h which may be used to define the halo voids distribution. Nevertheless, a comparison between voids in the DM and halo distributions should account for the halo biasing relation.

A result of the underdensity of voids and of their super-Hubble expansion rate is that the mass of haloes which form within them is generally smaller. Moreover, even though we will not work with galaxy models, we underline that the galaxy

luminosities are generally smaller inside voids, without significantly changing the ranking of galaxy luminosity (Benson et al., 2003).

Jennings et al. (2013) also compare their model with the halo voids distribution in simulation. They apply the same searching criteria to a halo distribution without finding an agreement with the Vdn model, suggesting there may not always be a one to one correspondence between voids traced by haloes and voids in the DM density field.

Furlanetto and Piran (2006) propose a scale dependent modification of the barrier $\delta_v = -2.71$ fixed by shell crossing. They argue that, since the population of haloes is dependent on the environment, voids in the DM and in the DM-haloes density fields may not share the same underdensity degree.

They demonstrate that the larger the void in the DM distribution, the less biased it is in the DM-halo distribution. To adress this issue, let us now consider the halo population inside voids. It is reasonable to assume that within voids haloes are formed by the exact same processes as outside them. This means that we can make use of the PS formalism to model their distribution.

If we address haloes in voids as sub-structures of larger evolved perturbations, the conditional mass function (eq. 2.50) gives

$$n_h(m_h|\delta_v, R_v) = \sqrt{\frac{2}{\pi}} \frac{\rho_v}{m_h^2} \left| \frac{\mathrm{d}\ \ln\sigma}{\mathrm{d}\ \ln m_h} \right| \frac{\sigma^2(\delta_c - \delta_v)}{(\sigma^2 - \sigma_v^2)^{3/2}} \exp\left[-\frac{(\delta_c - \delta_v)^2}{2(\sigma^2 - \sigma_v^2)} \right] , \tag{2.61}$$

where m_h is the halo mass, δ_v the linearized underdensity threshold fixed by shell crossing, $\sigma(m_h)$ is the mass variance of the halo and $\sigma_v(M_v)$, with $M_v \propto R_v^3$, the mass variance of a spherical void with radius R_v and density $\rho_v = (1 + \delta_v^{NL})\overline{\rho}$ from eq. (2.31), if the mean density of the Universe is $\overline{\rho}$.

Eq. (2.63) then gives the total number of haloes which form within a void of mean density ρ_v. Note that we cannot use the medium density of the Universe here, since it would overpredict haloes by a factor $(1 + \delta_v^{NL})$. This equation is defined only for values of $\sigma_v^2(M_v) < \sigma^2(m_h)$. The physical reason is simple: since M_v fixes the total amount of matter available within the void, no haloes with $m_h > M_c$ can develop.

The total comoving number density of haloes, with mass larger than some m_{min}, within a void is therefore given by the cumulative value of the mass function (2.63)

$$n_h^c(m_{min}|\delta_v, R_v) = \int_{m_{min}}^{\infty} \mathrm{d}m_h n_h(m_h|\delta_v, R_v) \ . \tag{2.62}$$

Thus, the ratio between $n_h^c(m_{min}|\delta_v, R_v)$ and the total comoving number density of haloes in the Universe, should give a measure of how biased tracers sample the density field within the void. We can therefore obtain an expression for the scale dependent underdensity threshold for voids traced by the DM-halo distribution (Furlanetto and Piran, 2006):

$$1 + \delta_h^{NL}(m_{min}, \delta_v, R_v) = (1 + \delta_v^{NL}) \frac{n_h^c(m_{min}|\delta_v, R_v)}{n_h^c(m_{min})} \ , \tag{2.63}$$

which depends on the mass resolution of the sample m_{min} and on the void scale R_v. The value δ_v is, again, fixed by shell crossing.

Chapter 3

Methods

In this chapter we will describe the procedure used to obtain the results presented later in this work. We make use of already existing Open Source softwares, which we expanded to develop a procedure aimed mainly at the search for spherical not-overlapped voids.

This chapter is organized as follows: in Section 3.1 we present the CosmoBolognaLib (Marulli et al., 2016) and the main numerical methods used in this work; in Section 3.2 we discuss the characteristics of the cosmological simulations used in this work; the void finding algorithms used to obtain void centres is described in Section 3.3. In Section 3.4 we describe an algorithm we have developed to obtain catalogues of underdense regions which are coherent with the theoretical definition of voids given in the previous Chapter.

3.1 CosmoBolognaLib: C++ libraries for cosmological calculations

The CosmoBolognaLib (hereafter CBL) is a large set of Open Source C++ numerical libraries for cosmological calculations. CBL is a *living project* aimed at defining a common numerical environment for cosmological investigations of the large scale structure of the Universe. This software is particularly suited to handle with catalogues of astronomical objects, both real and simulated. A large set of processing and analyzing functions is also implemented in the software. Thanks mainly to the adopted object-oriented programming technique, the CBL is flexible enough to be extended to other cases, in particular to handle astronomical objects, such as voids, for wich the software was not specifically designed.

In particular, we have upgraded the CBL to handle with catalogues of voids obtained with the two most popular public void finders, operation made possible by the implementation of a specific class of objects specifically designed to store relevant properties of voids (class `cosmobl::Void`[1]). The upload of all the developed functions is still not finished but will soon be completed.

[1] `cosmobl` is the global namespace of the CBL.

3.2 N-body simulations

A N-body simulation is, in general, a numerical simulation of a dynamical system of particles.In cosmology, the interest in N-body simulations is related to the study of the large scale structure formation and evolution. Thanks to the possibility of setting the cosmology-defining parameters as the user wish, N-body simulations are foundamental benchmarks for theoretical results.

As anticipated in Section 1.3.4, the main advantage in using a N-body simulation is that it allows to study the non-linear regimes of perturbation growth without the necessary simplifications adopted to reach analytical solutions.

In the most simple N-body simulations, only gravitational effects are taken into account. An N-body problem can be summarized by the general system

$$\begin{cases} \mathbf{F}_i = GM_i \sum_{i \neq j} \frac{M_j}{r_{ij}^2} \hat{r}_{ij} \\ \frac{\mathrm{d}\,\mathbf{v}_i}{\mathrm{d}\,t} = \frac{\mathbf{F}_i}{M_i} \\ \frac{\mathrm{d}\,\mathbf{x}_i}{\mathrm{d}\,t} = \mathbf{v}_i \end{cases} , \tag{3.1}$$

where $\mathbf{F}_i$ is the i-th component of the force, the M_i are the particle masses, $\mathbf{x}_i$ are the comoving coordinates of the i-th particle and $\mathbf{v}_i$ its velocity, r_{ij} is the distance between particle i and j, $\hat{r}_{ij}$ the corresponding versor. Basically the system (3.1) is equivalent to (1.38). The Euler equation of motion can be re-written as

$$\frac{\mathrm{d}\,\mathbf{v}_i}{\mathrm{d}\,t} + 2\frac{\dot{a}}{a}\mathbf{v}_i = -\frac{1}{a^2}\nabla\phi = -\frac{G}{a^3}\sum_{i,j\neq i} m_j \frac{\mathbf{x}_i - \mathbf{x}_j}{|\mathbf{x}_i - \mathbf{x}_j|^3} = \frac{\mathbf{F}_i}{a^3} , \tag{3.2}$$

where a is the scale factor, G the gravitational constant and a dot indicates a derivation w.r.t. time. The Poisson equation with the Second Friedmann Equation becomes

$$\nabla^2\phi = 4\pi G\bar{\rho}(t)a^2\delta = \frac{3}{2}H_0^2\Omega_0\frac{\delta}{a} , \tag{3.3}$$

where $\bar{\rho}$ is the medium non-relativistic matter density of the Universe, δ the local density contrast, H_0 is the Hubble parameter and Ω_0 the non-relativistic matter density parameter.

A N-body code consists in the numerical integration through discretized time steps, δt. The algorithm is divided in two main sections that are reiterated for each time step. The first section calculates the total gravitational force, $\mathbf{F}_i$, acting on each particle at some given instant t. Afterwards, the numerical integration of the equation of motion begins and the new positions, $\mathbf{x}_i^{t+\delta t}$, and velocities, $\mathbf{v}_i^{t+\delta t}$, are obtained for each particle. At each cycle the time is updated, $t = t + \delta t$. The choice of δt is not completely arbitrary but has to follow a suitable criterion. From the set of possible time steps that satisfy the requirements the largest one is chosen. There is a wide number of possible criteria to choose which can be suitable for different aproaches, nevertheless they can be divided into three main categories: *(i)* total energy conservation; *(ii)* convergence of final positions and velocities and *(iii)* reproducibility of the initial conditions (Bagla and Padmanabhan, 1997). The outputs of such an algorithm consist of a set of snapshots of the particles positions. Each snapshot contains data about some istant in the time evolution of the system of particles.

In cosmological simulations, the periodic boundary conditions and a very large number of particles are the key considerations in developing the detailed algorithm for a N-body code. Of two main N-body code components, the integration of the equation of motion is a process of order $O(N_p)$, where N_p is the total number of particles, while the calculation of force depends on the used method. There is a large variety of techniques, which differ in CPU efficiency, implementation simplicity, mass and spatial resolution, to compute the gravitational field. These techniques are:

- *Particle-Particle* (PP): the force acting on the i-th particle is given by the direct summation of all the forces generated by all other particles in the simulation. Since the number of operations required is proportional to N_p^2, this class of methods is the most time consuming, but also the most accurate, because the force is computed for each particle.

- *Particle Mesh* (PM): the Poisson equation is solved in the Fourier domain and the potential/force is computed on a fixed grid. All the relevant quantities (e.g. the density field or the force itself) are interpolated in a grid in order to move the particles. Even though this may be the fastest of the possible methods, its resolution is limited by the smothing scale of the grid.

- *Particle-Particle Particle Mesh* ($\mathrm{P^3M}$): this scheme enhances the resolution of a PM method by adding a short-range direct-summation force to the force interpolated on the mesh inside a sphere of fixed radius R around each particle (Efstathiou et al., 1985). The number of operations added by this correction is proportional to $N_p n(R)$, where $n(R)$ is the average number of particles within R.

- *Tree code*: the simulation volume is divided in a hierarchical tree of subregions. Each of them is characterized by the centre of mass and the mass of all the hosted particles. Since the force from particles in distant regions is approximated by the force from the centre of mass for particles in that region, the number of operations required for calculating the force acting on each particle is reduced. Resolution is reached by continuing the subdivision of cells until, on the lowest level of the tree, in each cell is included at most one particle. The total number of required operations is proportional to $N_p \ln(L_{box}/l_{min})$, where L_{box} is the side length of the box and l_{min} the smallest inter-particle separation for a given particle distribution.

All of these methods have advantages and disadvantages.

In the following we present our set of cosmological simulations obtained through N-body algorithms.

3.2.1 The set of cosmological simulations

The simulations used in this work have been obtained with a GADGET-2 based code (see e.g. Springel, 2005; Baldi, 2012a, and references therein). The main structure of GADGET-2 is that of a Tree code where gravitational interactions are computed with a hierarchical tree multipole expansion. In this work we have not considered

the hydrodynamical interactions between particles, but only their gravitational interaction.

We use a set of simulations with different resolutions, cosmological parameters and sizes to work both with dark matter particles and haloes. The general characteristics of the used simulations are reported in Table 3.1, togheter with the box

Table 3.1: Our set of cosmological simulations with relevant physical quantities. The first letter in the name indicates the cosmological model, the second is for the side length of the box ([h]uge = 1 Gpc, [l]arge = 500 Mpc, [m]edium = 256 Mpc, [s]mall = 128 Mpc, [t]iny = 64 Mpc) while the third letter is for the spatial resolution (in terms of m.p.s., which is measured in Mpc: [l]ow = 2, [m]edium = 1, [h]igh = 0.5, [v]ery high = 0.25).

name	$\mathbf{L_{box}}$ [Mpc/h]	$\mathbf{N_p}$	**m.p.s.** [Mpc/h]	**mass** [$10^{10}\ M_\odot/h$]	**Cosmology**
Lll	500	256^3	≈ 2.00	56.06	ΛCDM
Lsm	128	128^3	1.00	7.52	
Lmm	256	256^3	1.00	7.52	
Lsh	128	256^3	0.50	0.94	
Ltv	64	256^3	0.25	0.12	
Lhm	1000	$2 \cdot 1024^3$	1.00	5.84	
				1.17	
Emm	256	256^3	1.00	7.52	EdS
Omm	256	256^3	1.00	7.52	Open

size of each simulation and the number of included particles, we list the spatial resolution in terms of the *medium particle distance* (m.p.s.),

$$\text{m.p.s.} = \left(\frac{V}{N}\right)^{1/3},$$

and the mass resolution. The "Cosmology" column refers to the chosen set of cosmological parameters, which are listed in Table 3.2.

Table 3.2: The different sets of adopted cosmological parameters: $h_0 = H_0/100 km\, s^{-1}\, Mpc^{-3}$ with H_0 the Hubble parameter at redshift zero; Ω_{CDM}, Ω_b and Ω_Λ the present day density parameters of CDM, baryons and cosmological constant, respectively; $\mathcal{A}_s$ and n_s are the scalar amplitude and the spectral index of the initial power spectrum, respectively.

	h_0	Ω_{CDM}	Ω_b	Ω_Λ	$\mathcal{A}_s$	n_s
ΛCDM	0.7	0.226	0.0451	0.7289	$2.194 \cdot 10^{-9}$	0.96
EdS	0.7	0.95	0.05	0.0	$6.408 \cdot 10^{-10}$	0.96
Open	0.6731	0.226	0.0451	0.0	$7.6648 \cdot 10^{-9}$	0.96

By means of a FoF algorithm, we obtained a set of halo catalogues from the `Lhm` simulation box, that will be used and described in Chapter 5.

3.3 Void Finders

There is not general concordance in the definition of voids. Indeed, many different techniques have been proposed and applied over the years (see Colberg et al., 2008, for a cross-comparison of different void finders).

Void finders can be classified in three broad classes, depending on the type of criterion adopted (Lavaux and Wandelt, 2010). The first class is based on a density criterion and defines voids as regions empty of tracers or with local density well below the mean (e.g. El-Ad and Piran, 1997; Hoyle and Vogeley, 2002). Tracers (typically galaxies) are divided in two classes: *wall tracers* and *field tracers*: while the first lie in strongly overdense regions the latter in "mildly underdense" regions. The separation between the two depends on an ad hoc parameter which specifies the local density of tracers. Voids are then regions empty of wall tracers.

The second class uses geometric criteria and identifies voids as geometrical structures like spherical cells or polyedra (e.g. Platen et al., 2007; Neyrinck, 2008; Sutter et al., 2015b). In particular void finders belonging to this class look for particular features in the continuous three-dimensional distribution of matter. It is then necessary to reconstruct the density field from the sampling tracers distribution. Local minima in the density field are then used to reconstruct the void distribution. Further in this Section we will explain in more detail this technique, which is the one adopted in this thesis.

The third class is that of finders based on dynamical criteria in which galaxies are considered as test particles of cosmic velocity field as well as tracers of the underlying mass distribution (e.g. Hahn et al., 2007; Lavaux and Wandelt, 2010; Elyiv et al., 2015).

The first two classes identify voids from tracers' Eulerian positions. This strategy has disadvantages especially when using sparse samples of objects as mass tracers of the density field. On the first place, when galaxies are used as mass tracers, some biasing prescription has to be adopted to specify the relation between the galaxy and the mass density field. This has been demonstrated to play a significant role when using voids to distinguish different cosmological scenarios, in particular different theories of gravity or couplings in the dark sector, as these models might feature a non-standard evolution of the halo bias (Pollina et al., 2016). Moreover the Eulerian nature of these methods is prone to significant shot noise which grows rapidly as the resolution of tracers decreases. This is arguably the most relevant problem as, to make matters worse, voids are naturally underdense and so even more affected. Nevertheless, methods of the second class are largely the most used and at least a couple of open source functioning codes are available.

The methods of the third class are less affected by shot noise, as they can be defined in Lagrangian coordinates, and are found to ease a theoretical interpretation (Lavaux and Wandelt, 2010; Elyiv et al., 2015). However the voids selected cannot be directly compared with those obtained by the other methods.

In our work we need a method which is able to select objects in simulations as similar as possible to the given theoretical definition of a void used in the theoretical formalism presented in Section 2.3 (i.e. an underdensity which is evolved through *shell-crossing*). We mainly work with DM particles as tracers, which are less sparse than haloes or galaxies, making them less affected by shot noise. For the first part

of the work we do not have particular reasons to do things the hard way, so we make use of an already available void finder of the second class[2].

VIDE, the Void IDentification and Examination Toolkit (Sutter et al., 2015b), is an open source Python/C++ code for finding cosmic voids in galaxy redshift surveys and N-body simulations, charachterizing their properties, and providing a platform for more detailed analysis. At its core, VIDE uses a substantially enhanced version of ZOBOV (ZOnes Bordering On Voidness), a popular publicly available code (Neyrinck, 2008) that finds density depressions in a set of points, without any free parameters or assumptions about shape. Conceptually a ZOBOV void is simply a density minimum with a depression around it. The density field is estimated using a *Voronoi Tesselation* technique (hereafter VT, see van de Weygaert and Schaap, 2007, for a review). VT divides space into cells around each particle, with each cell defined as the region which is closer to a certain particle than to any other. The numerical density field is then defined in each i-th particle position as the inverse of the corresponding cell's volume: $n(i) = 1/V(i)$.

The density minima are found as the cells whose density is lower than the density of every other adjacent cell, which are then linked to this minima to construct *zones*. This step is made for both the sake of calculus speed and information compression. Nevertheless these zones can conceivably be called voids.

Finally, zones are joined together in voids with a *watershed method* similar to that applied by Platen et al. (2007). With the watershed concept a void is defined with analogy to a catchment basin in the density field. For each zone the "water level" is set to the minimum density of the field and then gradually raised until a density maximum is encountered, virtually causing water to overflow into another basin. If this maximum is below a certain fraction of the average density of tracers then the two zones are considered substructures of the same void, else they will be considered as two separate voids. The choice of the discriminating density is a free parameter, which can be selected by the user. These procedure leads to a dependencies' tree of zones which is stored in three output files: two files are binaries containing the used tracer coordinates and zone adjacencies, while the third file contains the output void catalogue.

For our work the main advantages in using VIDE, instead of running directly ZOBOV, are in terms of functionality in pre-processing. In fact VIDE offers the possibility of working both with DM particles and haloes, and it can also randomly subsample the input catalogues. The outputs of VIDE include all the original outputs of the ZOBOV algorithm and add a series of post-processed catalogues, with a set of additional information on the level of voids in the dependencies' tree.

3.4 From centres to spherical non-overlapped voids

As we are searching for voids which have gone through shell crossing, we have to proceed with an accurate post-processing of the catalogues provided by the void finder. The aim is to obtain a catalogue of voids which are as close as possible to the spherical underdensities gone through shell crossing, described in Section 2.1.

[2]Nevertheless we will work also with sparse tracers and so further investigations using a more suitable void finder would be a desirable future development of this work.

In fact, most of the studies carried on void sizes statistics lack of coherence between the voids investigated and the theoretical model applied.

In particular, we wish to make a direct comparison to the predictions of the excursion set formalism which assumes that these underdense regions are non-overlapping spheres of a given underdensity corresponding to a region at the moment of shell crossing. The moment of shell crossing, as pointed out in Section 2.1.2, is the key feature which defines a nonlinear void in the matter distribution today.

The catalogues provided by ZOBOV and VIDE retain information about the voids' centres position, their statistical relevance, radius and core density. In these catalogues all the voids and subvoids found by the algorithm are stored, without restrictions on the shape, on possible overlappings or on density. Since these objects are merely underdense regions and then do not fall into our definition of voids, we have implemented an algorithm which overcomes the inconsistency between the voids found in simulated tracers distributions and the theoretical definition of voids given in Section 2.1.2. The algorithm is organized as follows:

- we clean the catalogue provided by the void finder in order to speed up the search and prune not relevant objects;
- we rescale the void size according to the shell crossing threshold and to the sphericity requirement;
- we erase all the overlapping objects on the basis of a "larger priority" criterion.

The outline of this method has been firstly presented in JLH to be applied to ZOBOV catalogues. We have extended to work with all void catalogues which list the void centre position, its central density (i.e. the value of the density within the inner region of the void) and the embedded density contrast. We extend their approach to VIDE catalogues for reasons that will be clearified in this Section.

In the rest of this Section we first present a quick discussion on how to define the centre of a void, we then proceed in the detailed description of the algorithm.

3.4.1 Void centres

While there is a general agreement on the theoretical definition of a void, this is not the case when it comes to simulations or surveys. The existence of several alternative definitions of void centre reflects this uncertainty.

Especially when tracers are sparse, the void centre definition is crucial to the estimate of the central void density and, thus, to the efficiency of our cleaning algorithm. Moreover Elyiv et al. (2015) suggest that this is an issue also in the reconstruction of the density profile which, even though not explicitly treated in this work, is obtained for each void by the rescaling algorithm that will be described later.

The void centre saved in the ZOBOV output text file is defined as the position of the least dense particle of the zone found by the VT. We call this centre the *core particle centre*, $\mathbf{r}_C$. We argue that such a definition, based on the position of a tracer, is not the most reasonable choice in our case. We want to characterize the inner region of the void which, by definition, has to be the less dense. While this might not be a problem when the density field is well sampled, it becomes an issue

when we deal with sparse samples, such as the DM halo distribution. In fact, we treat all the tracers as particles of the same mass $\overline{m}$, which depends on the mean density of matter, $\overline{\rho}_m$. It is set by the critical density of the Universe, ρ_c, and by the matter density parameter Ω_m:

$$\overline{\rho}_m = \Omega_m \rho_c = \frac{\overline{m} N}{V} \ . \tag{3.4}$$

Since $\overline{m}$ depends on the total number of tracers, N, and on the total volume of the simulation, V, the sparser the sample the greater the mass traced by each particle.

In JLH only DM distributions with m.p.s. smaller than 2 Mpc are taken into account, so the centre definition does not interfere with the results. Nevertheless, we do consider the halo distribution, so $\mathbf{r}_C$ may not always be the most convenient choice. When dealing with sparse samples, we prefer to use the *volume weighted centre* of the void, as it is defined by the VIDE algorithm:

$$\mathbf{r}_V = \frac{\sum_{j=1}^{N} \mathbf{x}_j \, V_j}{\sum_{j=1}^{N} V_j} \ ,$$

where N is the total number of particles embedded in the void, $\mathbf{x}_j$ their coordinates and V_j the volume of the corresponding Voronoi Cell. This centre does not necessarily coincide with the position of a tracer.

We have verified that in the DM distribution our results are not sensitive to a swap between $\mathbf{r}_C$ and $\mathbf{r}_V$ in the void centre definition. We will then use $\mathbf{r}_C$ when possible (i.e. when we have tracers with m.p.s. smaller than 2 Mpc) since the algorithm is simpler to implement and the results can be directly compared with that of JLH. In other cases (i.e. for sparse samples, such as halo catalogues) we will use the volume weighted centre $\mathbf{r}_V$.

3.4.2 The cleaning algorithm

As already mentioned, we follow the procedure described by JLH to clean up the output catalogues of ZOBOV and VIDE. We have implemented a sorting algorithm that rejects voids according to their statistical relevance. The statistical relevance of a void is measured in terms of its *density contrast*, defined as the ratio between the central value and the maximum value of the radial density profile. ZOBOV outputs are already ordered following this criterion, so this step is not applied.

We then apply two cutting criteria:

- the central density should be less than $(1 + {\delta_v}^{NL})\,\overline{\rho}$, since we are searching voids that have a maximum embedded density equal to $(1 + {\delta_v}^{NL})\,\overline{\rho}$ (see eq. 2.31);
- the effective radius of the void should belong to the interval $[r_{min}, r_{max}]$. This step allows us to prune the catalogue from spurious small voids and from the first output in the ZOBOV text file which is a void which takes up the entire simulation[3]. The lower limit, r_{min}, is chosen as a multiple of the m.p.s. while the upper limit, r_{max}, is a fraction of the box size.

[3]This object is also documented by Neyrinck in the original ZOBOV paper who suggests to manually remove it.

It has to be noted that there are no particular restrictions to follow when selecting the limits of the accepted radii interval. It is reasonable to select an interval which has some multiple of the m.p.s. as a lower limit and some fraction of the maximum size of the simulation box as an upper limit. This free choice, however, may cause an "over-cutting" of the catalogue. In fact, the biggest voids in the results are sensible to the upper limit of the interval.

Since ZOBOV finds voids by blending a set of adjacent zones with a hierarchy where the most underdense zone defines the void central region and the others are considered sub-voids, the former is only listed once in the catalogue while the others are provided as offspring of the larger structure. If the cutting algorithm rejects the parent void, moving lower in the catalogue it will eventually reconsider all the zones except the deepest one. We implemented in the code the possibility to save this deepest zone if the void where it is embedded got pruned by the selection criteria. Of course this zone is saved only if it satisfies itself the cutting criteria. One should refer to the original JLH paper for more details.

Even though the centre definition problem is mentioned in JLH too, we think they underestimate the effect it may cause on the cleaning algorithm. In fact, as they use a set of N-body simulations with low m.p.s., the results are obviously not affected by this choice. This statement fades when dealing with sparse samples.

The density associated to the core particle may be too high and not representative of the true central density of a void, resulting in erasing from the catalogue more candidates than necessary. For sparse samples we then use the central density given by VIDE, namely the density of a spherical region centred in $\mathbf{r}_V$ and with radius $0.25\,R_{eff}$.

3.4.3 The rescaling algorithm

Our rescaling algorithm makes large use of the CBL, especially to deal with catalogues and with the tracers' distribution. The cleaned catalogue of voids is stored in a object of the class Catalogue. We have implemented a new class of objects (`cosmobl::Void`) which allows direct access to a set of defining properties of voids (e.g. comoving coordinates, radius, statistical relevance, central density).

We can divide the procedure in two main parts: the former produces the proper rescaling while the latter checks for overlappings and eventually rejects unsuitable voids.

The rescaling procedure extracts one by one the voids in the cleaned catalogue and then procedes as follows:

- It firstly generates a chain-mesh (class `cosmobl::ChainMesh`, Marulli et al., 2016) which allows to quickly search objects in the tracers distribution. This is done by dividing the simulation into cubic cells and then considering only tracers which belong to a set of cells close to the void center. The maximum distance considered is set as some multiple K of the effective radius provided by ZOBOV R_{eff}. This section provides a list of particles of the simulation at a maximum distance $K \cdot R_{eff}$

- It reconstructs the density profile of the spherical region centred in the void centre and with radius $K \cdot R_{eff}$, tracer by tracer. The largest radius which

encloses a density $(1+\delta_v^{NL})\,\overline{\rho}$ is selected as the new rescaled radius of the i-th void.

- When all the voids in the catalogue have been rescaled, a new catalogue with the updated radii is saved.

This part of the algorithm provides a set of voids that are spherical and enclose the shell crossing density, as requested by the spherical evolution model. Moreover they do represent the first crossing of the underdensity, threshold barrier δ_v, as requested by the excursion set model, since the maximum scale which encloses the requested density is selected as new radius.

Overlapping objects are then removed by scanning the catalogue of rescaled voids. A chain-mesh of the voids centres is constructed around each void, with a maximum distance given by the largest radius found by the rescaling algorithm. When two void centres are closer than the sum of the associated radii, the less underdense is rejected.

When all the voids have been scanned, the final catalogue is saved.

Chapter 4

Vdn model validation

The homogeneous and isotropic solutions of Einstein equations of the General Relativity (1.2) lead, in the last decades, to the development of the Standard Cosmological Model. According to this model, primordial fluctuations in the density field gradually evolve into today's structures (Section 2.1).

Cosmic voids represent a promising cosmological probes for a wide range of phenomena (see the introduction of Chapter 2). While the statistical distribution of galaxy clusters as a function of mass and redshift, is commonly used to derive cosmological constraints, a similar approach with voids is hard to apply. Indeed the generally accepted SvdW model fails in reproducing the number density of voids detected in cosmological simulations. Since voids found in simulations are usually under-predicted by the SvdW model, previous works could only fit the observed distributions of voids with a SvdW size function in very relaxed conditions (i.e. letting the lower value of the threshold δ_v free). This undermines the possibility of using voids as cosmological probes.

In their paper, JLH suggested that the Vdn model is capable of reproducing the voids size distribution in N-body simulations. This requires a selection of objects which are compatible with the spherical evolution theory.

In this Chapter we apply our algorithm to the snapshots of the set of N-body simulations descibed in Table 3.1, and show the results. We compare the results with both the Vdn and the SvdW models, for different cosmological scenarios and at different redshifts (Section 4.1). In Section 4.2 we extend the results obtained for a narrow range of radii to a broader interval. The results confirm that the Vdn model reproduces the data, but the resolution of the used tracers plays an important role (Section 4.3). In Section 4.4 we discuss the general predictions of the model when the cosmological parameters are varied.

4.1 Void size function in different cosmologies

In the previous Chapter, we outlined a procedure that should not depend on the cosmological model and that, in principle, should map the void distribution in cosmological N-body simulations. In particular, we want to test if the distribution of spherical non-overlapping voids, found with the void finder we have implemented (Section 3.4), is well predicted by the Vdn-model.

To this end we fix a reference simulation for different cosmological models to try out our machinery, namely the `Lmm` for a ΛCDM model, the `Emm` for the EdS model and the `Omm` for the Open model (see Table 3.1 for the simulation characteristics and Table 3.2 for the cosmological parameters).

In this first part of our analysis, we test the predictions of both the Vdn and SvdW model. While we fix the underdensity threshold at the level defined by the shell-crossing condition (i.e. $\delta_v^{NL} = -0.79$ and $\delta_v = -2.71$ as in eq.s (2.32) and (2.31), respectively), we let the overdensity threshold, δ_c, vary in the interval included between 1.06 (i.e. the turn-around linear density contrast for a spherical perturbation) and 1.69 (i.e. the linear density of a collapsing spherical perturbation). This is done to account for the discrepancies in the small radii region of the size function, which may be caused by having considered un-correlated steps in the excursion-set model (Jennings et al., 2013). In fact, Paranjape et al. (2012b) provide an extension of the excursion set double-threshold approach, with a better treatment of the void-in-cloud process. They argue that, to model the process of void squeezing, resulting from the collapse of larger overdense perturbations where a small void may be enclosed, the correlation between the steps of each random walk has to be taken into account.

Nevertheless, only the smaller scales are interested by these problematics (Paranjape et al., 2012a), and we do not account for the cited modification of the model, leaving it as a future development of our work.

We start showing the results for the ΛCDM model at increasing redshifts, $z = 0$, $z = 0.5$, $z = 1$ and $z = 1.5$, in Figures from 4.1 to 4.4 respectively.

We have divided the radial coordinate in equispaced logarithmic bins. In the top panel we show the differential void size function, defined as the number of voids with radius included in each logarithmic radial bin per unit of comoving volume and divided by the logarithmic extent of the bin. We cut the distribution measured in the simulation at twice the m.p.s. of the given simulation box (i.e. 2 Mpc/h).

We have overplotted the predictions of the two models: the Vdn model (black region) and the SvdW model (red region). The extension of the shaded region derives from the uncertainty on the definition of the overdensity threshold.

To highlight the difference between the theoretical models and the measured distribution we also show in the two lower panels the logarithmic difference between the simulation data and the two models. As we can see, the SvdW model (central panel) systematically overpredicts the total number of voids at all the scales, while the Vdn model (bottom panel) gives a better prediction of the distribution. Note that the scale on the y-axis of the two lower panels in all the Figures are different.

Since voids are wide, underdense, unlocalized regions, they particularly suffer limitations in the spatial resolution of the sample. At scales comparable to the spatial resolution of the simulation, voids are not well represented by the tracers sampling of the underlying density field. This causes a decreasing trend on the number counts, which can be guessed at the smallest scales in the shown plots, in spite of the applied cut.

At large scales, instead, the measurements are limited by the simulation box extension. Since neither VIDE nor our algorithm authomatically account for the periodical conditions of the N-body simulation, the largest peripheral objects of the

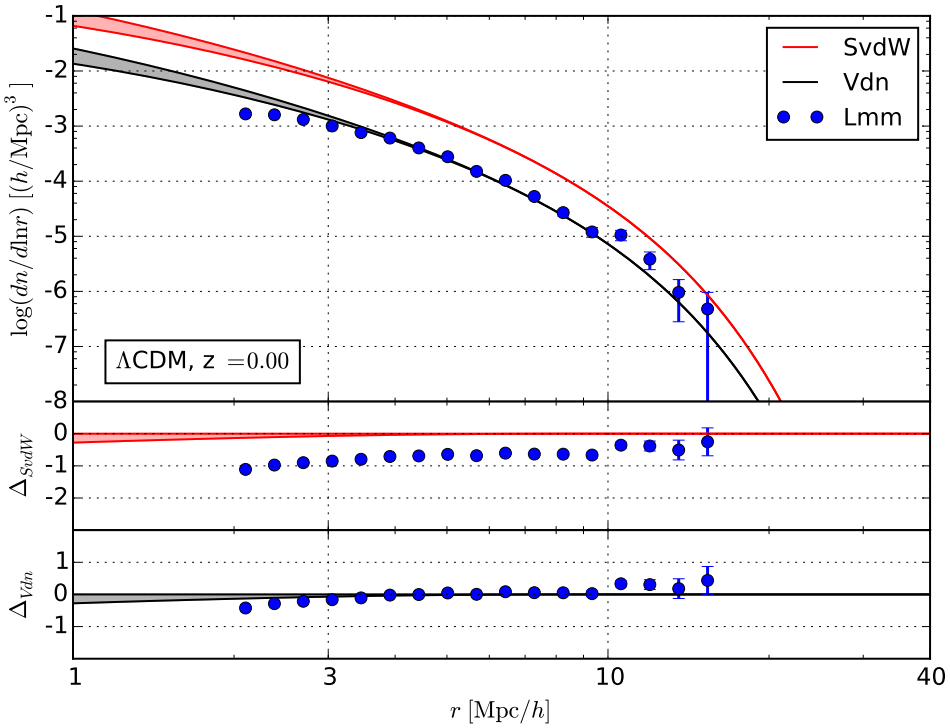

Figure 4.1: Top panel: Number density of voids at $z = 0$ in the `Lmm` *simulation box. Central panel: difference between the logarithmic values of the number counts and the SvdW model. Bottom panel: same as central panel but w.r.t. the Vdn model.*

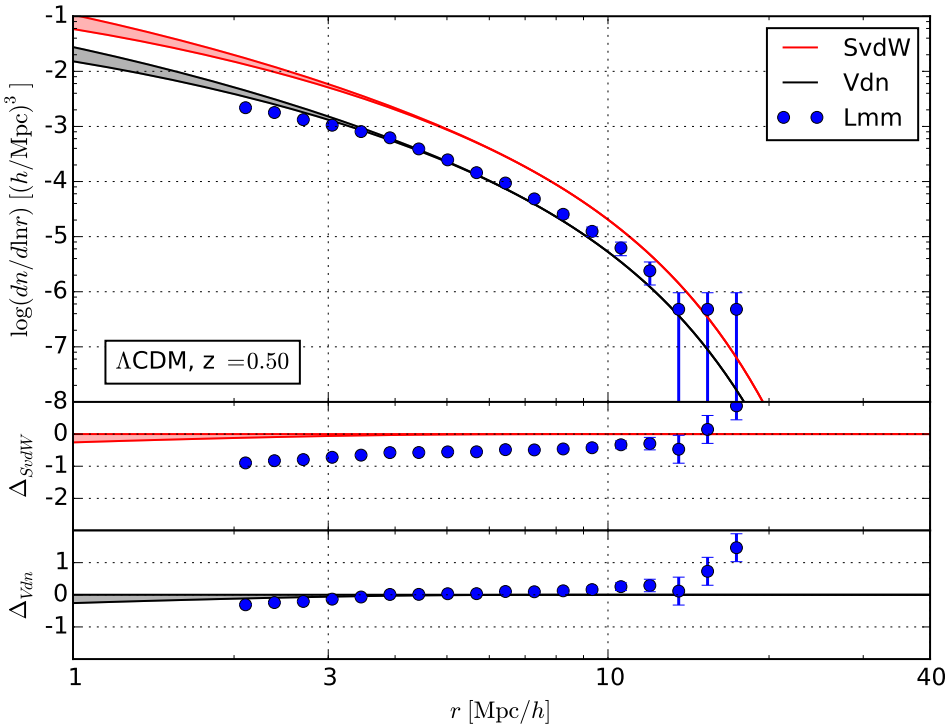

Figure 4.2: Same as Figure 4.1 at redshift $z = 0.5$.

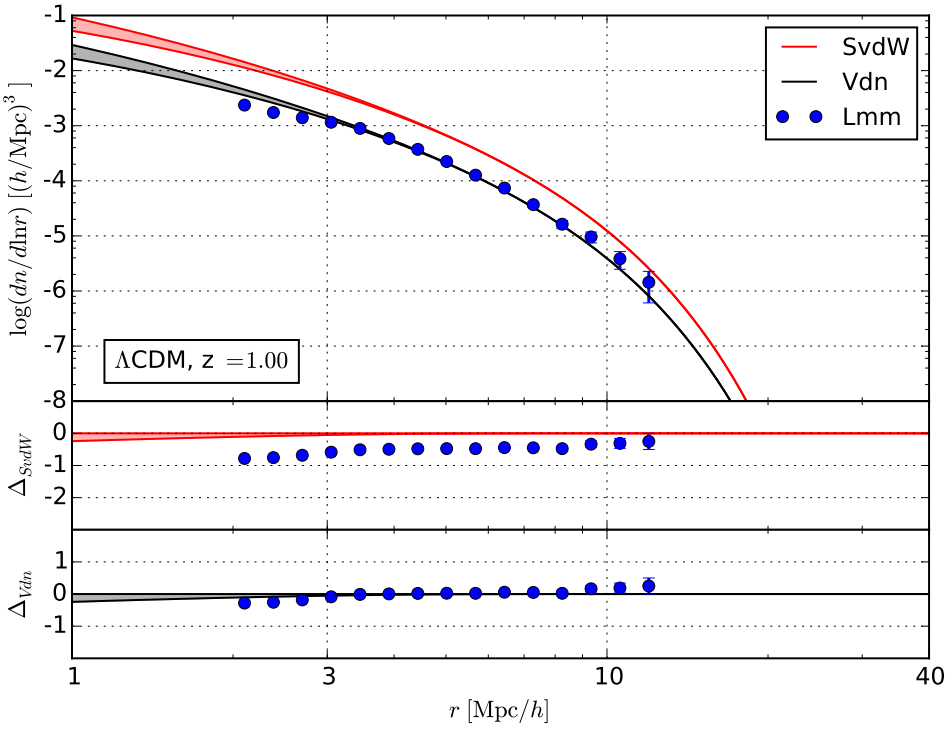

Figure 4.3: Same as Figure 4.1 at redshift $z = 1$.

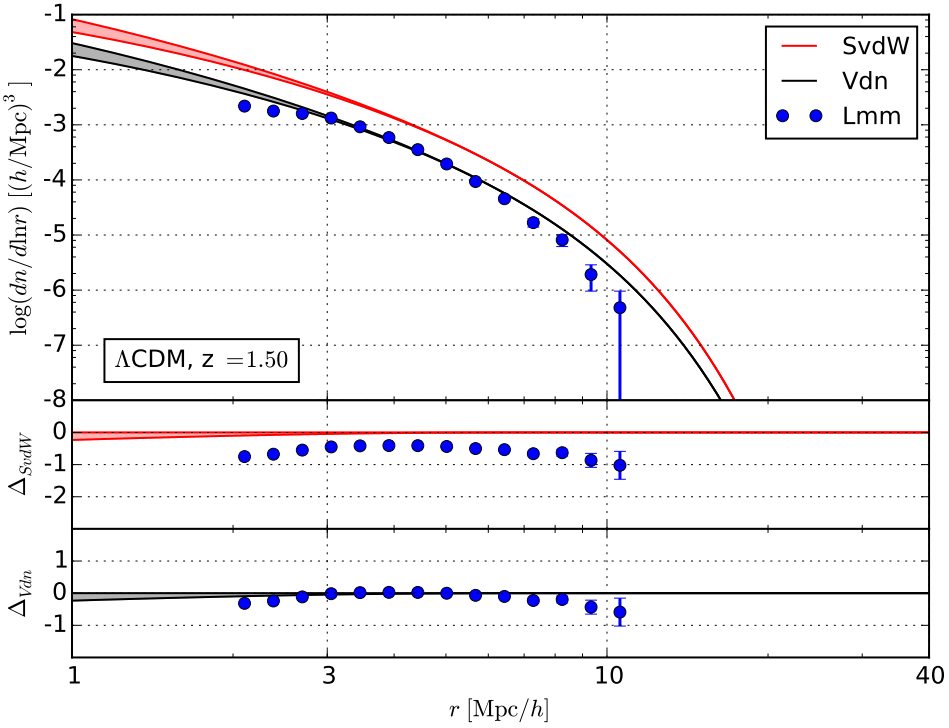

Figure 4.4: Same as Figure 4.1 at redshift $z = 1.5$.

simulation box may trespass on the boundaries of the box. Given the periodicity of the simulation box, what goes out of the box at one side, re-enters it at the other side. As a consequence of their widespread nature, outer voids are divided in smaller voids next to the boundaries of the simulation box. Contrary to the limits at small radii which cannot be overcome without increasing the resolution of the sample, this inaccuracy in the void counts at large radii can be faced with an upgrade of the void finding algorithm.

Considered this issue we can conclude that the Vdn model optimally predicts the data at all redshifts. Nevertheless at the higher redshift considered (Figure 4.4) we note a decrease in the number of voids w.r.t. the number predicted at large radii by the Vdn model. We reserve the study of this point to a future investigation.

For what concerns the time evolution of the void size function, the density of larger voids decreases with increasing redshift. On the other hand, the smaller scales are less affected. This suggests that the evolution of voids in ΛCDM models is already advanced at high redshift. The Vdn model well predicts this behaviour.

To extend our results to different cosmologies, we apply the same procedure just described for the ΛCDM model to other two scenarios.

We apply our algorithm and obtain the void distributions for the Emm simulation box (see Table 3.1). We show our results in Figures from 4.5 to 4.8 for the same redshift values considered in the ΛCDM model. Even though in this case the void

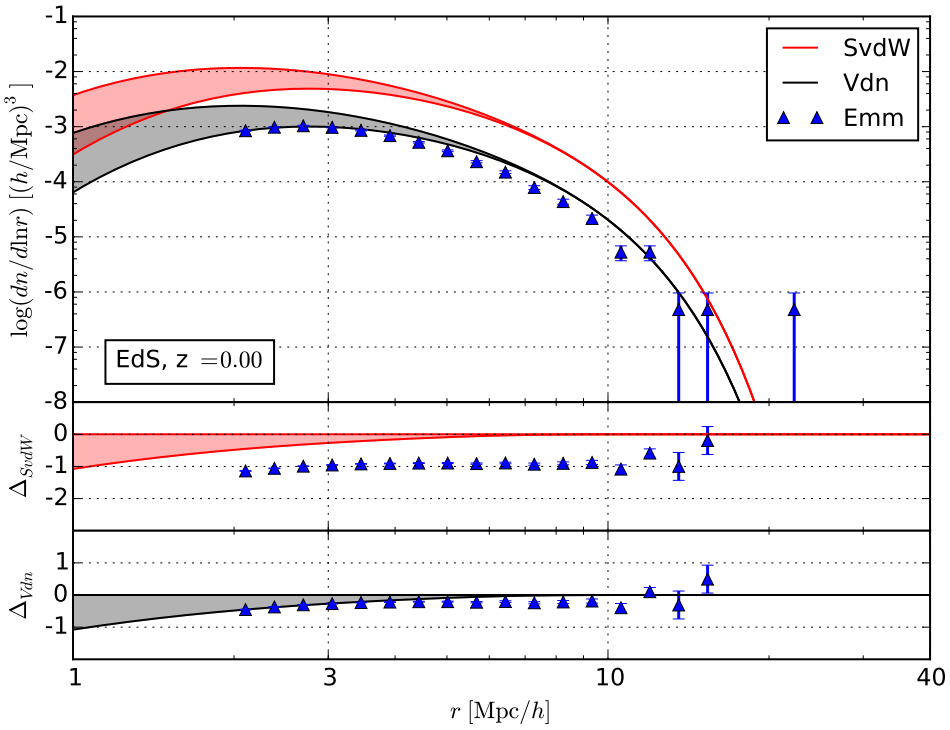

Figure 4.5: Top panel: Number density of voids at $z = 0$ in the Emm simulation box. Central panel: difference between the logarithmic values of the number counts and the SvdW model. Bottom panel: same as central panel but w.r.t. the Vdn model.

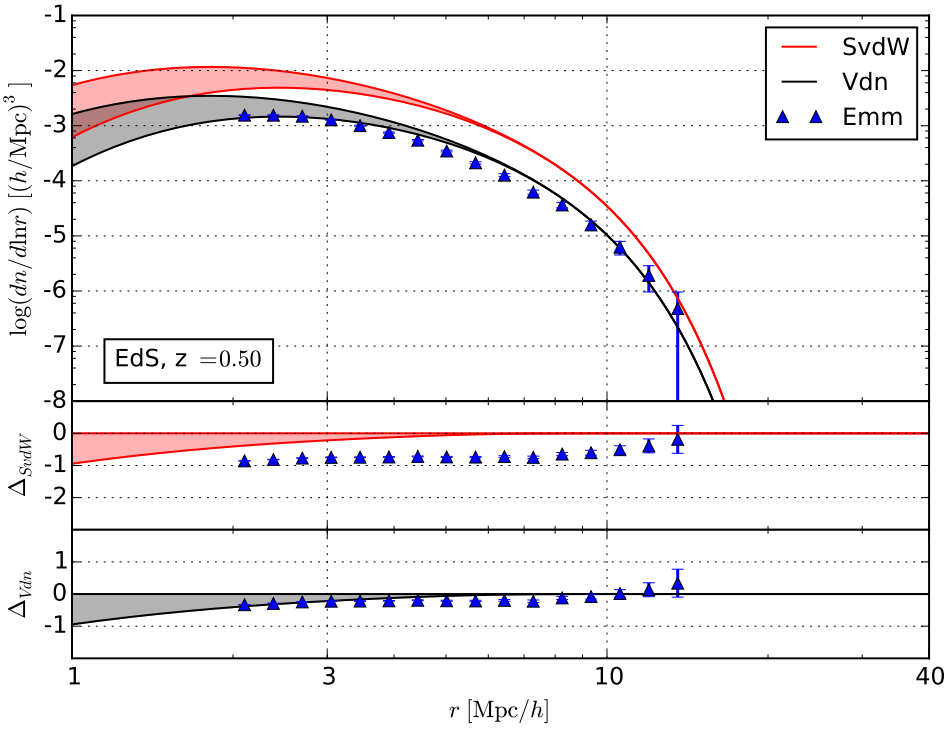

Figure 4.6: Same as Figure 4.5 at redshift $z = 0.5$.

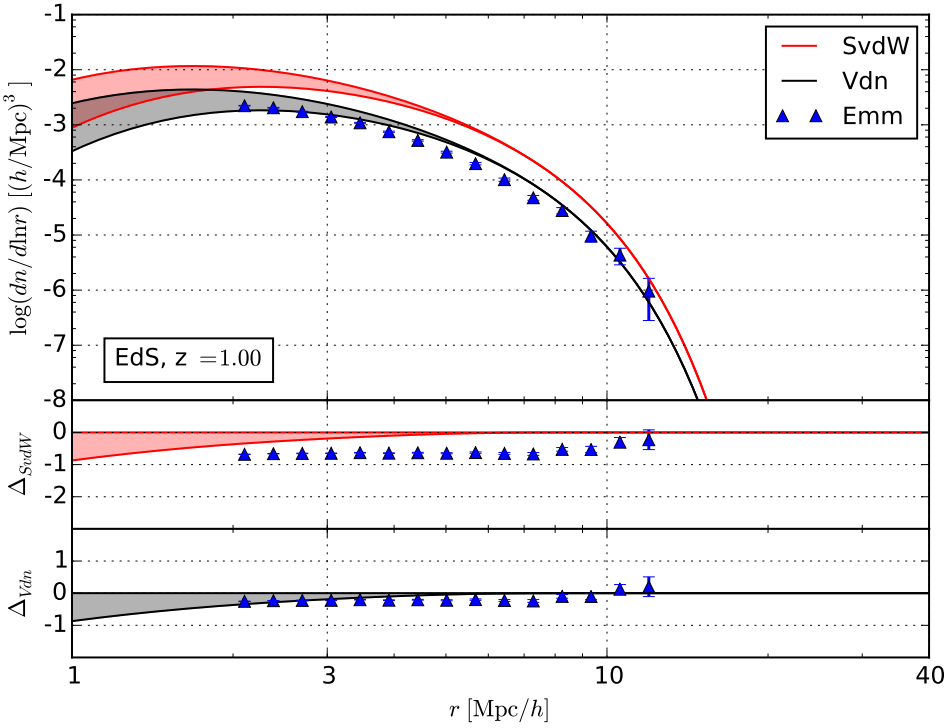

Figure 4.7: Same as Figure 4.5 at redshift $z = 1$.

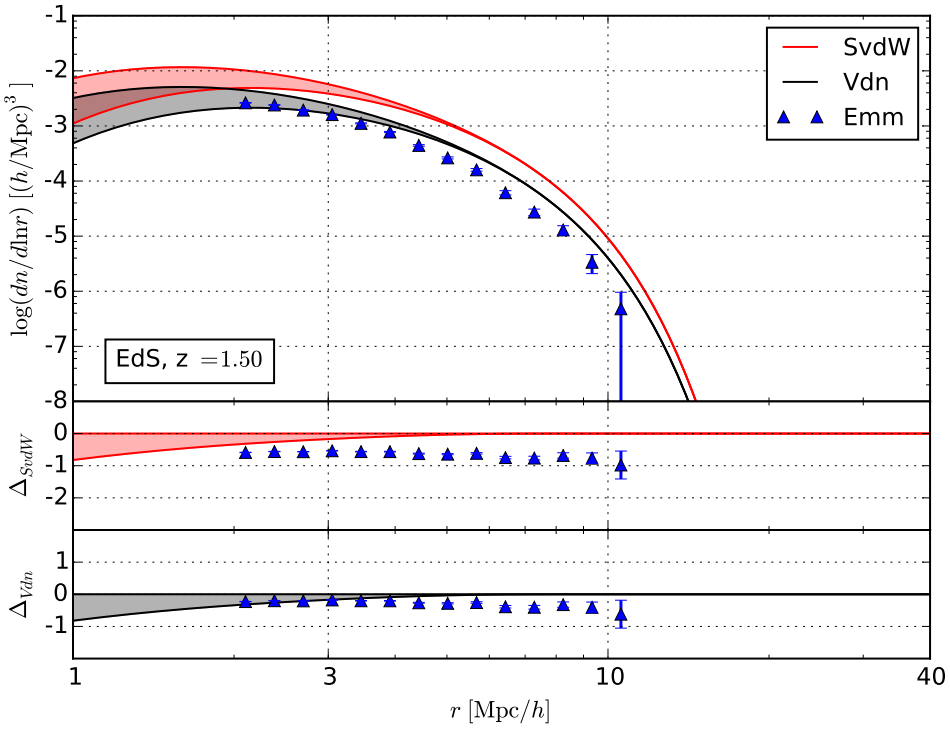

Figure 4.8: Same as Figure 4.5 at redshift $z = 1.5$.

counts seem to be on average slightly overpredicted by the Vdn model, its predictions are largely better than the SvdW model which, once again, completely fails at predicting the size distribution at all redshift. We can notice the systematical decreasing trend in the counts at small radii which, in the EdS case, are marginally included in the grey reliability range of values. Furthermore, also the behaviour at large radii w.r.t. the model is again similar to that of the ΛCDM model, including the peculiar deficiency at the largest redshift.

Also in this case the redshift evolution of the size distribution is acceptably reproduced by the Vdn model. We again notice how the largest voids in the catalogue shrink from small to high redshift.

The last cosmological model we examine is an Open model with parameters provided in Table 3.2. Figures from 4.9 to 4.12 show the void counts and void size function models for the same set of redshift values considered before.

We notice that the general increase in the void density at large radii is, for this particular cosmological model, more emphasised. Nevertheless, an open model typically is more evolved than a flat model, such as the other two models inspected. Given that an underdense region receives a burst in its outward acceleration from the general expansion of the background Universe, this shall result in broader voids.

We hypothesize that the Vdn model's under-prediction of the measurements, for this particular cosmology, is not physical but once again an effect of not having considered the boundaries' periodicity of the simulation boxes.

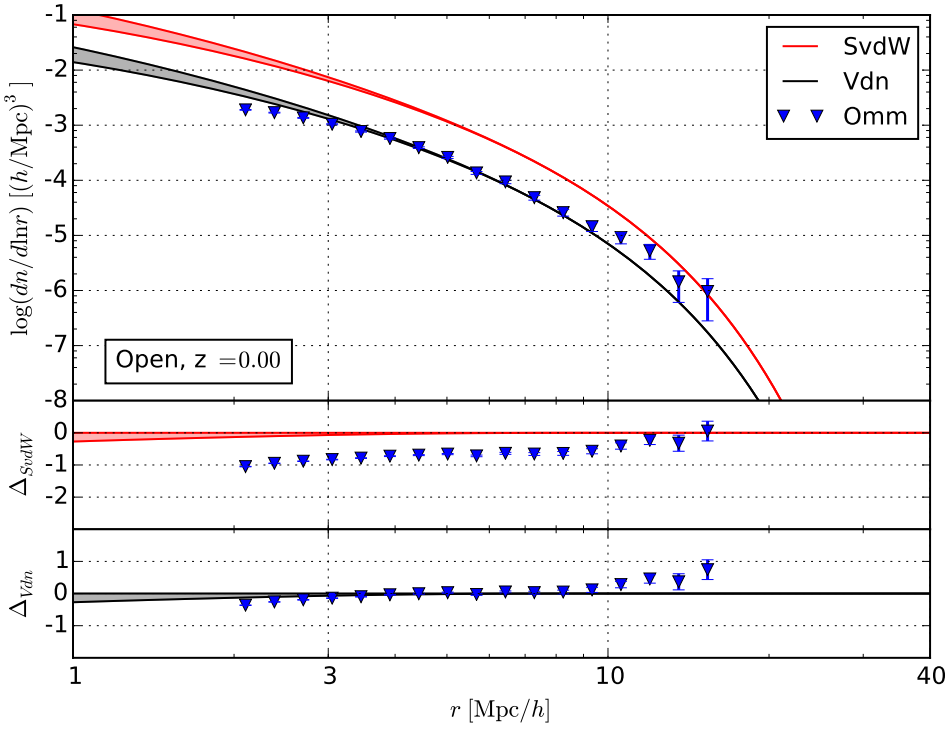

Figure 4.9: Top panel: Number density of voids at $z = 0$ in the Omm simulation box. Central panel: difference between the logarithmic values of the number counts and the SvdW model. Bottom panel: same as central panel but w.r.t. the Vdn model.

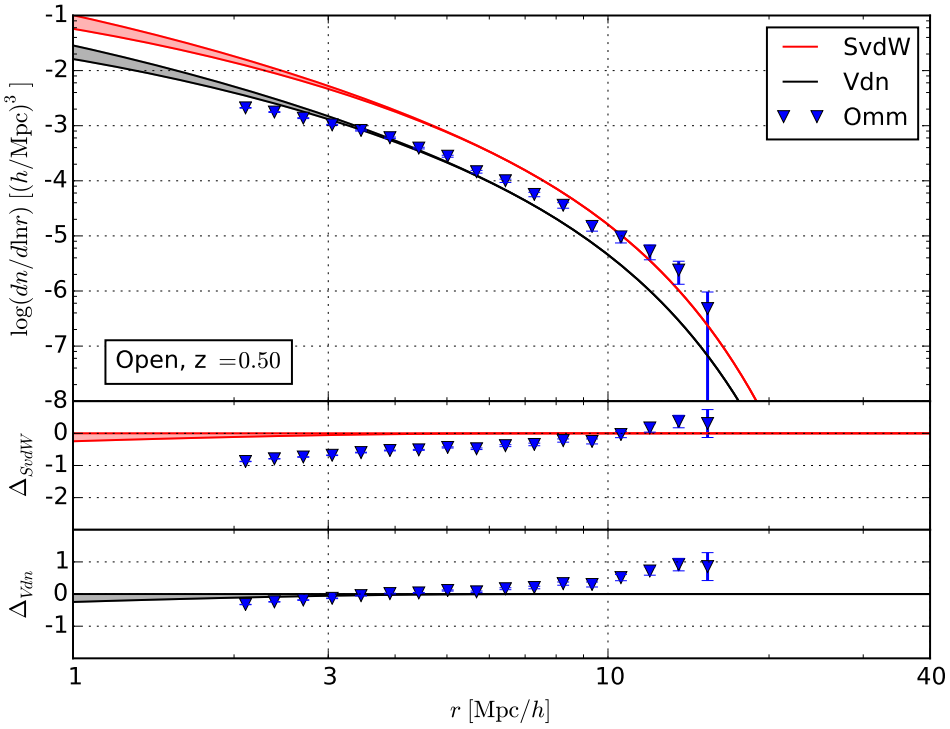

Figure 4.10: Same as Figure 4.9 at redshift $z = 0.5$.

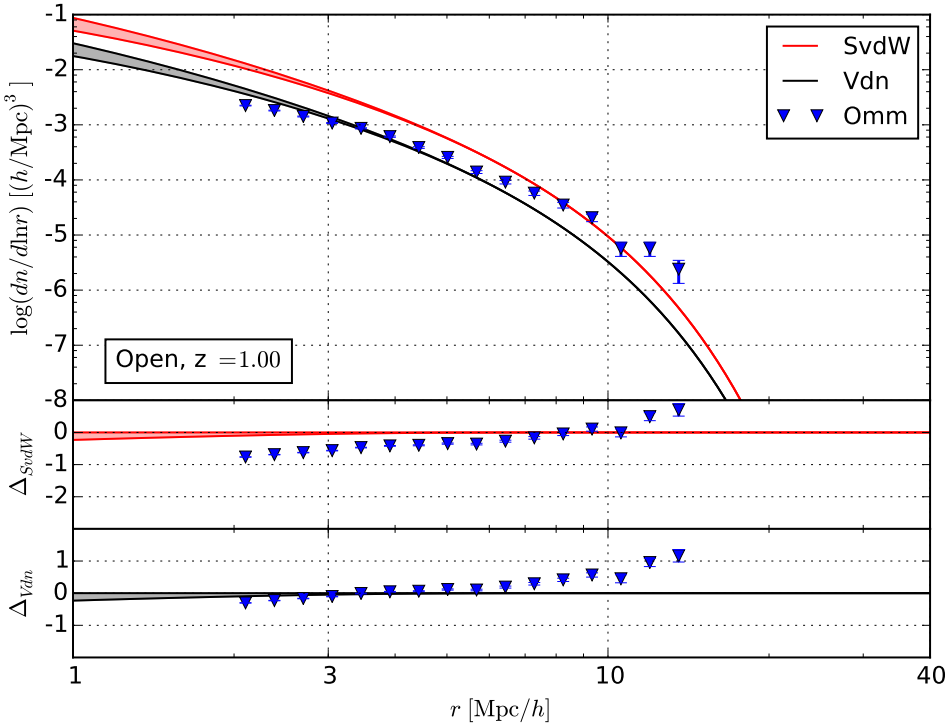

Figure 4.11: Same as Figure 4.9 at redshift $z = 1$.

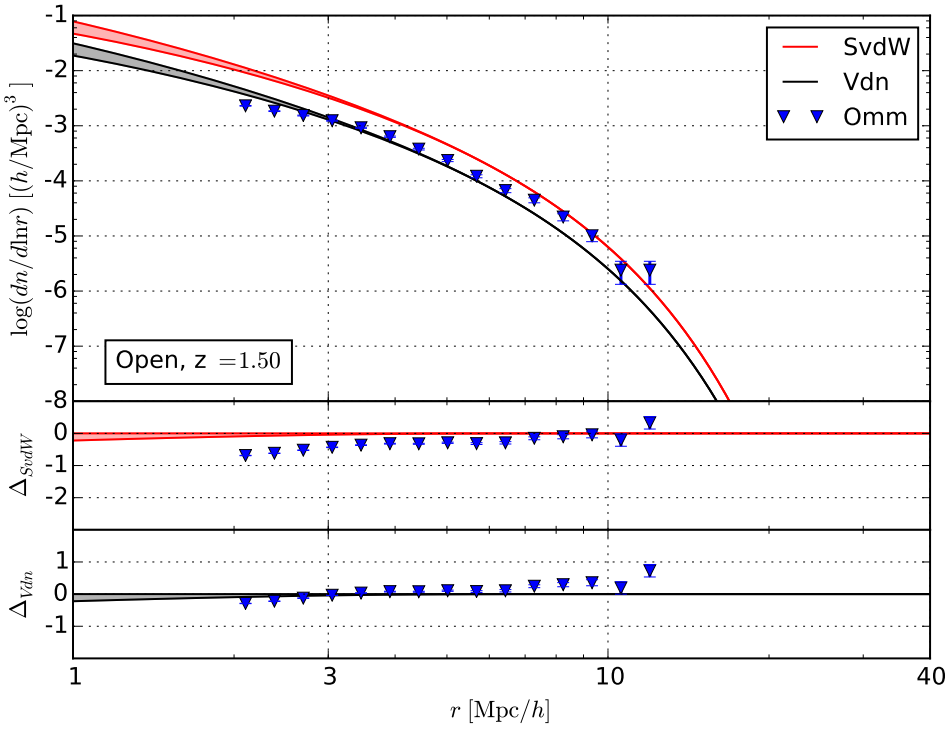

Figure 4.12: Same as Figure 4.9 at redshift $z = 1.5$.

4.2 Extending the model validation over a larger range of radii

We have established that, for standard cosmologies, the Vdn model predicts the void size distribution over a considerable range of redshifts. Nevertheless we only have shown measurements in a limited radii range. Extending these limits is desirable for several reasons, first of all to strengthen our results but also because at different values of the redshift the theoretical relation depends differently on the defining parameters.

For now let us just note that a different behaviour at varying radii is expected by the analytical form of the theoretical model of eq. (2.54), which is indeed approximated in eq. (2.55) by a piecewise function (the dependence on the radii is implicitly contained in $x \propto \sigma = \sigma(r)$). Furthermore, we have already had the possibility to appreciate how the model behaves as a consequence of the modification of the overdensity threshold δ_c. In all the figures shown, only the small scales are affected by such a variation in one of the parameters. We will examine deeper the parameter dependence of the Vdn model in the last Section of this Chapter.

We have already described how the reliability of the data depends both on the m.p.s. of the simulation at small scales, and on the simulation box size at large scales. The only possibility of mapping the statistical distribution of voids in simulations over a broader range of radii is to obtain void catalogues in several simulation boxes, with different m.p.s. and sizes.

Ideally we would obtain only one large simulation with a great number of particles. This would provide a test bench with a small m.p.s., reliable up to the largest scales. However, given the extremely long computing time and the great amount of memory required[1], such a simulation is not simple to be obtained.

Nevertheless a similar result can be obtained with a set of catalogues covering a larger range of sizes and resolutions. We then fix the ΛCDM cosmological model as our test bench and run our algorithm for all the simulation boxes of Table 3.1 (except the `Lhm` box), with box side length from 64 Mpc/h to 500 Mpc/h and with m.p.s. from 0.25 Mpc/h to 2 Mpc/h.

In Figure 4.13 we have superimposed the results coming from all the simulation boxes. Having established the generally better reliability of the Vdn model w.r.t. the SvdW model, we only show the logarithmic differences w.r.t. the Vdn model. It is possible to notice how the new results do actually cover a wider range. For comparative purposes we also show the void distribution measured for the `Lmm` simulation box (blue dots).

At small radii we extend our measurements from the resolution limit of the `Lmm` box (2 Mpc/h) to that of the `Ltv` box (green dots, 0.5 Mpc/h). Having mapped the void size distribution at smaller radii, we can observe how the decreasing trend of the blue points, at radii smaller than approximately 3.5 Mpc/h, is not present in the simulation boxes with smaller m.p.s., as we were expecting. Nevertheless it appears again around 3 times the m.p.s. of each simulation.

Furthermore, at large radii we extend the mapping from 15 Mpc/h to 20 Mpc/h.

[1] Each particle is described by three coordinates which can be stored as `float` numbers. Since each `float` requires at least 4 bytes, a simulation, like the `Lhm`, with approximately two billions particles requires at least 24 Gigabytes of memory to store all the data.

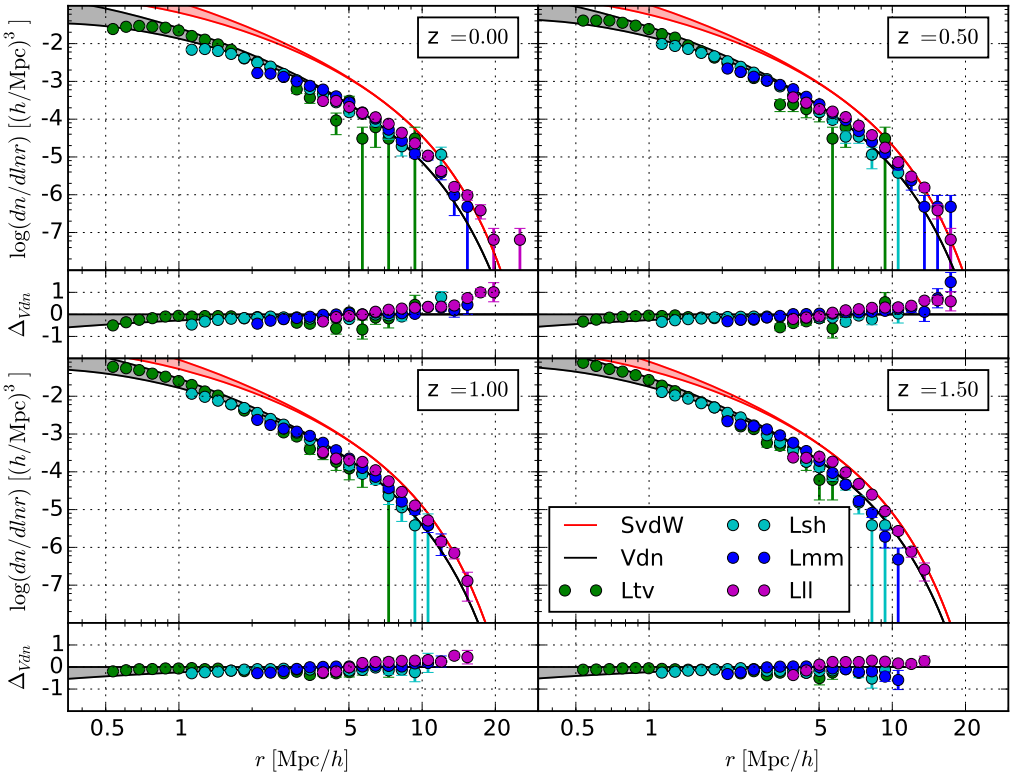

Figure 4.13: Top panel: differential void size distribution for all the simulation boxes considered, compared with the Vdn model (grey shaded region) and the SvdW model (red shaded region). Bottom panel: logarithmic difference between the simulated measurements of void size distribution and the Vdn model. The results for four different values of redshift are shown.

Nonetheless, the void distribution of the `Lll` simulation box (magenta dots) shows a stronger increasing trend than the others, w.r.t. the Vdn model which underestimates the counts. This discreapancy might be a side effect due to the large m.p.s. of the `Lll` box, which has the smallest resolution between our set of simulations.

We encountered similar difficulties at the very beginning of our research work, when we were analysing the large `Lhm` simulation box. Since it was necessary to sub-sample the particle distribution to overcome the already exposed computational issues, the resulting sample had smaller resolution and was not well predicted by the model. The development of smaller samples, with higher resolution, let us bypass these difficulties.

Nevertheless, these problematics deserve more investigations, that will be addressed in the following Section.

4.3 Problems related to resolution and subsampling

In previous works (Sutter et al., 2014a; Pisani et al., 2015; Nadathur and Hotchkiss, 2015) the voids distribution has been studied by subsampling the DM particles to match a number density of tracers comparable to that expected for tracers like galaxies in future surveys. None of them find agreement between the SvdW model and simulated data, unless the underdensity threshold, δ_v, was considered as a free parameter to be fitted. This results in a threshold which is approximately between -0.45 and -0.5 when tracers are subsampled to match an average number density of $n = 4 \cdot 10^{-3}$ $(h/\mathrm{Mpc})^3$.

Doubts concerning the application of a subsampling procedure when searching voids with a VT method have been raised by Achitouv et al. (2015). They argue that the procedure of JLH may be correct, since it searches for spherical underdensities similar to the theoretical definition of voids, but it may suffer from the un-sphericities generated by the voids evolution.

Nevertheless the authors concentrate only on the density field mapping by the VT algorithm with void finders such as ZOBOV or VIDE. Their point is that, since voids should be separeted by thin ridges, the subsampling of the tracers may result in a loss of individuality. We object that they do not consider the natural tendency of underdensities to become more spherical while they evolve (Icke, 1984) which is instead considered by the excursion set model.

Furthermore, even if unspherical voids formed, our void finder should suffer less from this issue. In fact, it finds all the largest spheres that do not overlap and fit inside an underdense region. Nonetheless, the problems with the `Lll` simulation box encountered in the previous Section, and that only sketched out with the `Lhm` box, together with the exposed arguments raised by other authors, push towards the necessity of testing the actual effect of subsampling.

At this goal, we analyse two simulation boxes with the same side length (128 Mpc/h) but different resolutions: the `Lsm` simulation box, whith m.p.s.= 1 Mpc/h, and the `Lsh` simulation box, whith m.p.s.= 0.5 Mpc/h. The `Lsh` has been already proven to be well represented by the Vdn model. On the other hand the `Lsm` has the same resolution of the `Lmm`, which is also well represented by the Vdn model. We can then randomly subsample the `Lsh` to match the same number density of the

`Lsm`:

$$n'_{Lsh} = ss \cdot n_{Lsh} \equiv n_{Lsm} \ ,$$

where n'_{Lsh} is the number density of subsampled tracers in the `Lsh` box, which multiplied by a subsampling factor, ss, has to match n_{Lsm}. The results for the number density of voids are shown for different redshifts in Figure 4.14. What can

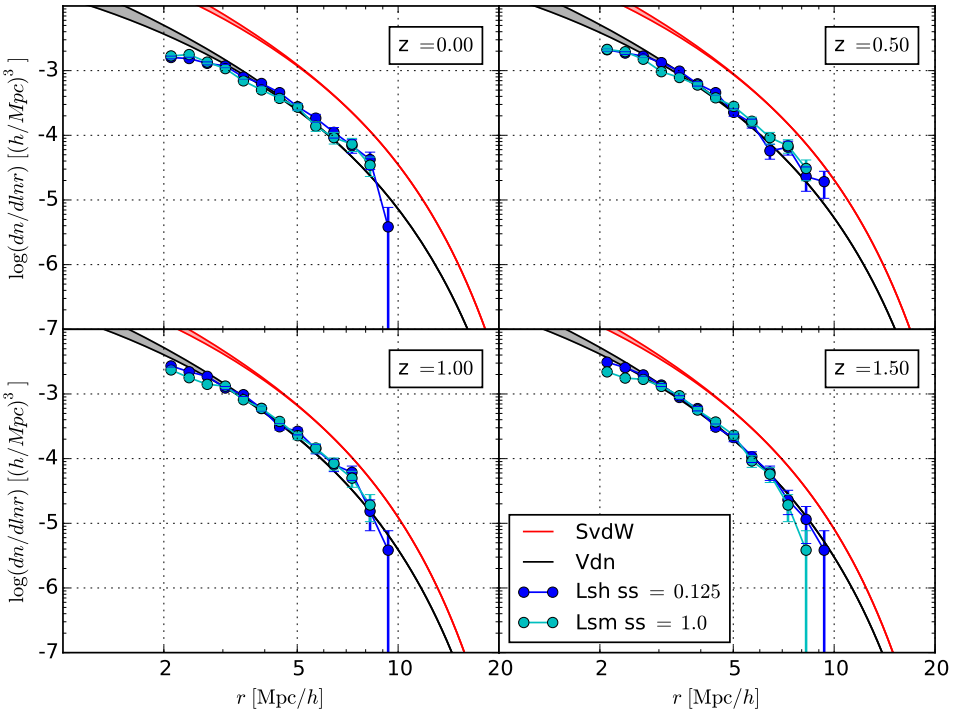

Figure 4.14: Comparison at four values of redshift between the void size distribution in the `Lsh` *simulation, with a subsampling of* $ss = 0.125$*, and the void size distribution of the full sample* `Lsm` *simulation box. The two shaded regions illustrate the predictions of the SvdW model (red) and of the Vdn model (black).*

be noted is that the two data distributions are almost perfectly overlapped at all scales and redshifts, with minimal deviations only at very low and very large values. We conclude that the subsampling of tracers does not directly cause a deterioration in the data, if a sufficiently high resolution is maintained.

We have already shown that the data from the simulation box with the lowest resolution (`Lll`) tend to be underpredicted by the Vdn model at the largest radii. Such a behaviour could be corrected by assuming a higher value of the threshold δ_v, as done by the cited authors, since, in the interested range, the model mainly depends on the δ_v value. However, in our theoretical approximation we want the threshold to be fixed, since it is this parameter that more than any other distinguishes what should be defined as "void" by merely underdense regions (Section 2.1.2).

Given that also a full sample may encounter an overcrowding of large voids when its resolution is not high enough, we conclude that the most relevant property of a sample when dealing with voids is its spatial resolution.

We show the effect of diminishing the resolution of the sample by subsampling the `Lsh` simulation box to match the number density of DM particles in the `Lll` box. The results are shown in Figure 4.15 for all redshifts.

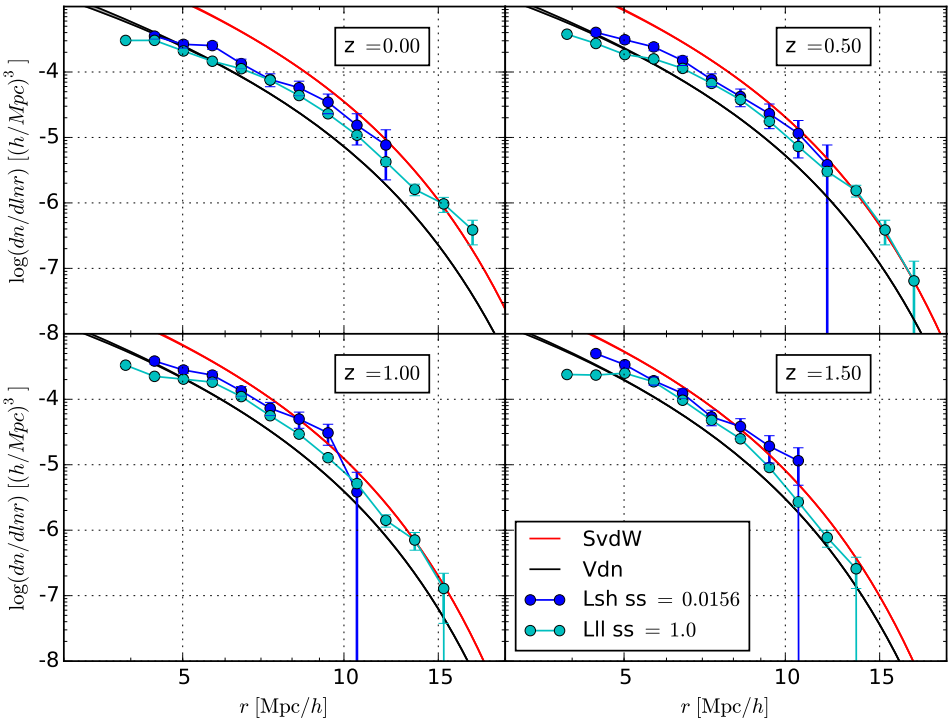

Figure 4.15: Comparison at four values of redshift between the void size distribution in the `Lll` simulation, with a subsampling of $ss = 0.0156$, and the void size distribution of the full sample `Lsm` simulation box. The two shaded regions illustrate the prediction of the SvdW model (red) and of the Vdn model (black).

The blue dots, representing the voids distribution in the subsampled `Lsh`, are underpredicted by the Vdn model while the SvdW tends to get closer. This behaviour is similar to what happens to the `Lll` taken full sample.

It has been advocated (Achitouv et al., 2015; Nadathur and Hotchkiss, 2015) that VT void finders may be incompatible with the theoretical models exposed in Section 2.3, since they rely on different definitions of voids (i.e. while the theory defines voids dynamically, as objects which have gone through shell-crossing, void finders only search for underdensities). Nonetheless, we are confident that our procedure overcomes this incosistency, thanks to the requirements of sphericity, no-overlap and shell-crossing fixed density contrast.

Subsampling the tracers does not deteriorate the data obtained if a sufficiently high resolution is maintained (Figure 4.14). Conversely, if the subsampling is too extreme, the resolution of the sample becomes too low and the Vdn model does not predict the void distribution anymore (Figure 4.15). We conclude that the not correct prediction of the void counts measured in some simulated samples is an effect of the the low resolution rather than the subsampling (as argued by some authors). Tracing the density field with sparse samples may in fact result in the accidental merger between adjacent voids which generates larger voids. This is

not in conflict with the theoretical model of void size function since the excursion set theory for voids does in fact consider the effect of mergers between voids in an exactly identical way as the PS formalism accounts for the hierarchical merger history of haloes. Nevertheless the systematical growth in the counts of large voids for under-resoluted samples damages the possibility of obtaining faithful predictions with the theoretical Vdn model. We have however proven that with a m.p.s. of at least 1 Mpc/h the Vdn model succeeds in predicting the size distribution of voids for different redshifts and without particular preferencies for the cosmological parameters of the used simulated samples.

4.4 Cosmological dependencies of the model

The promising results of the previous Sections push toward the conclusion that the theoretical Vdn model successfully predicts the void size distribution. We are therefore interested in how a variation of the cosmological parameters may modify the predictions of the model. As examples, in this Section we compare the model's behaviour in different scenarios by simply varying the value of some parameters, namely Ω_M, σ_8 and w_0. Since we have already commented how the model evolves in time in Section 4.1, we fix the redshift value at $z = 0$, and let only the other parameters vary.

In the first place, we examine how the model is affected by a variation in the matter content of the Universe. Figure 4.16 illustrates the resulting differential VSF.

We assume flatness (i.e. total density parameter $\Omega_T = 1$) and let the matter density parameter Ω_M vary, together with the dark energy density parameter, so that $\Omega_\Lambda = 1 - \Omega_M$. All the other cosmological parameters are fixed at the values of the ΛCDM cosmology in Table 3.2, which we assume as a fiducial model. In the lower box we show the logarithmic differences between the different models and the fiducial one.

We note that for cosmologies with a higher mass content the density of small voids decreases, while the number of large voids increases. This should not surprise since the larger is the matter density parameter, the quicker the overdensities collapse into haloes. Because of the *void-in-cloud* process, this results in the crushing of small voids. On the other hand, since the underdense regions are faster evacuated by matter, which preferentially flows into the collapsing haloes, the voids which survive the void-in-cloud process are systematically larger as a result of the enhanced merger rate between adjacent underdensities.

This result is also highlighted by the cumulative volume fraction in voids, $\mathcal{F}(r)$, (Figure 4.17), which measures the fraction of the Universe volume occupied by voids with radii larger than a given value r. The Vdn model requires that the total volume fraction shall be a conserved quantity and effectively, all the models reach out a constant value approximately around the 38% of the total volume. Models with a higher matter content however reach this value at larger radii w.r.t. the others, suggesting that small voids do not play a significant role.

We have also considered how the theoretical model is affected by the variation of the matter fluctuations power spectrum normalisation, as expressed in terms of r.m.s. of the mass variance on scales of 8 Mpc/h in the present Universe, σ_8. Given

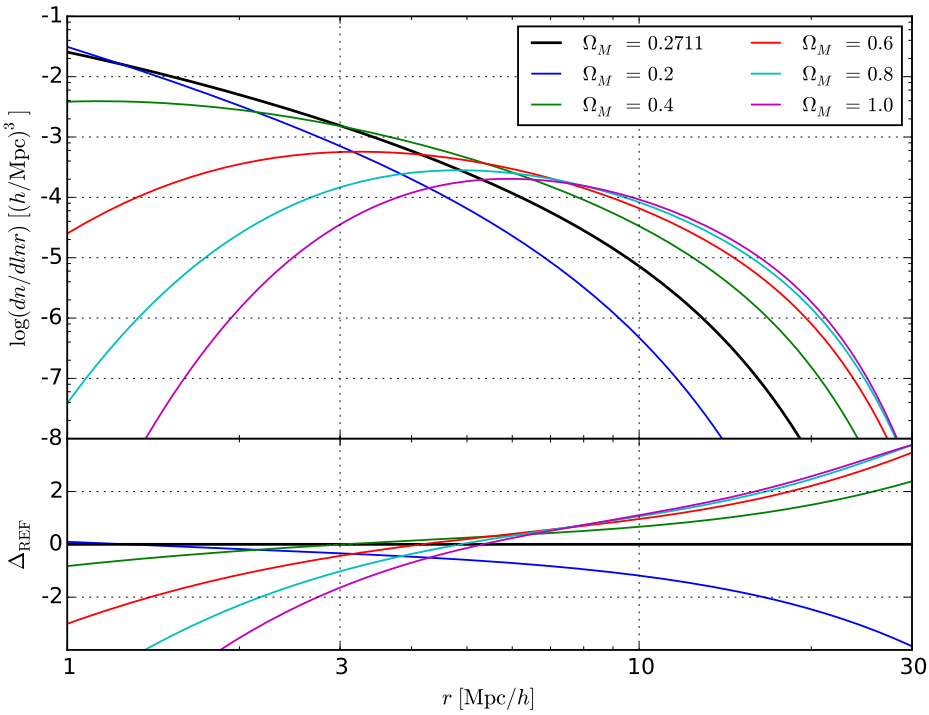

Figure 4.16: Predictions of the Vdn model with varying matter content of the Universe under the assumption of flatness. The black line represents our fiducial model (ΛCDM in Table 3.2). In the top panel the differential Void Size Function is shown. In the bottom panel we plot the differences between the logarithmic values of the curves w.r.t. the fiducial model ($\Omega_M = 0.2711$).

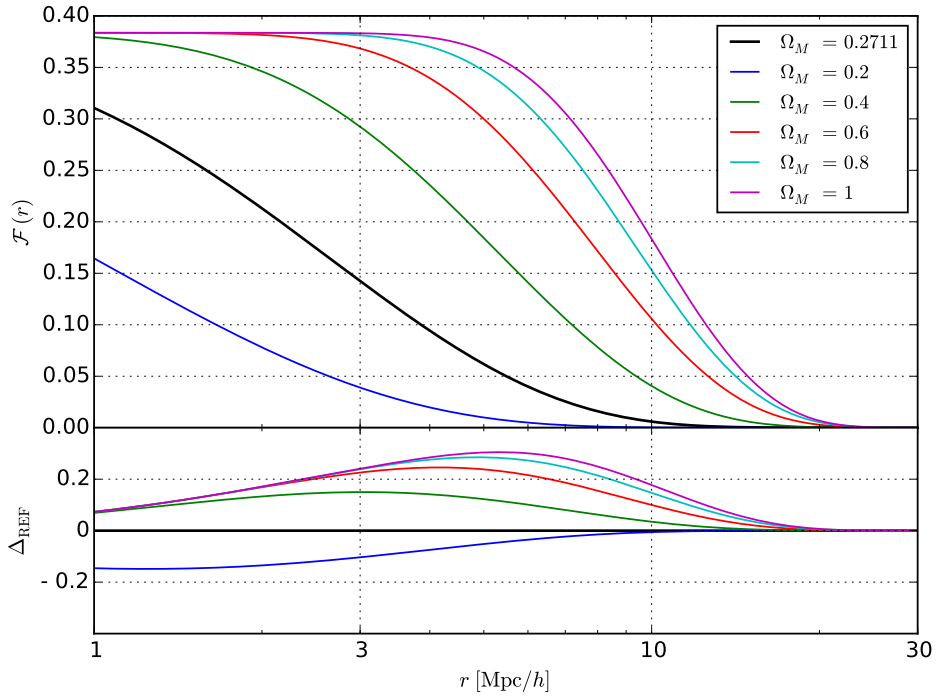

Figure 4.17: Volume fraction occupied by voids with varying matter content of the Universe under the assumption of flatness. The black line represents our fiducial model (ΛCDM in Table 3.2). In the top panel the cumulative volume fraction in voids (eq. (2.58)). In the bottom panel the difference between the curves w.r.t. the fiducial model ($\Omega_M = 0.2711$) is shown.

their extremely large scale, cosmic voids cover the smallest wavelengths of the power spectrum and should then provide useful constraints on its amplitude in the region where it approaches the self-similarly evolved part. The local abundance of voids should then deliver an essential cosmological information, since consistency with this constraint is one of the most important tests a model can pass.

In Figure 4.18 we show the curves obtained varying σ_8 in a range between 0.6 and 1. We notice that they are expecially sensitive at the higher values of the radii. In the lower box we stress the logarithmic differences between the set of models and the fiducial model. The intersection of the curves marks approximately the point where the void-in-cloud process becomes relevant (i.e. approximately 2 Mpc/h in such a ΛCDM model). It is then clear that, to put constraints on the value of σ_8 using the cosmic voids distribution, it requires a catalogue of the largest observable voids (i.e. hopefully with radii larger the 10 Mpc/h).

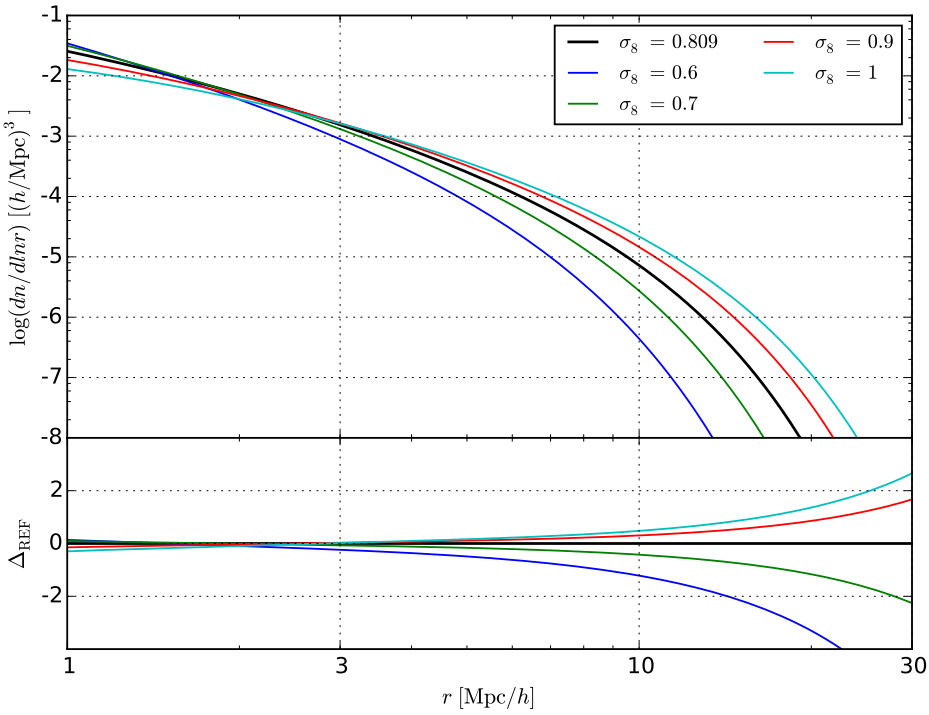

Figure 4.18: Same as Figure 4.16 for varying normalisation of the power spectrum of matter fluctuations, σ_8.

Lastly, we explore the effects on the void size function of a variation of the equation of state parameter of the DE component, w_0. This parameter governs the rate at which the DE energy density evolves and is required to be a smaller than $-1/3$ in order to induce an accelerated expansion. For a perfect, unchanging vacuum energy, like the cosmological constant Λ, we must have $w_0 = -1$, meaning that the pressure is equal in amplitude and opposite in sign to the energy density. If the real measure of w_0 is slightly different from this reference value, we would expect extremely different scenarios for the future evolution of the Universe. Thus measuring these possible deviations of the parameter w_0 is crucial for our understanding

the nature of the DE component and the future evolution of our Universe.

In Figure 4.19 we illustrate how the predictions of the Vdn model vary as a function of w_0. The lower panel stresses the difference between the models, showing that the size function of voids is more sensitive to an increase of w_0 w.r.t. the reference value (black curve). Note that varying by a factor of two towards smaller values of w_0 (blue curve) has an effect comparable to a variation by a factor of 0.25 (red curve) towards larger values of w_0.

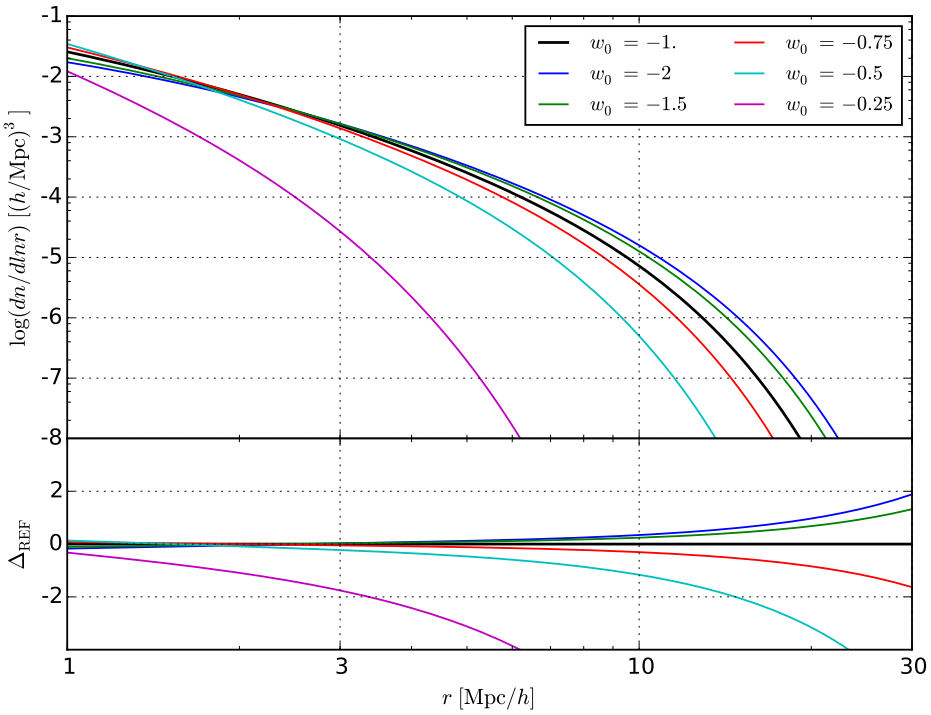

Figure 4.19: Same as Figure 4.16 for varying DE eq. of state parameter, w_0.

Chapter 5

Voids in the halo distribution

Having proved the efficiency of the Vdn model for the void size function in the DM particles, we focus now our research to a ΛCDM cosmology at redshift $z = 0$. The ultimate goal of this part is to find a model capable of predicting the void size distribution for real tracers of the underlying density field, such as galaxies or clusters of galaxies.

As we have seen in the previous Chapter, a statistical study of voids in terms of their size, first of all requires a sample of tracers with a sufficient resolution. In the last decades the resolution and spatial extension of surveys grew considerably. Voids can be identified in these surveys as regions of space containing fewer galaxies than average.

Several void catalogues have been compiled with data from the Sloan Digital Sky Survey (SDSS) Data Release 7 (DR7) (Pan et al., 2012; Sutter et al., 2012; Ceccarelli et al., 2013; Nadathur and Hotchkiss, 2014), but the methods used to define and identify voids in the original galaxy distribution vary in each case. Recently the DR11 and DR12 releases from the SDSS-III Baryon Oscillation Spectroscopic Survey (BOSS) have covered much larger volumes of the Universe, spanning a larger redshift range. Based on these data a large void catalogue, compiled by Nadathur (2016), was made pubicly available. The vast majority of these catalogues are obtained using the ZOBOV void finder or similar algorithms. In particular Nadathur (2016) use the ZOBOV void finder regulated to not merge any adjacent zone (see Section 3.3), thus retaining as much information as possible on the underlying topology of the reconstructed density field.

In Table 5.1 we present an example of the main characteristics of the data-set selected to compile some of the cited catalogues. For each of these samples we show only the properties useful for our discussion. Note that all these samples have mean number density $\langle n \rangle$ smaller than that of our N-body simulations with the smallest resolution (compare with Table 3.1). Furthermore, all of these samples are composed with tracers which are biased w.r.t. the DM density field. Since we found that a minimum spatial resolution of m.p.s.≈ 1 Mpc$/h$ is necessary to apply safely our method, it is clear that a generalization to the case of sparse and biased samples is necessary.

In this Chapter we discuss some possible modifications to our fiducial model of void size function, in order to apply it also to the case of biased tracers. In fact, it has been observed (Benson et al., 2003) that voids found in the galaxy distribution

Survey	Magnitude limit	Δz	$\langle n \rangle$ $(\mathrm{Mpc}/h)^{-3}$	m.p.s. Mpc/h
SDSS DR7	$M_r < -18.90$	$[0.00, 0.05]$	$2.4 \cdot 10^{-2}$	3.47
	$M_r < -20.40$	$[0.00, 0.10]$	$8.7 \cdot 10^{-3}$	4.86
	$M_r < -21.35$	$[0.00, 0.15]$	$2.5 \cdot 10^{-3}$	7.37
	$M_r < -22.05$	$[0.00, 0.20]$	$5.1 \cdot 10^{-4}$	12.52
	$M_g < -21.20$	$[0.16, 0.36]$	$9.4 \cdot 10^{-5}$	21.99
	$M_g < -21.80$	$[0.16, 0.44]$	$2.6 \cdot 10^{-5}$	33.76
SDSS-III DR11	-	$[0.15, 0.43]$	$\approx 2 \cdot 10^{-4}$	17.10
(BOSS)	-	$[0.43, 0.70]$	$1.5 \cdot 10^{-4}$	18.71

Table 5.1: Some useful properties of the volume-limited samples used to compile the currently available void catalogues. We show the magnitude limit for all the catalogues for which is available. We also show the limiting redshifts Δz. The average number density of galaxies $\langle n \rangle$ and the mean particle separation are showed in comoving units.

are systematically larger than DM voids. Moreover the size of voids shall increase with the total mass of the tracers (Furlanetto and Piran, 2006). We make use of a set of halo catalogues obtained by means of a FoF algorithm[1] applied to the N-body simulation with volume 1 $(\mathrm{Gpc}/h)^3$ (`Lhm` in Table 3.1).

The halo distribution mimics the distribution of biased tracers like galaxies obtained from real observations. Our analysis aims at filling the gap between what we can learn about voids from simulations and the actual voids identified in galaxy samples.

We apply the VIDE algorithm to four halo mock catalogues at redshift $z = 0$. These catalogues are obtained by cutting the original one at different minimum masses. In Table 5.2 we show the properties of the set of catalogues to discuss the role of bias. Since the tracers of the density field are in this case not only biased but also extremely sparse, as anticipated in Section 3.4.1, we will use the barycentre of the voids, $\mathbf{r}_V$, instead of the core particle centre, $\mathbf{r}_C$, as centre of the voids.

To get an idea of how the internal structure of voids traced by haloes is designed, in Section 5.1 we compare their stacked density profiles with the stacked profiles of the underlying DM density field. This result leads to two possible approaches which we examine in the following Section 5.2. With our analysis we find that the halo-void radii are connected to the DM-void radii and to the underlying DM density contrast by two different scaling relations (Section 5.3). By means of the last of these two relations we develop a semi-analytical model for a halo-VSF in Section 5.4.

[1] A Friend of Friend (FoF) algorithm defines haloes in a simulation by grouping all the DM particles which are separated by a distance smaller than $\beta \cdot$mps, with β a free paramater (frequently ≈ 0.2) and mps the mean particle separation. The resulting halo catalogue should only depend on mps.

catalogue	M_{min} $M_\odot/h$	b_{eff}	N_H	$\langle n \rangle$ $(\mathrm{Mpc}/h)^{-3}$	mps Mpc$/h$
`no-cut`	$2 \cdot 10^{12}$	1.197	1582064	$1.58 \cdot 10^{-3}$	8.58
`low-cut`	$5 \cdot 10^{12}$	1.386	761577	$7.62 \cdot 10^{-4}$	10.95
`mid-cut`	$1 \cdot 10^{13}$	1.583	394012	$3.94 \cdot 10^{-4}$	13.64
`high-cut`	$1 \cdot 10^{14}$	2.771	25615	$2.56 \cdot 10^{-5}$	33.92

Table 5.2: Characteristic quantities for our set of halo catalogues. We show the minimum halo mass M_{min}, the effective bias of the sample b_{eff}, calculated with an internal function of the CBL, the total number of haloes N_H, the mean numerical density $\langle n \rangle$ and the mean particle separation mps.

5.1 Stacked Profiles

It is generally accepted that, in the standard ΛCDM cosmology, voids have a *self-similar* internal structure, which do not depend on the tracer used to define them (see e.g. Ricciardelli et al., 2013, 2014; Hamaus et al., 2014).

The self-similarity assures that the internal structure of voids can be charachterized by a *single parameter*: the void effective radius R_{eff}. Despite the wide range of values this parameter covers, it is possible to average the internal density distribution to recover a common behaviour for voids with approximately the same R_{eff}.

It has been noticed (Nadathur et al., 2015) that, to recover self-similarity, some criteria have to be applied to the selection of objects to be classified as voids. Nevertheless the most crucial role seems to be played by the cut on the minimum density of the candidate, ρ_{min}, which has to be sufficiently low (i.e. $\rho_{min} \lesssim 0.3 \langle \rho \rangle$). Since such an underdensity degree is largely achieved by our default cleaning algorithm, we are confident that the objects we select should behave in a self-similar manner. Therefore, we apply our cleaning algorithm to the `no-cut` halo-void catalogue, obtaining a collection of approximately 7000 voids with VIDE-radii between ≈ 9 Mpc$/h$ and ≈ 63 Mpc$/h$.

To reconstruct the density field for both the DM particles and haloes, we have to implement an additional step in the code. In fact, the original `Lhm` simulation counts approximately two billions particles. Since, we want to mantain the maximum possible resolution of DM particles, we have to load all the particle positions of the simulation snapshot, which is extremely expensive in terms of required memory.

We overcome this issue by dividing the simulation box into a mosaic of smaller sub-boxes. Since voids are extended regions, even though their centres are found in a sub-box, their edges may trespass on its boundaries. To prevent this eventual incompleteness in the density mapping, we let the cubic tiles of the mosaic overlap by 100 Mpc$/h$. Voids which belong to the overlapping region between two or more sub-boxes are considered as belonging only to the sub-box where they are further from the edges.

By means of this procedure, we store in binary files the position of each particle and, separately, of each halo embedded in a spherical region of radius thrice the VIDE-radius of the parent halo void. For each DM-density profile we also find the

scale of the corresponding shell crossing radius, r_{eff} (i.e. the radius which embeds a density $\rho = 0.2 \cdot \langle\rho\rangle$). This is the true radius of the considered void, since it describes a spherical region gone through shell crossing.

To obtain the stacked profiles we divide our sample into four equally large groups of voids with radii between 17 Mpc/h (i.e. twice the m.p.s. of the halo sample) and 61 Mpc/h. Since our definition of voids is based on a spherical approximation, and on the requirement of shell-crossing, we compute the spherically-averaged density profiles of all the voids included in each of these bins. The number density is estimated within a series of logarithmically equispaced spherical shells centred in the barycentre of each void

$$\overline{n}_i = \frac{1}{N} \sum_{j=0}^{N} \frac{n_i^j}{\langle n \rangle} \ , \tag{5.1}$$

where $\overline{n}_i$ is the mean density of voids included in the i-th bin, N is the number of voids in the i-th bin, n_i^j is the density of the j-th void in the i-th bin and $\langle n \rangle$ is the mean number density of the considered simulation. This is done for both the DM particles and the haloes, expressing the tracers distance from the centre in terms of r_{eff}. The results are shown in Figure 5.1.

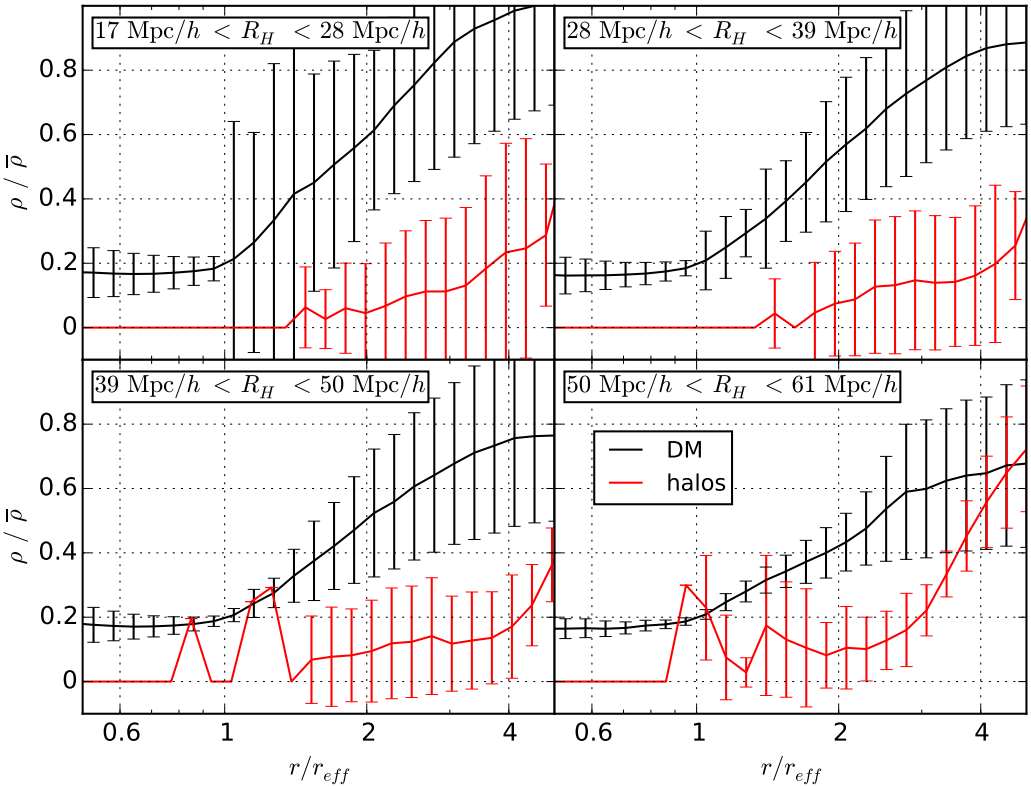

Figure 5.1: Stacked density profiles for the DM particles (black curves) and haloes (red curves). The x-axis scale is the distance from the centre in terms of the shell-crossing effective radius r_{eff}, while the y-axis is the density in terms of the mean density of tracers. In each of the four boxes we show the interval of radii of the averaged objects, in terms of the radius provided by the application of VIDE to the halo catalogue.

The error bars represent the scatter in the density profiles, whose general behaviour is to decrease approaching the shell-crossing effective radius in the DM particle voids. Even though this results from our definition of voids, the smaller the considered voids, the larger is the scatter in their profile immediately outside the effective radius r_{eff}. The fluctuations in the halo-void density profiles are caused by the low statistics in the inner region of voids, as a consequence of the larger sparsity of the tracers. This behaviour is more relevant in larger voids, which are less numerous than smaller voids.

As it is clearly shown in Figure 5.1, voids identified in the halo distribution are systematically more underdense than voids in the DM particle distribution in the inner regions. Nevertheless, voids in the halo distribution seem to generally encompass voids in the DM distribution. This result demonstrates that, if we want to predict the statistical distribution of voids with realistic tracers, it is necessary to adapt the theoretical model.

5.2 Towards a model for the VSF in the halo distribution

From the stacked profiles found in the previous Section we can conclude that halo-voids do not share the same radius of their DM particle-defined counterpart. Moreover, the void density profile traced in the halo distribution shows a density distribution that, not only reaches smaller values in the inner regions, but is also shifted to larger radii. Furthermore, the more we get closer to the void centre, the less are the tracers and thus the density goes to zero.

To predict the void distribution as a function of size in the distribution of realistic tracers, we have to re-elaborate the adopted formalism. To this end, we consider two possible modifications which rely on two different criteria based on the matter distribution within a void. Refering to the profiles in Figure 5.1, let us consider the position in the DM particle density profile where the shell-crossing threshold is reached, i.e.

$$P_r = (r/r_{eff}, \rho/\overline{\rho}) = (1, 0.2) \ ,$$

as our reference point. Ideally, we can either move downward along the y-axis from P_r until the halo density profile is reached, or we can move rightward along the y-axis. With the first approach we are trying to find in the halo distribution the boundaries of the DM void. Instead the second approach relies in the definition of a halo void enclosing the DM void which has generated it.

The first approach, which relies on the definition of a tracer-dependent threshold based on the formalism anticipated in Section 2.4, is exposed and commented in Section 5.2.1. The second approach will be discussed in Section 5.2.2, but it will also lead to the results presented in Sections 5.3 and 5.4.

5.2.1 The analytical tracer-dependent threshold

If we want to compare halo-voids to the predictions of the Vdn model we have to detect in the halo distribution the same voids we detect in the DM particle distribution. In fact, the Vdn model has been developed for the DM density field

where the underdensity threshold is fixed by the shell-crossing condition. Therefore, if we want to keep the value $\delta_v = -2.71$ fixed in the model, we have to find the corresponding value of the threshold in the case of biased tracers.

By means of the formalism exposed in Section 2.4, the threshold for haloes depends on the ratio between the halo distribution inside a void with a given average density w.r.t. the average halo distribution of the sample.

We recall eq. (2.63) here for convenience:

$$\frac{\rho_v}{\langle\rho\rangle} \equiv 1 + \delta_h^{NL}(m_{min}, \delta_v, R_v) = (1 + \delta_v^{NL}) \frac{n_h^c(m_{min}|\delta_v, R_v)}{n_h^c(m_{min})} \ , \tag{5.2}$$

where $n_h^c(m_{min})$ and $n_h^c(m_{min}|\delta_v, R_v)$, are the cumulative halo mass function and the conditional cumulative halo mass function, respectively.

Eq. (5.2) shows that a void with underdensity δ_v^{NL} in the DM distribution would have a scale-dependent underdensity $\delta_h^{NL} \propto R_v$ in the halo distribution. We can then calculate a coherent halo threshold for each catalogue in our set. The results of this measure are reported in Figure 5.2 and show that the required threshold, corresponding to the shell-crossing condition, gradually decreases when we cut the catalogue at larger masses. We only show non-zero values, as at a total mass $M_v(R_v)$

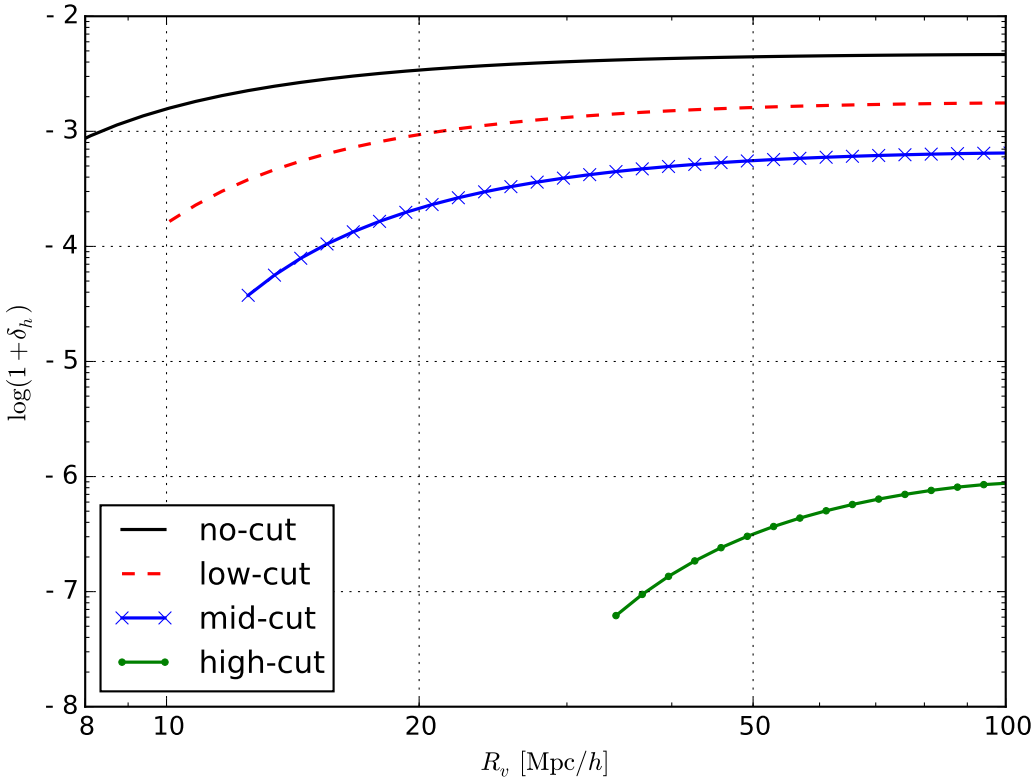

Figure 5.2: Density constrast $\rho_v/\langle\rho\rangle \equiv (1 + \delta_h)$ (eq. (5.2)) within halo-voids as a function of radius. All curves assume $\delta_v = -2.71$, and are referred to our halo catalogues, with the M_{min} shown in Table 5.2. Only the non-zero values are shown.

the function (5.2) is not defined (see Section 2.4).

A relevant feature to note is that the scale-dependent threshold tends asymptotically towards a costant value at large radii, while it decreases at small radii. For the `no-cut` haloes this value is approximately $5 \cdot 10^{-3}$, which justifies the downward shift of the density profiles in Figure 5.1.

Nonetheless, such small values mean that there is no hope of using this method to identify voids in the halo distribution and thus comparing the halo-voids statistical distribution with an unmodified Vdn model. In fact all the found values are at least two orders of magnitudes smaller than the shell-crossing threshold.

A simple example can better explain our point. Let us consider a large void, of 10 Mpc/h in radius, in the DM particle distribution. The `no-cut` catalogue is contained in 1 $(\mathrm{Gpc}/h)^3$ volume and has a mean number density of approximately $\langle n_{\texttt{no-cut}} \rangle \approx 0.0016\ (\mathrm{Mpc}/h)^{-3}$, meaning that the average expected underdensity within the considered void in the halo distribution is

$$\overline{n}_v(10\ \mathrm{Mpc}/h) = (1 + \delta_h(10\ \mathrm{Mpc}/h)) \cdot \langle n_{\texttt{no-cut}} \rangle \approx 10^{-6}\ (\mathrm{Mpc}/h)^{-3}\ .$$

Our spherical void has a volume of approximately $4 \cdot 10^3\ (\mathrm{Mpc}/h)^3$ so we do not have any chance of mapping its internal density in the halo distribution within such a small radius.

5.2.2 Shell-crossing fixed threshold

The alternative to the definition of a tracer-dependent threshold is to fix the underdensity threshold at some reasonably high value and search for the dependence between the halo-void properties and the DM underlying void.

We choose to define the halo-void radius as the point where the mean enclosed density is $\rho_{v,H} = 0.21 \cdot \overline{\rho}$, the same value found in eq. (2.31) and fixed by the shell-crossing condition. We apply our algorithm, obtaining a catalogue of spherical non-overlapping voids for each of the halo catalogues of Table 5.2. In Figure 5.3 we show the number counts of voids found by means of this procedure. All the points are showed, without applying any cut except that automatically applied by VIDE, which does not consider objects smaller than the m.p.s. of the tracers' distribution.

It is easy to appreciate how the total number of voids decreases as we search them in more biased tracers. The decreasing behaviour when the m.p.s. is approached is not physical but it is caused by the spatial resolution limit of the tracers used in the analysis.

Since the low statistics may cause issues in the following part of this work, we point out that in the `high-cut` void catalogue (magenta points in Figure 5.3) only a total of 148 voids are found by our algorithm. This, once again, highlights how crucial the resolution of the sample may be: voids traced by sparser samples are systematically larger and thus their number decreases considerably. To highlight the radial shift between the four distributions, in Figure 5.4 we have normalized the number counts by the total number of objects. All the distributions have a similar shape, even though the `high-cut` catalogue shows wide fluctuations.

The next step of our analysis is to obtain the stacked density profiles of voids in all our halo-distributions. For each considered void-catalogue, we divide the voids in four radial bins between the m.p.s. of the parent halo-catalogue and approximately the radius of the largest void found inside it.

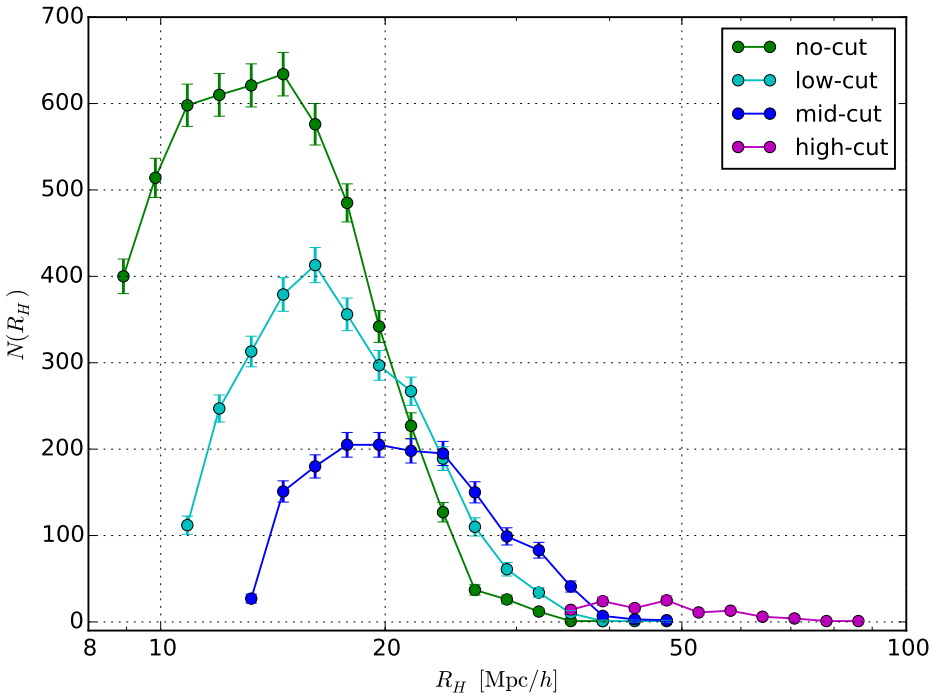

Figure 5.3: Total number of halo-voids found by our algorithm as a function of radius. The error bars show the Poisson error in the counts.

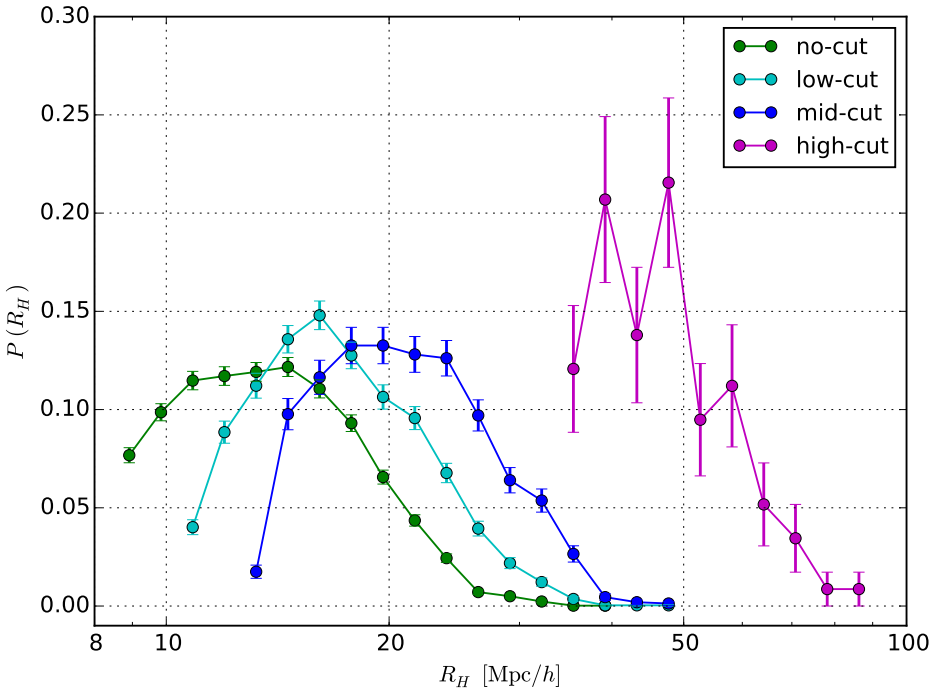

Figure 5.4: Same as Figure 5.3 normalized to the total number of voids in the cleaned VIDE catalogue.

The stacking procedure applied is the same described in Section 5.1 and the results are shown in Figures from 5.5 to 5.7 for `no-cut`, low-cut and `mid-cut` catalogues, respectively. We have superimposed the underlying DM-void density profiles (black curves) found in the DM particles distribution to the halo density profiles (red curves). To ease the comparison, both the profiles are shown in units of the respective effective radius. We make sure to radially divide each catalogue in order to include at least 50 objects in each bin.

The profiles obtained for the `high-cut` catalogue showed fluctuations of the order of the average density, which affects our capability to analyse them. This is an effect of the low number of voids which compose the catalogue. Since widening the radial bins is not an option (e.g. if the averaged void profiles span a wider range of radii the hypothesis of grouping "similar" objects fades) we decide from here after to not consider this catalogue anymore.

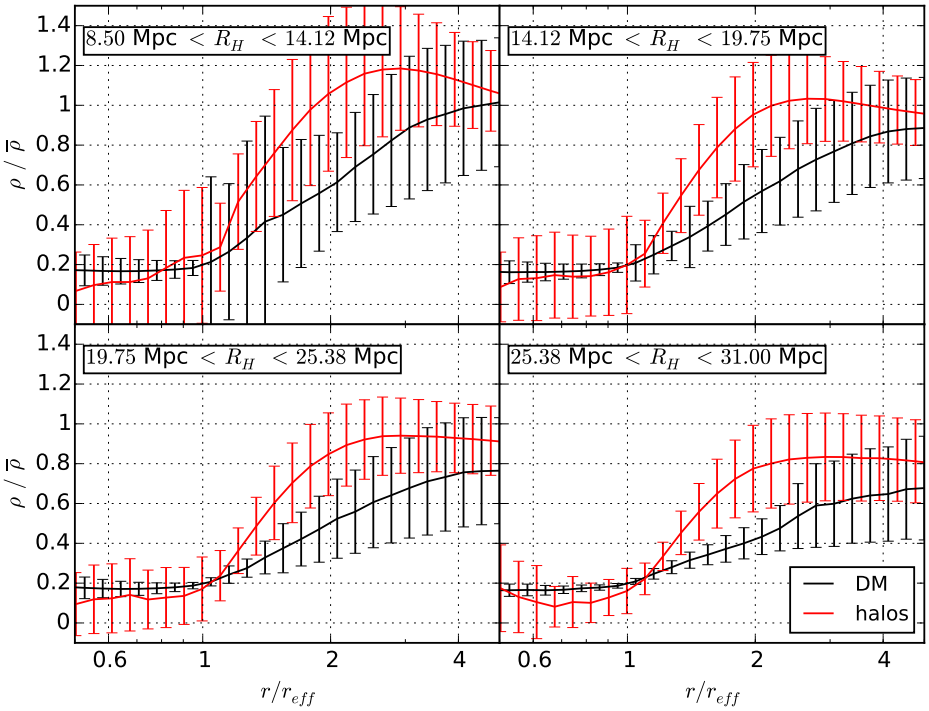

Figure 5.5: Stacked density profile for the DM particles (black curves) and haloes of the `no-cut` *catalogue (red curves). The scale is in units of the effective radius defined in the two different tracers distributions.*

We notice that in the inner regions, $r < r_{eff}$ the red profiles are similar or lower than the black profiles, while in the outer regions, $r > r_{eff}$, they grow faster than their counterpart. The red and black profiles systematically tend to rejoin at very large radii ($r > 3\, r_{eff}$), as expected.

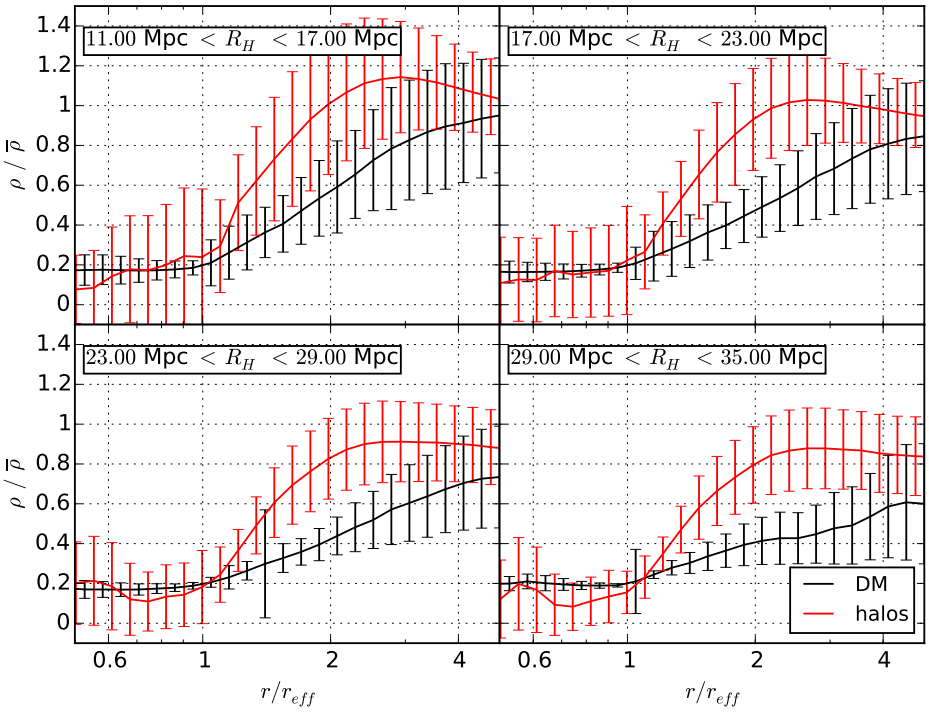

Figure 5.6: Same as Figure 5.5 for the `low-cut` *catalogue.*

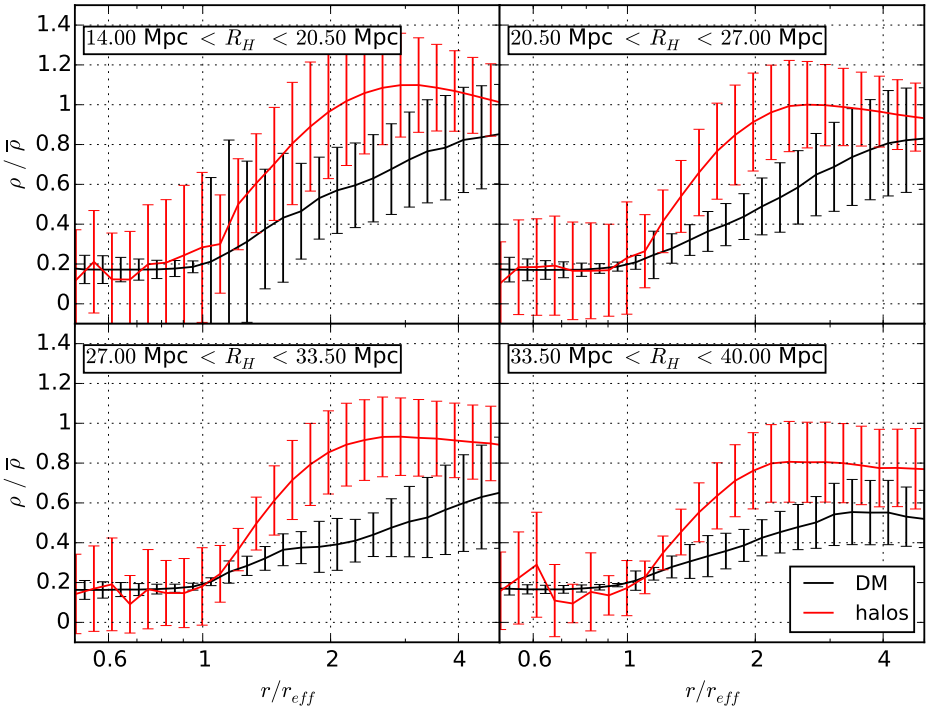

Figure 5.7: Same as Figure 5.5 for the `mid-cut` *catalogue.*

5.3 Scaling relations

We search for dependencies between the characterizing quantities of the found halo-voids and the underlying DM voids, in the set of catalogues built in the previous Section. As it has already been said, voids can be completely characterized by their radius, which is therefore the chosen quantity for the analysis we want to conduct.

We find two broad scaling relations which connect the halo-void radii found with the method described in Section 5.2.2 with two relevant properties of the DM voids in the DM density field: the DM void radii r_{eff} and the average DM density.

5.3.1 Halo-radius vs DM-radius

Having fixed the radius of each halo-void as the radius, R_H, of the sphere which includes a density contrast equal to that predicted by the shell-crossing condition, we find that a relation with the radius of the underlying DM void r_{eff} exists for each inspected halo catalogue.

This is a linear relation in the logarithmic values of the radii of the form

$$\log R_H = A + B \log R_{DM} \ , \tag{5.3}$$

with the error on the fit given by the root mean square of the residuals. Eq. (5.3) implies that the effective radii in the two tracers are connected by a power law.

In Figures from 5.8 to 5.10 these relations for the three remaining catalogues are shown. In each of these Figures the linear logarithmic relations are shown with the corresponding values of the coefficients and with the relative errors. The circles are the median values in each bin while the error bars show the positions of the first and third quartiles.

To obtain the fit we consider voids in the same range of radii used to plot the stacked profiles but we do not consider radial bins with less than 5 values.

As expected from our analysis of the stacked profiles, in all the considered catalogues the halo-void radii are larger than their counterpart in the DM particle distribution. Still, there is a clear trend which may justify further investigations in this direction.

The coefficients in these relations (summarized in Table 5.3) show no particular trend which may suggest a dependence on the effective bias of the halo catalogue used to define them. Nevertheless, this may be a result of the large scatter of the data.

Name	A	B	error
`no-cut`	-0.97	1.16	0.0087
`low-cut`	-0.45	0.62	0.0048
`mid-cut`	-0.56	0.63	0.0051

Table 5.3: Best fit values of the R_H-R_{DM} relation's coefficients (A and B in eq. (5.3)) found for the three different catalogues and the corresponding error obtained as the root mean square of the residuals.

The large scatter of the found relations could possibly be a result of a not perfect overlap between the centre of the voids traced by haloes and the underlying DM

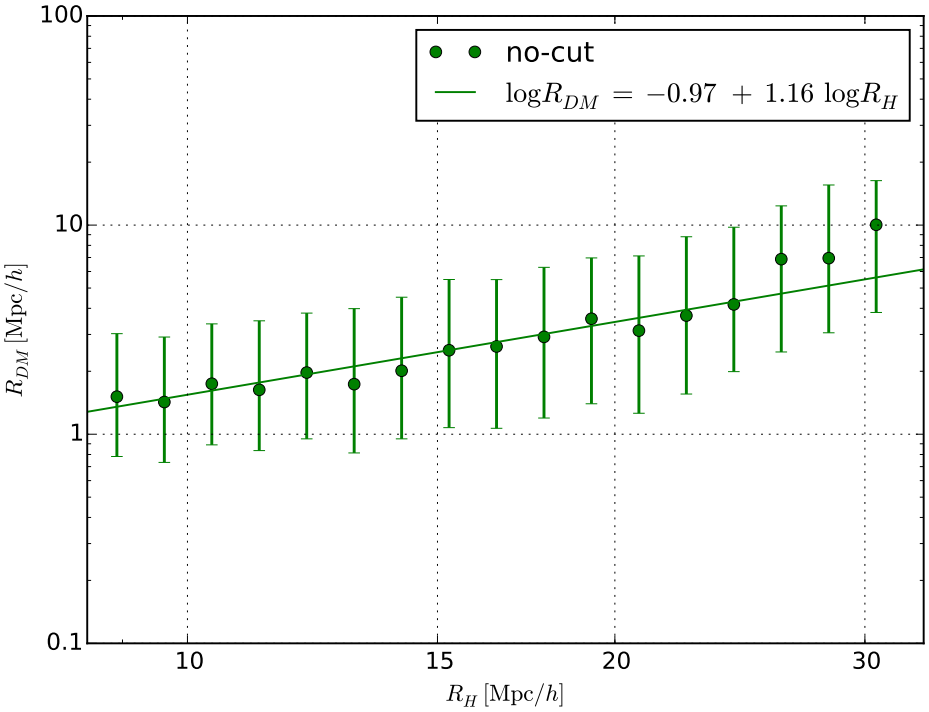

Figure 5.8: The relation between the effective radii obtained for voids in the halo distribution of the `no-cut` *catalogue (R_H) and the effective radii of the underlying DM voids in the DM distribution (R_{DM}). The points represent the median values while the error bars indicate the quartiles. We show the best-fit logarithmic linear relation (solid line) between our measurements.*

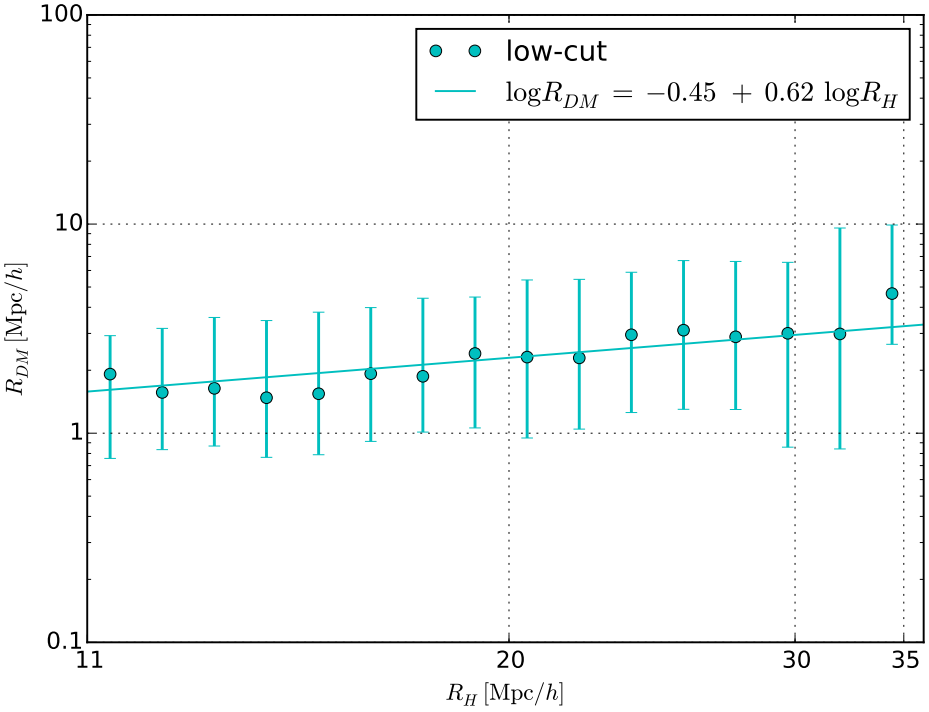

Figure 5.9: Same as Figure 5.8 for the `low-cut` *catalogue.*

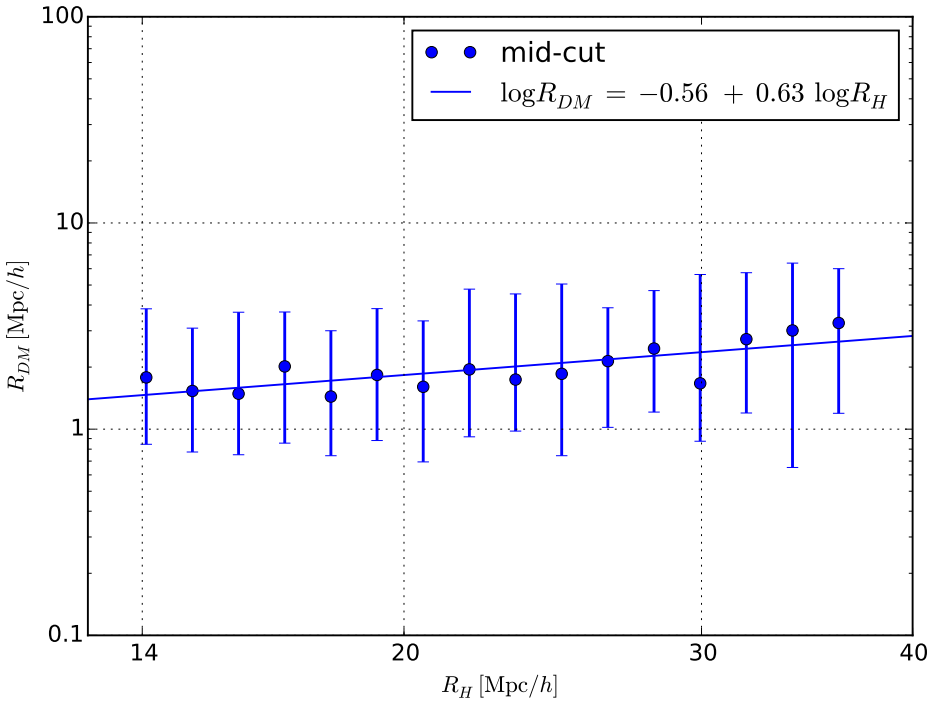

Figure 5.10: Same as Figure 5.8 for the `mid-cut` *catalogue.*

void centre. We leave the analysis of this eventuality for future expansion of our work.

5.3.2 Halo-radius VS DM-density

The second relation we find connects the halo-void radius R_H with the underlying un-biased DM density contrast. Once again the relation found is linear in the logarithmic values of the two quantities (see previous Section), resulting in a power-law dependence of the kind

$$\delta_{DM} = C + D \log R_H \ , \tag{5.4}$$

where the coefficients depend on the bias of the chosen catalogue.

We show the results in Figures from 5.11 to 5.13 for catalogues `no-cut`, `low-cut` and `mid-cut`, respectively.

The general trend is a decrease of the average encompassed density as the radius grows. This result indicates that larger voids in the halo distribution encompass deeper underdensities in the DM distribution.

We also note that the coefficients of these relations seem to vary in a self-consistent way (Table 5.4). While the intercept of the relation generally decreases with the effective bias of the sample, the slope is fixed at approximately -0.3. Given the extremely limited number of considered occurrencies we cannot confirm this behaviour and reserve further investigations for the future.

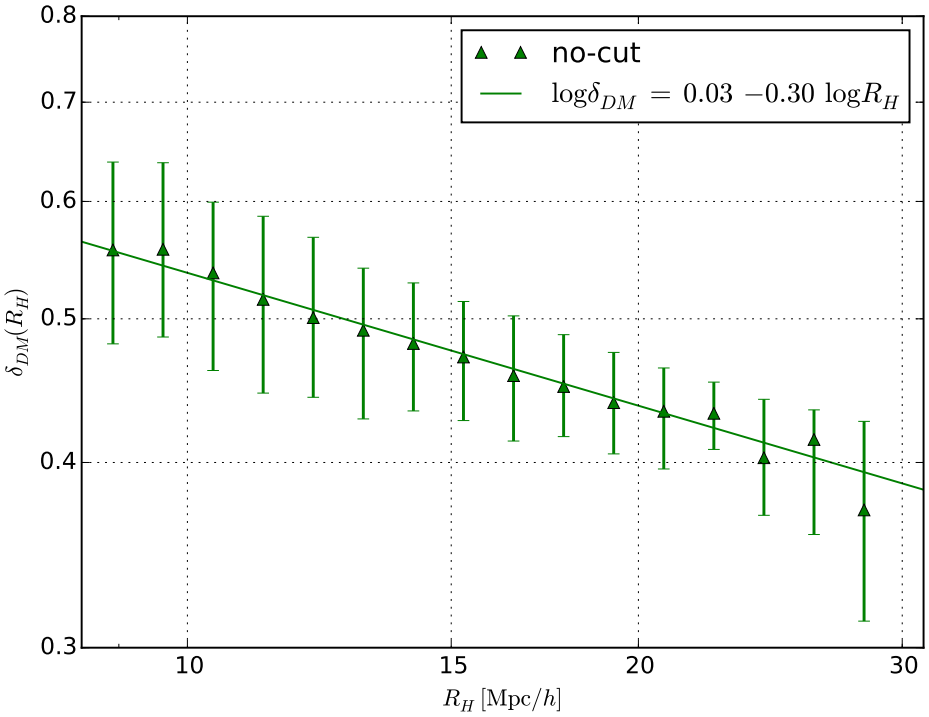

Figure 5.11: Relation between the effective radii obtained for voids in the halo distribution of the `no-cut` *catalogue (R_H) versus density contrast of the underlying DM distribution (δ_{DM}). The triangles represent the median values while the error bars indicate the quartiles. We show the best-fit relation (solid line) between our measures.*

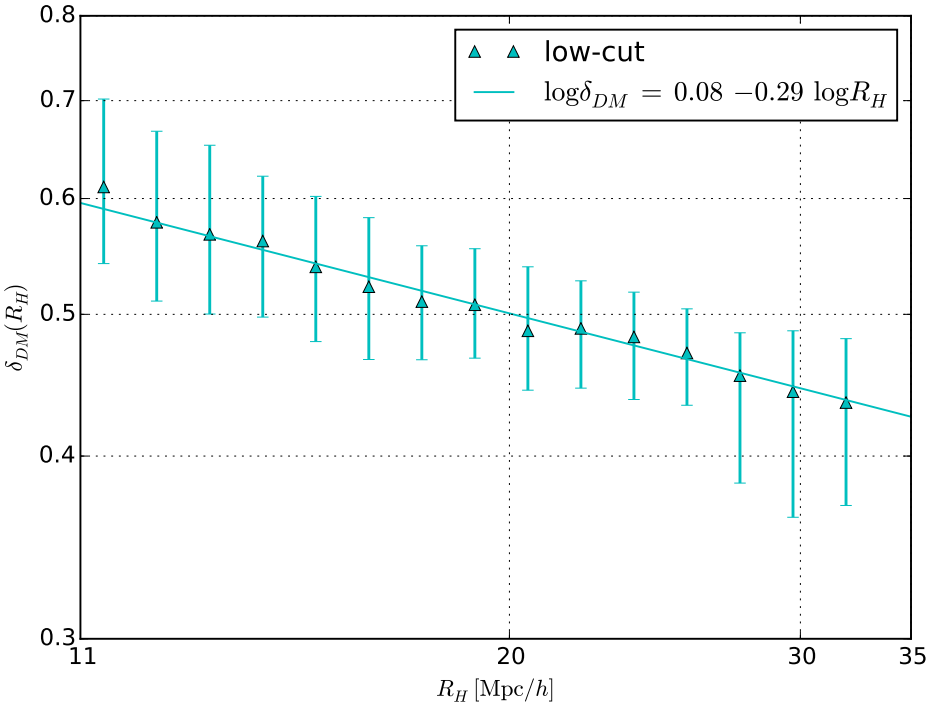

Figure 5.12: Same as Figure 5.11 for the `mid-cut` *catalogue.*

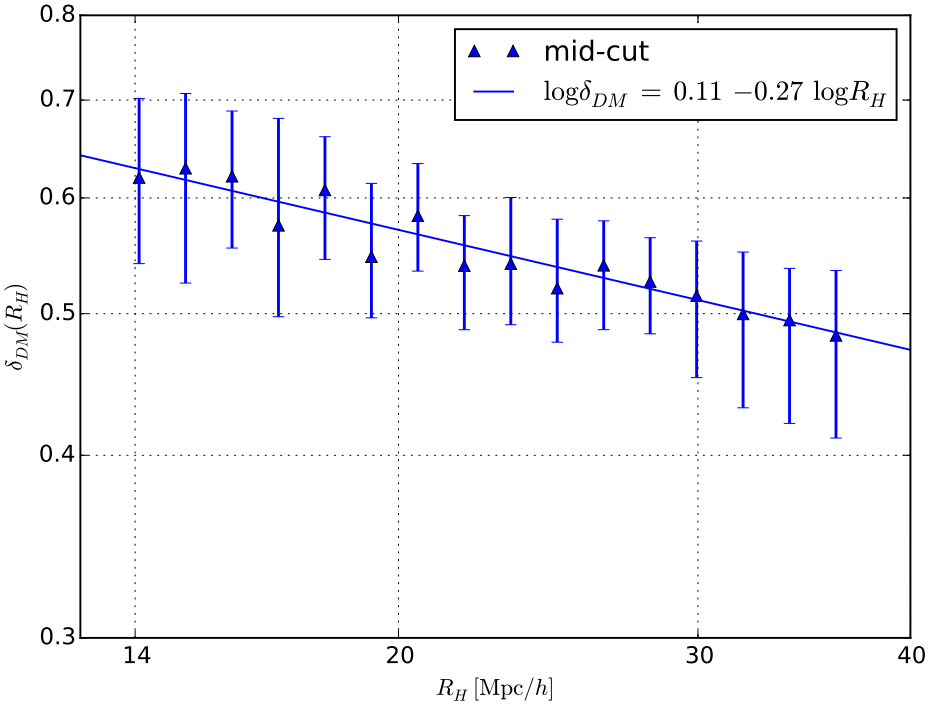

Figure 5.13: Same as Figure 5.11 for the `low-cut` *catalogue.*

Name	C	D	error
`no-cut`	0.03	-0.30	$8.847 \cdot 10^{-5}$
`low-cut`	0.08	-0.29	$3.983 \cdot 10^{-5}$
`mid-cut`	0.11	-0.27	$1.373 \cdot 10^{-4}$

Table 5.4: Best fit values of the R_H-δ_{DM} relation's coefficients (C and D in eq. (5.3)) and corresponding error in the fit (root mean square of the residuals).

5.4 Building a semi-analytical model for halo voids

The relation between R_H and δ_{DM} may be useful to define a semi-analytical model for the void size function. In fact, with the Vdn model we can predict the statistical distribution of voids in the DM density field but not in biased distributions. The void distributions of the three examined catalogues are extremely underpredicted by the Vdn model with a shell-crossing threshold of $\delta_v^{NL} = 0.21$.

We argue that, to recover the agreement between the measured distribution and the theoretical model some prescriptions have to be applied. We procede as follows:

1. From the scaling relation found in the previous Section (eq. (2.54)), we define the mean DM density contrast enclosed in each halo void with a given size as the halo scale-dependent underdensity threshold.

2. This varying threshold, found in the non-linear matter distribution, is linearly extrapolated, applying the fit found by Bernardeau (1994),

$$\delta_v \to \mathcal{C}\left[1 - (\delta_{DM}(R_H))^{-1/\mathcal{C}}\right] = {\delta_{DM}}^L(R_H) \ , \tag{5.5}$$

   where $\mathcal{C} = 1.594$ and L specifies a linearized value. Eq. (5.5) is proven to reproduce the linear evolution values of the density contrast with error bars within an error of 0.2% up to $\delta = -1$ (Bernardeau, 1994; Jennings et al., 2013).

3. We then replace the DM underdensity threshold δ_v in the void size function (eq. (2.54)), with the scale-dependent $\delta_{DM}^L(R_H)$ and then calculate the predictions of this modified Vdn model (red curve in Figures from 5.14 to 5.16).

4. We define a confidence range of values approximately around the new semi-analytical model (shaded regions in Figures from 5.14 to 5.16). This is achieved by repeating steps from 1 to 3 using point by point the fitted quartiles values, instead of the fitting values of the density contrast.

We can now compare the void distribution in the catalogues obtained in Section 5.2.2 (black dots in Figures from 5.14 to 5.16) with the void size function semi-analytical model obtained with the outlined procedure. In fact, since the void size function has been derived for fluctuations in the un-biased DM density field, using the (linearly extrapolated) DM density contrast of Section 5.3.2 is consistent with the theoretical model. Note that all the Figures shown are cut at around twice the m.p.s. of the halo distribution, at the small radii. Conversely, at large radii, we cut our results at the upper limit of the reliability interval used to define the scaling relation. All points included in this radial range are at least marginally included in the confidence range of values (defined in point 4.).

In conclusion, our analysis portrays an uncharted but viable approach to the statistical modelling of the halo void size distribution. Our results underline the existence of a non-trivial dependence between the tracers used to define voids, their sizes and the DM density contrast within them. The raised questions require further investigations, such as examining their characteristics at varying redshift and at varying cosmological parameters (with an analysis similar to that outlined in Section 4.4).

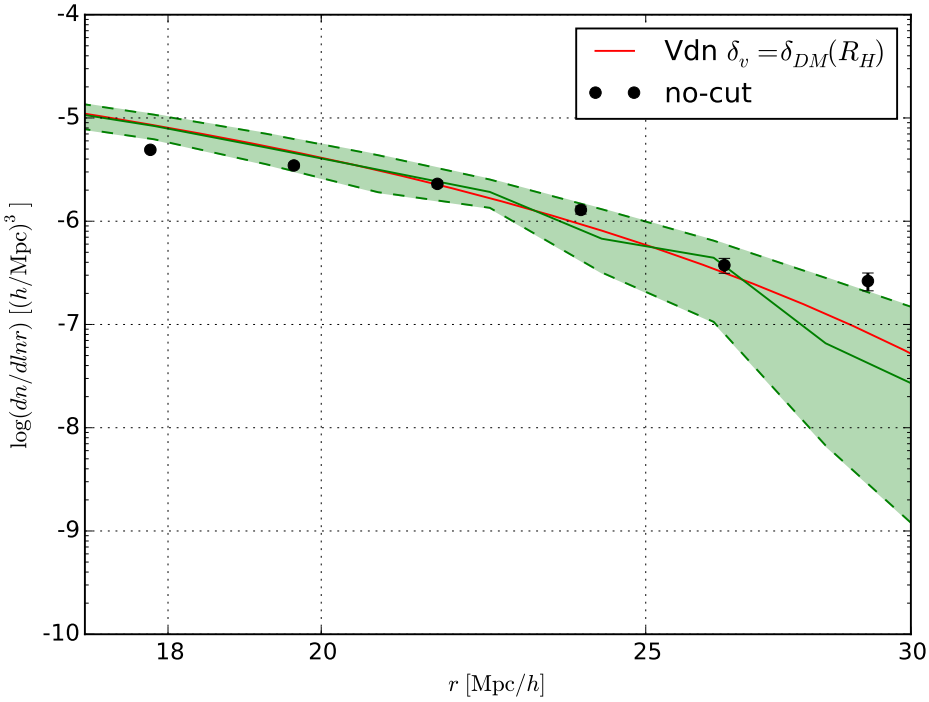

Figure 5.14: The size distribution of halo-voids in the **no-cut** *catalogue is shown (black dots) versus the semi-analytical model described in Section 5.4 (red curve). The coloured shaded region indicates the confidence range of the model, and is the defined as the value of the Vdn model with a density threshold given by the median (solid line) and the first and third quartiles (dashed lines) used to fit the corresponding linear relation.*

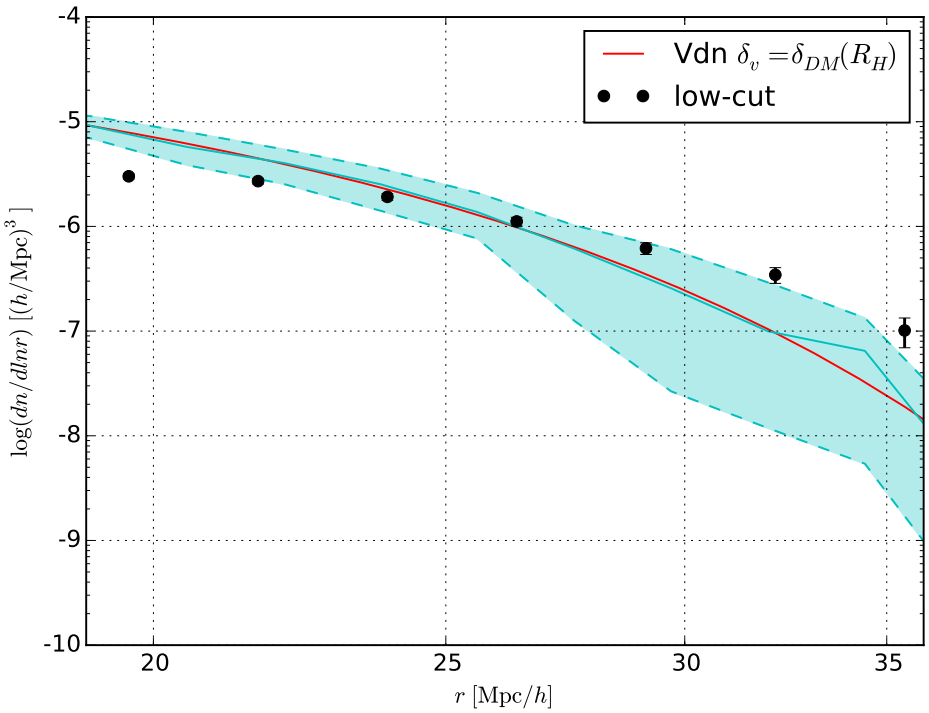

Figure 5.15: Same as Figure 5.14 for the `low-cut` *catalogue.*

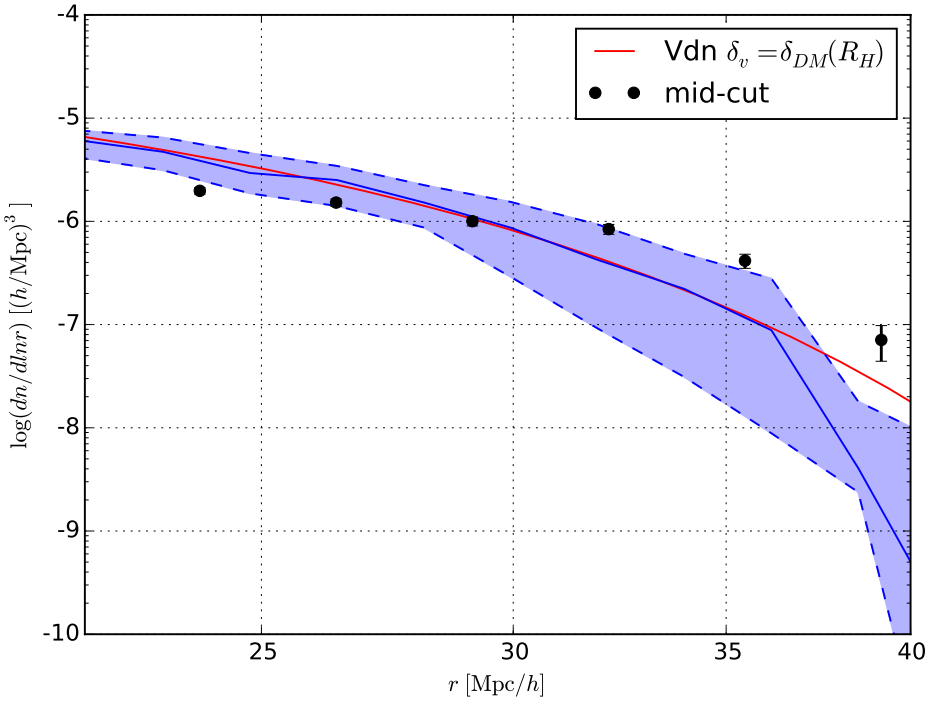

Figure 5.16: Same as Figure 5.14 for the `mid-cut` *catalogue.*

Chapter 6

Discussion and conclusions

In this Chapter we sum up the results gathered in our work, by highlighting the relevance of the outlined picture, and then making an insight into some possible future developments. We start the discussion recalling the cosmological context in which our work sits, recovering the main open questions of the modern cosmological framework. We then describe how cosmic voids could become, in the near future, an important probe of cosmological theories. We depict how the results obtained in this work could constitute an improvement of the already existing methods to inspect the voids statistics. Finally, we provide some possible improvements for the proposed line of work, with an eye on the possibilities that future surveys will provide.

The currently accepted *flat*-ΛCDM Standard Cosmological Model (exposed in Section 1.2) is based on the solution of the Einstein's field equations under the conditions of homogeneity and isotropy imposed by the assumption of the CP (Section 1.1). It accounts for the existence of two major energy components, CDM and DE, and for the flatness of the metric. This model is characterised by a set of parameters that describe its energy content, the primordial density fluctuation distribution which seeded the today's structures and the expansion rate of its overall scale.

Despite the theoretical framework of the Standard Cosmological Model is well outlined and generally accepted by the scientific community, there are still an amount of unresolved issues. They span from the accuracy of the inflation hypothesis, which at the first place determines the initial conditions of today's structure seeds, to the nature of the DM component. Moreover, the true cause of the accelerating expansion observed now is yet poorly understood. The introduction of a cosmological constant term Λ in the Einstein's equation, in fact, only accounts for this trend without giving any other information on its physical meaning.

The Standard Model is characterised by a set of observable cosmological parameters which are constrained by constantly improving experiments. The appearance of today's large scale structure depends on the initial conditions and on the metric and energy properties of the Universe. Thus, from the observable characteristics of these structures it should be possible to put constraints on the cosmological model's defining parameters. The measurements of the matter fluctuations, w.r.t. the prescripted homogeneity and isotropy, obtained from CMB anisotropies and

wide galaxy redshift surveys are capable of probing the longer wavelengths (low Fourier wavenumbers k) of the matter power spectrum (Section 1.3.3). The study of the largest structure statistics, namely clusters and voids, is then a promising probe of the mildly non-linear regime. The final goal is in fact to extract the underlying linear gravitational metric fluctuation field Φ which characterizes the initial conditions and the cosmic parameters that define its evolution.

Even though with the CMB probes of the density field it is possible to derive the cosmic parameters with spectacular accuracy (e.g. Planck Collaboration et al., 2015a), there are significant near-degeneracies among cosmological parameters such as Λ and the curvature, which require other experiments, such as wide galaxy surveys. In particular, voids emerge out of the density troughs on the primordial Gaussian field of density fluctuations, and given their large scales (i.e. larger than tens of Mpc), they are between the largest observable structure in the Universe, thus with the lowest values of wave number k in the Fourier space. It is then possible to use their statistics for a wide range of cosmological investigations. They have been shown to have great potential for constraining DE and testing theories of gravity via several measurements, which include geometrical tests, lensing measurements and the check on the forecasts of theories of gravity and cosmology (see the introduction of Chapter 2 and references therein).

Nevertheless voids are for their nature only mildly non-linear, since they are physically limited to reach at most a density contrast of $\delta = -1$. This suggests that their evolution should be easier to reconstruct than that of positive perturbations which can reach extremely non-linear amplitudes. The spherical evolution approximation (Section 2.1) should optimally describe the development of cosmic voids, since the natural tendency of underdensities is to get more spherical as they evolve (Icke, 1984). Cosmic voids evolve with an accelerated rate of expansion w.r.t. the Hubble flow, until they reach shell-crossing, and then reconnect with the general Universe expansion (Blumenthal et al., 1992).

Despite their recently-gathered popularity as cosmological tools, a gap of knowledge between the evolution of individual voids through simulations and observations versus theory still persists. In particular, the accepted statistical description of the void number density as a function of their size (SvdW model, Sheth and van de Weygaert, 2004) lacks of a complete validation.

The main goal of this thesis work was to cover this gap, both developing a coherent procedure to identify voids in simulated distribution, and reconsidering the accepted model definition by inspecting the possibilities offered by a recently proposed refinement (Vdn model, Jennings et al., 2013). The improvement of this model is the inclusion of the natural tendency of voids to overlap and merge. This is achieved by requiring that, in the transition from the linear theory prediction to the non-linear actual situation, the total volume occupied by voids is conserved, rather than their total number. Since their results are the only one promising in terms of the possibilities they offer to actually attain cosmological constraints from the size distribution of voids, we define the Vdn model as our fiducial model. Both the Vdn model and the SvdW model depend on cosmological parameters and cosmic time.

As a first step we developed and implemented an alternative to the existing

public void finders (Section 3.4). We are confident that our algorithm better locates objects which can be called "voids", by means of the cited theoretical definition which is based on the requirement of shell-crossing. The developed code makes large use of the CBL library (Section 3.1) and is intended as a future expansion of these public libraries[1]. Even though it can be upgraded, we find the void centre definition of the existing public void finders an acceptable approximation and we use this information to locate the central position of candidate underdensities in DM N-body simulations. We then search for the largest spherical region embedding a mean density contrast of 0.2 $\overline{\rho}$, with $\overline{\rho}$ the mean density of the considered simulation, and we finally make sure the found regions do not overlap. The condition on the mean density ensures that the selected underdensities have actually passed through shell crossing while the second condition guarantees the volume conservation.

Having established a procedure capable of building a reliable void catalogue, we want to test if, when compared to models, the theoretical predictions are accurate. We have used a set of N-body simulations with different characteristics in terms of cosmological model, box size and resolution. For each of these simulations we gathered the snapshots at four different cosmic times. In Section 4.1 we demostrate that the Vdn model effectively predicts the distribution of cosmic voids in the DM density field in an accurate way and is almost equally reliable for all the cosmological models considered. We think that this result may help to better understand the possibilities and limits offered by an excursion set model in predicting the size function of cosmic voids. Even though refinements of the void finder are desirable, doubts on the actual effectiveness of the existing model have been removed. Furthermore, with our analysis we have shown that also the redshift dependency of the Vdn model is trustworthy, even though these results are preliminary, we have found an acceptable agreement with the model for all the redshifts and for all the cosmological models considered.

Prompted by these results, we have fixed the fiducial cosmological model and selected a set of simulations with varying resolutions and sizes as an additional test. Applying our algorithm to this new set of data enlarges the scales over which the model can be tested and validated. We find that the effectiveness of the model in predicting the number of voids contained in a given volume fades when the resolution of the catalogue is decreased. Nevertheless, as long as the tracers sample has a resolution of at least 1 Mpc/h the counts of voids found are correctly predicted by the Vdn model.

To complete our analysis, we have studied the effect of a variation in the defining cosmological parameters in the Vdn model, with the purpose of roughly determining if the model is actually sensitive to such a variation. Moreover, this gives us a hint on what are the scales at which the model is more sensitive to a variation of parameters and thus, provide prescriptions on its possible future applications to real data. The general conclusion is that the most relevant change is caused by the total mass contained in the Universe (parameterised by Ω_M). Nonetheless also a variation in the power spectrum normalization seems to affect in a significant way the predicted abundance of large voids.

Since the largest scales are preferentially interested by these variations in the

[1] A part of the code has already been integrated in the latest public version of CosmoBolognaLib.

cosmological parameters, and considered the limitations which the total volume covered by a certain sample puts on the precision of the statistics of large voids, it is necessary to recover data from large surveys to obtain reliable constraints on cosmology by the comparison with the studied model.

We believe that any more accurate consideration on the actual possibilities of the model would be pointless without its validation in distributions of realistic matter tracers. All our considerations have been made addressing only unbiased tracers but, with the goal of obtaining a new reliable cosmological probe from the void size function, validating its reliability on real matter tracers is a necessary development. Cosmological surveys do in fact provide information on the observable galaxies in the Universe, which populate the DM collapsed haloes. Both haloes and galaxies are biased tracers of the underlying DM density field (Section 1.3.4). As a testing ground for real galaxy redshift surveys, we study the void statistics in a halo mock catalogue built by means of a FoF algorithm from a large simulation with side length of 1 Gpc/h (Chapter 5).

Galaxy voids are predicted to present significant differencies from DM voids (Furlanetto and Piran, 2006), taking into account biased tracers is then a crucial step in the study of the statistical distribution of voids. The approaches explored are two:

- The threshold to be used in our algorithm (Section 3.4) is fixed at the value required by shell crossing for DM voids. We have considered a possible scale-dependent modification of this value when dealing with biased tracers.
- We have conversely fixed the threshold at the shell-crossing value and researched for the scaling relation between the radius of the halo voids found by our algorithm and their embedded DM density contrast.

In the first of these approaches, we tested the predictions of the analytical scale dependent threshold suggested by Furlanetto and Piran (2006). They propose a threshold δ_H which indirectly depends on the bias of the tracers and varies with radius. Having obtained the void centres for a set of halo catalogues with different levels of bias and then outlined the density profile in both the biased and the unbiased tracers (haloes and DM particles respectively), we compared the degree of underdensity in the inner regions with the theoretical scale-dependent threshold δ_H. We conclude that the derived extremely low values of δ_H together with the low resolution of the halo mock catalogues, especially in the inner regions of voids, do not permit to use this approach.

Moving to the second possible approach, we fixed the threshold used to define voids at the shell crossing level ($\delta_v^{NL} = 0.21$) independently of the used tracers. We applied our void finder to build a set of spherical, non-overlapped halo void catalogues for different values of the bias parameter. We then map the DM density field within each void. This procedure allowed us to obtain two relevant pieces of information about halo voids, DM voids and the underlying DM density field:

- there is an existing scaling relation between the radius of the halo-void and that of the underlying DM void;
- the density contrast of DM scales with the radius of the halo void in which it is embedded.

The first of these findings allow us to confirm that halo voids do result from the expansion of a shell crossing DM void.

We use the second result to define a semi-analytical model which predicts the halo void size distribution. The Vdn model depends on a parameter which quantifies the underdensity degree, δ_v. Since there is no particular reason for the reliability of the model to depend on the chosen threshold, we use the found relation to build a model for halo voids depending on the two fitting parameters of the relation. DM voids are fitted by a Vdn model with a fixed value of δ_v, conversely our model for halo voids has a varying threshold.

Comparing the defined semi-analytical model with the halo void catalogues previously obtained, we find that their void distribution is well reproduced. This test is reliable for each of the three catalogues we have analysed. Therefore we conclude that the probed procedure is actually valid, even though it requires refinements.

Similar scaling relations are present in each of the catalogues inspected, but the values of their defining parameters seem not universal. We think that the scaling relations may depend by the effective bias of the sample, but larger sets of simulations are necessary to confirm this hypothesis.

Our current understanding of the found scaling relations is not accurate enough to employ them in a real situation. We therefore intend to refine the data and the deriving relations to improve our comprehension of their occurence. These are the possible future developments of our work:

- to verify if it is possible to overcome the resolution issues outlined in Chapter 4 at both large and small scales. This would require larger simulations with higher resolution and to consider a better treatment of the periodicity and boundaries crossing of voids.

- We also hope to better mark the position of both DM void and halo void centres, possibly considering their respective position in the same simulation. We also plan to insert the possibility of a re-definition of the void centre in order to better measure its position in particularly sparse samples, accounting for the dinamical properties of voids. This is particularly desirable for biased sparse tracers.

- We think that the discovery of the existence of the scaling relations found in this work is a relevant improvement in our understanding of the statistical distribution of voids in biased samples. Thus, their dependence on bias should be investigated. They have to be surely tested at varying redshift and for different cosmological models. We are curious to inspect their eventual dependence on the cosmological model at fixed bias of the tracers.

- Last but not least, a theoretical explanation of these scaling relations is desirable.

In conclusion, all our results demonstrate that the voids size distribution could in the future become a valid cosmological probe.

Bibliography

I. Achitouv, M. Neyrinck, and A. Paranjape. Testing spherical evolution for modelling void abundances. *Monthly Notices of the RAS*, 451:3964–3974, August 2015. doi: 10.1093/mnras/stv1228.

S. Aiola, A. Kosowsky, and B. Wang. Gaussian approximation of peak values in the integrated Sachs-Wolfe effect. *Physical Review D*, 91(4), February 2015.

S. Alam, F. D. Albareti, C. Allende Prieto, F. Anders, S. F. Anderson, T. Anderton, B. H. Andrews, E. Armengaud, É. Aubourg, S. Bailey, and et al. The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III. *Astrophysical Journal*, 219:12, July 2015.

J. S. Bagla and T. Padmanabhan. Cosmological N-body simulations. *Pramana*, 49: 161, August 1997.

M. Baldi. The CoDECS project: a publicly available suite of cosmological N-body simulations for interacting dark energy models. *Monthly Notices of the RAS*, 422: 1028–1044, May 2012a. doi: 10.1111/j.1365-2966.2012.20675.x.

M. Baldi. Dark Energy simulations. *Physics of the Dark Universe*, 1:162–193, November 2012b.

A. Barreira, M. Cautun, B. Li, C. M. Baugh, and S. Pascoli. Weak lensing by voids in modified lensing potentials. *Journal of Cosmology and Astroparticle Physics*, 8, August 2015.

A. J. Benson, F. Hoyle, F. Torres, and M. S. Vogeley. Galaxy voids in cold dark matter universes. *Monthly Notices of the RAS*, 340:160–174, March 2003.

F. Bernardeau. The nonlinear evolution of rare events. *Astrophysical Journal*, 427: 51–71, May 1994. doi: 10.1086/174121.

G. Bertone, D. Hooper, and J. Silk. Particle dark matter: evidence, candidates and constraints. *Physics Reports*, 405:279–390, January 2005.

G. R. Blumenthal, L. N. da Costa, D. S. Goldwirth, M. Lecar, and T. Piran. The largest possible voids. *Astrophysical Journal*, 388:234–241, April 1992.

J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser. Excursion set mass functions for hierarchical Gaussian fluctuations. *Astrophysical Journal*, 379:440–460, October 1991.

J. R. Bond, L. Kofman, D. Pogosyan, and J. Wadsley. Theoretical Tools for Large Scale Structure. *ArXiv Astrophysics e-prints*, October 1998.

E. G. P. Bos, R. van de Weygaert, K. Dolag, and V. Pettorino. The darkness that shaped the void: dark energy and cosmic voids. *Monthly Notices of the RAS*, 426:440–461, October 2012.

Y.-C. Cai, M. C. Neyrinck, I. Szapudi, S. Cole, and C. S. Frenk. A Possible Cold Imprint of Voids on the Microwave Background Radiation. *Astrophysical Journal*, 786, May 2014.

Y.-C. Cai, N. Padilla, and B. Li. Testing gravity using cosmic voids. *Monthly Notices of the RAS*, 451, July 2015.

L. Ceccarelli, D. Paz, M. Lares, N. Padilla, and D. G. Lambas. Clues on void evolution - I. Large-scale galaxy distributions around voids. *Monthly Notices of the RAS*, 434:1435–1442, September 2013.

S. Chandrasekhar. Stochastic problems in physics and astronomy. *Rev. Mod. Phys.*, 15:1–89, Jan 1943. URL http://link.aps.org/doi/10.1103/RevModPhys.15.1.

J. Clampitt and B. Jain. Lensing measurements of the mass distribution in SDSS voids. *Monthly Notices of the RAS*, 454:3357–3365, December 2015.

Joseph Clampitt, Yan-Chuan Cai, and Baojiu Li. Voids in modified gravity: excursion set predictions. *Monthly Notices of the Royal Astronomical Society*, 431: 749–766, 5 2013.

J. M. Colberg, R. K. Sheth, A. Diaferio, L. Gao, and N. Yoshida. Voids in a ΛCDM universe. *Monthly Notices of the RAS*, 360:216–226, June 2005.

J. M. Colberg, F. Pearce, C. Foster, E. Platen, R. Brunino, M. Neyrinck, S. Basilakos, A. Fairall, H. Feldman, S. Gottlöber, O. Hahn, F. Hoyle, V. Müller, L. Nelson, M. Plionis, C. Porciani, S. Shandarin, M. S. Vogeley, and R. van de Weygaert. The Aspen-Amsterdam void finder comparison project. *Monthly Notices of the RAS*, 387:933–944, June 2008.

P. Coles and F. Lucchin. *Cosmology: The Origin and Evolution of Cosmic Structure, Second Edition*. Wiley, July 2002.

M. Colless, B. A. Peterson, C. Jackson, J. A. Peacock, S. Cole, P. Norberg, I. K. Baldry, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. Cannon, C. Collins, W. Couch, N. Cross, G. Dalton, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, W. Sutherland, and K. Taylor. The 2dF Galaxy Redshift Survey: Final Data Release. *ArXiv Astrophysics e-prints*, June 2003a.

M. Colless, B. A. Peterson, C. Jackson, J. A. Peacock, S. Cole, P. Norberg, I. K. Baldry, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. Cannon, C. Collins, W. Couch, N. Cross, G. Dalton, R. De Propris, S. P. Driver, G. Efstathiou, R. S.

Ellis, C. S. Frenk, K. Glazebrook, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, W. Sutherland, and K. Taylor. The 2dF Galaxy Redshift Survey: Final Data Release. *ArXiv Astrophysics e-prints*, June 2003b.

V. de Lapparent, M. J. Geller, and J. P. Huchra. A slice of the universe. *Astrophysical Journal*, 302:L1–L5, March 1986.

G. Efstathiou, M. Davis, S. D. M. White, and C. S. Frenk. Numerical techniques for large cosmological N-body simulations. *Astrophysical Journal*, 57:241–260, February 1985.

A. Einstein. Zur allgemeinen Relativitätstheorie. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 778-786.*, 1915.

A. Einstein. Kosmologische Betrachtungen zur allgemeinen Relativitäts theorie. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 142-152.*, 1917.

Hagai El-Ad and Tsvi Piran. Voids in the large-scale structure. *The Astrophysical Journal*, 491(2):421, 1997. URL http://stacks.iop.org/0004-637X/491/i=2/a=421.

A. Elyiv, F. Marulli, G. Pollina, M. Baldi, E. Branchini, A. Cimatti, and L. Moscardini. Cosmic voids detection without density measurements. *Monthly Notices of the RAS*, 448:642–653, March 2015.

S. Flender, S. Hotchkiss, and S. Nadathur. The stacked ISW signal of rare superstructures in ΛCDM. *Journal of Cosmology and Astroparticle Physics*, 2, February 2013.

A. Friedmann. Über die Krümmung des Raumes. *Zeitschrift fur Physik*, 10:377–386, 1922.

S. R. Furlanetto and T. Piran. The evidence of absence: galaxy voids in the excursion set formalism. *Monthly Notices of the RAS*, 366:467–479, February 2006.

B. R. Granett, M. C. Neyrinck, and I. Szapudi. An Imprint of Superstructures on the Microwave Background due to the Integrated Sachs-Wolfe Effect. *Astrophysical Journal, Letters*, 683, August 2008.

D. Gruen, O. Friedrich, A. Amara, D. Bacon, and ... Bonnett. Weak lensing by galaxy troughs in des science verification data. *Monthly Notices of the Royal Astronomical Society*, 455(3):3367–3380, 2016. URL http://mnras.oxfordjournals.org/content/455/3/3367.abstract.

J. E. Gunn and J. R. Gott, III. On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. *Astrophysical Journal*, 176:1, August 1972.

L. Guzzo and The Vipers Team. VIPERS: An Unprecedented View of Galaxies and Large-scale Structure Halfway Back in the Life of the Universe. *The Messenger*, 151:41–46, March 2013.

Oliver Hahn, Cristiano Porciani, C. Marcella Carollo, and Avishai Dekel. Properties of dark matter haloes in clusters, filaments, sheets and voids. *Monthly Notices of the Royal Astronomical Society*, 375(2):489–499, 2007. URL http://mnras.oxfordjournals.org/content/375/2/489.abstract.

N. Hamaus, P. M. Sutter, G. Lavaux, and B. D. Wandelt. Probing cosmology and gravity with redshift-space distortions around voids. *Journal of Cosmology and Astroparticle Physics*, 11, November 2015.

Nico Hamaus, P. M. Sutter, and Benjamin D. Wandelt. Universal density profile for cosmic voids. *Phys. Rev. Lett.*, 112:251302, Jun 2014. URL http://link.aps.org/doi/10.1103/PhysRevLett.112.251302.

Fiona Hoyle and Michael S. Vogeley. Voids in the point source catalogue survey and the updated zwicky catalog. *The Astrophysical Journal*, 566(2):641, 2002. URL http://stacks.iop.org/0004-637X/566/i=2/a=641.

E. Hubble. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15:168–173, March 1929.

V. Icke. Voids and filaments. *Monthly Notices of the RAS*, 206:1P–3P, January 1984.

S. Ilić, M. Langer, and M. Douspis. Detecting the integrated Sachs-Wolfe effect with stacked voids. *Astronomy and Astrophysics*, 556, August 2013.

E. Jennings, Y. Li, and W. Hu. The abundance of voids and the excursion set formalism. *Monthly Notices of the RAS*, 434:2167–2181, September 2013.

F.-S. Kitaura, C.-H. Chuang, Y. Liang, C. Zhao, C. Tao, S. Rodríguez-Torres, D. J. Eisenstein, H. Gil-Marín, J.-P. Kneib, C. McBride, W. J. Percival, A. J. Ross, A. G. Sánchez, J. Tinker, R. Tojeiro, M. Vargas-Magana, and G.-B. Zhao. Signatures of the Primordial Universe from Its Emptiness: Measurement of Baryon Acoustic Oscillations from Minima of the Density Field. *Physical Review Letters*, 116(17), April 2016.

E. Krause, T.-C. Chang, O. Doré, and K. Umetsu. The Weight of Emptiness: The Gravitational Lensing Signal of Stacked Voids. *Astrophysical Journalletters*, 762: L20, January 2013.

M. Kuhlen, M. Vogelsberger, and R. Angulo. Numerical simulations of the dark universe: State of the art and the next decade. *Physics of the Dark Universe*, 1: 50–93, November 2012. doi: 10.1016/j.dark.2012.10.002.

C. Lacey and S. Cole. Merger rates in hierarchical models of galaxy formation. *Monthly Notices of the RAS*, 262:627–649, June 1993.

C. Lacey and S. Cole. Merger Rates in Hierarchical Models of Galaxy Formation - Part Two - Comparison with N-Body Simulations. *Monthly Notices of the RAS*, 271:676, December 1994.

T. Y. Lam, J. Clampitt, Y.-C. Cai, and B. Li. Voids in modified gravity reloaded: Eulerian void assignment. *Monthly Notices of the RAS*, 450, July 2015.

G. Lavaux and B. D. Wandelt. Precision cosmology with voids: definition, methods, dynamics. *Monthly Notices of the RAS*, 403:1392–1408, April 2010.

G. Lavaux and B. D. Wandelt. Precision Cosmography with Stacked Voids. *Astrophysical Journal*, 754:109, August 2012.

Guilhem Lavaux and Benjamin D. Wandelt. Precision cosmology with voids: definition, methods, dynamics. *Monthly Notices of the Royal Astronomical Society*, 403 (3):1392–1408, 2010. URL http://mnras.oxfordjournals.org/content/403/3/1392.abstract.

B. Li, G.-B. Zhao, and K. Koyama. Haloes and voids in f(R) gravity. *Monthly Notices of the RAS*, 421, April 2012.

C. C. Lin, L. Mestel, and F. H. Shu. The Gravitational Collapse of a Uniform Spheroid. *Astrophysical Journal*, 142:1431, November 1965.

S. J. Maddox, G. Efstathiou, W. J. Sutherland, and J. Loveday. Galaxy correlations on large scales. *Monthly Notices of the RAS*, 242:43P–47P, January 1990.

F. Marulli, A. Veropalumbo, and M. Moresco. CosmoBolognaLib: C++ libraries for cosmological calculations. *Astronomy and Computing*, 14:35–42, January 2016.

E. Massara, F. Villaescusa-Navarro, M. Viel, and P. M. Sutter. Voids in massive neutrino cosmologies. *Journal of Cosmology and Astroparticle Physics*, 11:018, November 2015. doi: 10.1088/1475-7516/2015/11/018.

Peter Melchior, P. M. Sutter, Erin S. Sheldon, Elisabeth Krause, and Benjamin D. Wandelt. First measurement of gravitational lensing by cosmic voids in sdss. *Monthly Notices of the Royal Astronomical Society*, 440(4):2922–2927, 2014. URL http://mnras.oxfordjournals.org/content/440/4/2922.abstract.

D. Micheletti, A. Iovino, A. J. Hawken, B. R. Granett, M. Bolzonella, A. Cappi, L. Guzzo, U. Abbas, C. Adami, S. Arnouts, J. Bel, D. Bottini, E. Branchini, J. Coupon, O. Cucciati, I. Davidzon, G. De Lucia, S. de la Torre, A. Fritz, P. Franzetti, M. Fumana, B. Garilli, O. Ilbert, J. Krywult, V. Le Brun, O. Le Fèvre, D. Maccagni, K. Małek, F. Marulli, H. J. McCracken, M. Polletta, A. Pollo, C. Schimd, M. Scodeggio, L. A. M. Tasca, R. Tojeiro, D. Vergani, A. Zanichelli, A. Burden, C. Di Porto, A. Marchetti, C. Marinoni, Y. Mellier, T. Moutard, L. Moscardini, R. C. Nichol, J. A. Peacock, W. J. Percival, and G. Zamorani. The VIMOS Public Extragalactic Redshift Survey. Searching for cosmic voids. *Astronomy and Astrophysics*, 570:A106, October 2014. doi: 10.1051/0004-6361/201424107.

S. Nadathur. Testing cosmology with a catalogue of voids in the BOSS galaxy surveys. *Monthly Notices of the RAS*, June 2016.

S. Nadathur and S. Hotchkiss. A robust public catalogue of voids and superclusters in the SDSS Data Release 7 galaxy surveys. *Monthly Notices of the RAS*, 440: 1248–1262, May 2014.

S. Nadathur and S. Hotchkiss. The nature of voids - I. Watershed void finders and their connection with theoretical models. *Monthly Notices of the RAS*, 454: 2228–2241, December 2015.

S. Nadathur, S. Hotchkiss, and S. Sarkar. The integrated Sachs-Wolfe imprint of cosmic superstructures: a problem for ΛCDM. *Journal of Cosmology and Astroparticle Physics*, 6, June 2012.

S. Nadathur, S. Hotchkiss, J. M. Diego, I. T. Iliev, S. Gottlöber, W. A. Watson, and G. Yepes. Self-similarity and universality of void density profiles in simulation and SDSS data. *Monthly Notices of the RAS*, 449:3997–4009, June 2015. doi: 10.1093/mnras/stv513.

Mark C. Neyrinck. zobov: a parameter-free void-finding algorithm. *Monthly Notices of the Royal Astronomical Society*, 386(4):2101–2109, 2008. URL http://mnras.oxfordjournals.org/content/386/4/2101.abstract.

J. P. Ostriker. Discovery of "Dark Matter" In Clusters of Galaxies. *Astrophysical Journal*, 525:297, November 1999.

D. C. Pan, M. S. Vogeley, F. Hoyle, Y.-Y. Choi, and C. Park. Cosmic voids in Sloan Digital Sky Survey Data Release 7. *Monthly Notices of the RAS*, 421:926–934, April 2012.

A. Paranjape, T. Y. Lam, and R. K. Sheth. Halo abundances and counts-in-cells: the excursion set approach with correlated steps. *Monthly Notices of the RAS*, 420:1429–1441, February 2012a. doi: 10.1111/j.1365-2966.2011.20128.x.

A. Paranjape, T. Y. Lam, and R. K. Sheth. A hierarchy of voids: more ado about nothing. *Monthly Notices of the RAS*, 420:1648–1655, February 2012b. doi: 10.1111/j.1365-2966.2011.20154.x.

P. J. E. Peebles. *The large-scale structure of the universe*. Princeton University Press, 1980. 435 p., 1980.

A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophysical Journal*, 142:419–421, July 1965.

S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and T. S. C. Project. Measurements of Ω and Λ from 42 High-Redshift Supernovae. *Astrophysical Journal*, 517:565–586, June 1999.

A. Pisani, P. M. Sutter, N. Hamaus, E. Alizadeh, R. Biswas, B. D. Wandelt, and C. M. Hirata. Counting voids to probe dark energy. *Physical Review D*, 92(8): 083531, October 2015.

Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XIX. The integrated Sachs-Wolfe effect. *Astronomy and Astrophysics*, 571, November 2014.

Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. *ArXiv e-prints*, February 2015a.

Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, N. Bartolo, and et al. Planck 2015 results. XXI. The integrated Sachs-Wolfe effect. *ArXiv e-prints*, February 2015b.

E. Platen, R. van de Weygaert, and B. J. T. Jones. A cosmic watershed: the WVF void detection technique. *Monthly Notices of the RAS*, 380:551–570, September 2007.

G. Pollina, M. Baldi, F. Marulli, and L. Moscardini. Cosmic voids in coupled dark energy cosmologies: the impact of halo bias. *Monthly Notices of the RAS*, 455: 3075–3085, January 2016.

W. H. Press and P. Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *Astrophysical Journal*, 187:425–438, February 1974. doi: 10.1086/152650.

E. Ricciardelli, V. Quilis, and S. Planelles. The structure of cosmic voids in a cdm universe. *Monthly Notices of the Royal Astronomical Society*, 434(2):1192–1204, 2013. URL http://mnras.oxfordjournals.org/content/434/2/1192.abstract.

E. Ricciardelli, V. Quilis, and J. Varela. On the universality of void density profiles. *Monthly Notices of the Royal Astronomical Society*, 440(1):601–609, 2014. URL http://mnras.oxfordjournals.org/content/440/1/601.abstract.

A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich, R. L. Gilliland, C. J. Hogan, S. Jha, R. P. Kirshner, B. Leibundgut, M. M. Phillips, D. Reiss, B. P. Schmidt, R. A. Schommer, R. C. Smith, J. Spyromilio, C. Stubbs, N. B. Suntzeff, and J. Tonry. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *Astrophysical Journal*, 116:1009–1038, September 1998.

B. S. Ryden. Measuring Q 0 from the Distortion of Voids in Redshift Space. *Astrophysical Journal*, 452:25, October 1995.

R. K. Sachs and A. M. Wolfe. Perturbations of a Cosmological Model and Angular Variations of the Microwave Background. *Astrophysical Journal*, 147:73, January 1967.

C. Sánchez, J. Clampitt, A. Kovacs, B. Jain, J. García-Bellido, S. Nadathur, D. Gruen, N. Hamaus, D. Huterer, P. Vielzeuf, A. Amara, C. Bonnett, J. DeRose, W. G. Hartley, M. Jarvis, O. Lahav, R. Miquel, E. Rozo, E. S. Rykoff, E. Sheldon, R. H. Wechsler, J. Zuntz, T. M. C. Abbott, F. B. Abdalla, J. Annis, A. Benoit-Lévy, G. M. Bernstein, R. A. Bernstein, E. Bertin, D. Brooks, E. Buckley-Geer, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, M. Crocce, C. E. Cunha, C. B. D'Andrea, L. N. da Costa, S. Desai, H. T. Diehl, J. P. Dietrich, P. Doel, A. E. Evrard, A. Fausti Neto, B. Flaugher, P. Fosalba, J. Frieman, E. Gaztanaga, R. A. Gruendl, G. Gutierrez, K. Honscheid, D. J. James, E. Krause, K. Kuehn, M. Lima, M. A. G. Maia, J. L. Marshall, P. Melchior, A. A. Plazas, K. Reil, A. K. Romer, E. Sanchez, M. Schubnell, I. Sevilla-Noarbe, R. C. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, G. Tarle, D. Thomas, A. R. Walker, and J. Weller. Cosmic Voids and Void Lensing in the Dark Energy Survey Science Verification Data. *ArXiv e-prints*, May 2016.

B. P. Schmidt, N. B. Suntzeff, M. M. Phillips, R. A. Schommer, A. Clocchiatti, R. P. Kirshner, P. Garnavich, P. Challis, B. Leibundgut, J. Spyromilio, A. G. Riess, A. V. Filippenko, M. Hamuy, R. C. Smith, C. Hogan, C. Stubbs, A. Diercks, D. Reiss, R. Gilliland, J. Tonry, J. Maza, A. Dressler, J. Walsh, and R. Ciardullo. The High-Z Supernova Search: Measuring Cosmic Deceleration and Global Curvature of the Universe Using Type IA Supernovae. *Astrophysical Journal*, 507: 46–63, November 1998.

R. K. Sheth and G. Tormen. An excursion set model of hierarchical clustering: ellipsoidal collapse and the moving barrier. *Monthly Notices of the RAS*, 329: 61–75, January 2002.

R. K. Sheth and R. van de Weygaert. A hierarchy of voids: much ado about nothing. *Monthly Notices of the RAS*, 350:517–538, May 2004.

Ravi K. Sheth, H. J. Mo, and Giuseppe Tormen. Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, 323(1):1–12, 2001. URL http://mnras.oxfordjournals.org/content/323/1/1.abstract.

V. Springel. The cosmological simulation code GADGET-2. *Monthly Notices of the RAS*, 364:1105–1134, December 2005.

Yasushi Suto, Katsuhiko Sato, and Humitaka Sato. Expansion of voids in a matter-dominated universe. *Progress of Theoretical Physics*, 71(5):938–945, 1984. URL http://ptp.oxfordjournals.org/content/71/5/938.abstract.

P. M. Sutter, G. Lavaux, B. D. Wandelt, and D. H. Weinberg. A Public Void Catalog from the SDSS DR7 Galaxy Redshift Surveys Based on the Watershed Transform. *Astrophysical Journal*, 761:44, December 2012.

P. M. Sutter, G. Lavaux, N. Hamaus, B. D. Wandelt, D. H. Weinberg, and M. S. Warren. Sparse sampling, galaxy bias, and voids. *Monthly Notices of the RAS*, 442:462–471, July 2014a.

P. M. Sutter, A. Pisani, B. D. Wandelt, and D. H. Weinberg. A measurement of the Alcock-Paczyński effect using cosmic voids in the SDSS. *Monthly Notices of the RAS*, 443:2983–2990, October 2014b.

P. M. Sutter, E. Carlesi, B. D. Wandelt, and A. Knebe. On the observability of coupled dark energy with cosmic voids. *Monthly Notices of the RAS*, 446:L1–L5, January 2015a.

P. M. Sutter, G. Lavaux, N. Hamaus, A. Pisani, B. D. Wandelt, M. Warren, F. Villaescusa-Navarro, P. Zivick, Q. Mao, and B. B. Thompson. VIDE: The Void IDentification and Examination toolkit. *Astronomy and Computing*, 9:1–9, March 2015b.

M. Tegmark, M. R. Blanton, M. A. Strauss, F. Hoyle, D. Schlegel, R. Scoccimarro, M. S. Vogeley, D. H. Weinberg, I. Zehavi, A. Berlind, T. Budavari, A. Connolly, D. J. Eisenstein, D. Finkbeiner, J. A. Frieman, J. E. Gunn, A. J. S. Hamilton, L. Hui, B. Jain, D. Johnston, S. Kent, H. Lin, R. Nakajima, R. C. Nichol, J. P. Ostriker, A. Pope, R. Scranton, U. Seljak, R. K. Sheth, A. Stebbins, A. S. Szalay, I. Szapudi, L. Verde, Y. Xu, J. Annis, N. A. Bahcall, J. Brinkmann, S. Burles, F. J. Castander, I. Csabai, J. Loveday, M. Doi, M. Fukugita, J. R. Gott, III, G. Hennessy, D. W. Hogg, Ž. Ivezić, G. R. Knapp, D. Q. Lamb, B. C. Lee, R. H. Lupton, T. A. McKay, P. Kunszt, J. A. Munn, L. O'Connell, J. Peoples, J. R. Pier, M. Richmond, C. Rockosi, D. P. Schneider, C. Stoughton, D. L. Tucker, D. E. Vanden Berk, B. Yanny, D. G. York, and SDSS Collaboration. The Three-Dimensional Power Spectrum of Galaxies from the Sloan Digital Sky Survey. *Astrophysical Journal*, 606:702–740, May 2004.

R. van de Weygaert and W. Schaap. The Cosmic Web: Geometric Analysis. *ArXiv e-prints*, August 2007.

Rien Van de Weygaert and Erwin Platen. Cosmic voids: Structure, dynamics and galaxies. *International Journal of Modern Physics: Conference Series*, 01:41–66, 2011. URL http://www.worldscientific.com/doi/abs/10.1142/S2010194511000092.

Rien van de Weygaert and Eelco van Kampen. Voids in gravitational instability scenarios i. global density and velocity fields in an einsteinde sitter universe. *Monthly Notices of the Royal Astronomical Society*, 263(2):481–526, 1993. URL http://mnras.oxfordjournals.org/content/263/2/481.abstract.

G. Xu. A New Parallel N-Body Gravity Solver: TPM. *Astrophysical Journal*, 98: 355, May 1995.

L. F. Yang, M. C. Neyrinck, M. A. Aragón-Calvo, B. Falck, and J. Silk. Warmth elevating the depths: shallower voids with warm dark matter. *Monthly Notices of the RAS*, 451, August 2015.

Y. B. Zeldovich. A hypothesis, unifying the structure and the entropy of the Universe. *Monthly Notices of the RAS*, 160:1P, 1972.

A. R. Zentner. The Excursion Set Theory of Halo Mass Functions, Halo Clustering, and Halo Growth. *International Journal of Modern Physics D*, 16:763–815, 2007.

P. Zivick, P. M. Sutter, B. D. Wandelt, B. Li, and T. Y. Lam. Using cosmic voids to distinguish f(R) gravity in future galaxy surveys. *Monthly Notices of the RAS*, 451:4215–4222, August 2015. doi: 10.1093/mnras/stv1209.

F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, 1933.

F. Zwicky. On the Masses of Nebulae and of Clusters of Nebulae. *Astrophysical Journal*, 86:217, October 1937.