Cosmology from probe combination: synergies between the number counts of galaxy clusters and cosmic voids

Presentata da:

Davide Pelliciari

Relatori:

Ch.mo Prof. Federico Marulli

Ch.mo Prof. Lauro Moscardini

Dott. Sofia Contarini

Our truth is the intersection of independent lies
- Richard Levins

Abstract

The study of the distribution of matter on the largest scales can provide key information about the origin and the matter-energy content of the Universe, as well as its expansion rate and amplitude of today matter perturbations. This information can be interpreted within the *standard* Λ-cold dark matter model (ΛCDM), which is described by only six fundamental cosmological parameters. The latter can nowadays be constrained by means of a combined analysis of different cosmological probes. Indeed, constraints obtained by exploiting a single cosmological probe often show some degree of degeneracy, which can be potentially disentangled by means of the combination technique. Moreover, in the last decades Cosmology entered in a new *precision*-era, in which cosmological parameters are measured with sub-percent accuracy. Though this has strengthened our confidence about the ΛCDM model, it also led to some statistical tensions between constraints obtained with some probes at high and low redshift.

In this Thesis work we exploit the combination between the number counts of galaxy clusters and cosmic voids, which represent the most extreme objects of the matter density distribution of the today Universe. We analyse the statistics of these objects using the *Magneticum Pathfinder* hydrodynamical simulations, at redshifts $0.2 \leq z \leq 2$. Clusters of galaxies have been identified by means of a halo finder algorithm, which considers clusters as spherical overdensities. Cosmic voids have been identified in the simulated galaxy distributions using a geometrical void finder algorithm. Moreover, the void catalogues have been properly cleaned to match the definition of voids in the theoretical void size function model considered in this work. We compare the number counts of galaxy clusters with the theoretical halo mass function model proposed by Despali et al. (2016) and the number counts of cosmic voids with the Vdn size function model developed by Jennings et al. (2013), by performing a Bayesian MCMC analysis. In addition, to accurately compare the data to the theory, we calibrate these models to take into account the differences between the types of mass tracers used in this work and those used in the literature. Then we sample the posterior distributions of the

matter density parameter, Ω_m, and the today perturbation amplitude, σ_8, marginalising over the nuisance parameters of the considered models. We also analyse how less restrictive selections of the mass tracers can lead to biased cosmological constraints.

To obtain the combined constraints we apply three new numerical algorithms, implemented inside the CosmoBolognaLib C++/Python libraries, which allow us to perform the combination of two, or more, independent cosmological probes. In particular, we test the hypothesis of probe independence, verifying that the correlation between the different data-sets is statistically negligible. The implemented codes exploit three different algorithms for probe combination: *Posterior Product*, *Importance Sampling* and *Posterior as Prior*. The first method provides the standard technique for performing such analysis, i.e. by computing the simple product of the individual posterior distributions associated to each data-set. The second is based on the importance sampling technique, with which a given posterior distribution is sampled starting from another distribution (namely the one deriving from the other probe). According to this method, an *importance weight*, computed as the ratio of the given posterior distribution, is associated to each point in parameter space. In the third method, the combined confidence contours are obtained by considering the posterior of a given probe as the prior for the other, and viceversa.

The combined constraints obtained with the described methodologies are consistent with each other and are perfectly centred on the true cosmological parameters used to build the analysed simulations. Moreover, we find that the halo mass function and the void size function can be considered powerful complementary probes, providing a nearly perpendicular intersection in the Ω_m - σ_8 parameter space. This is a desired property in order to extract as much information as possible from the probe combination. The strongest constraints are obtained by using either the *Posterior Product* or the *Posterior as Prior* method, which both provide an improvement of about $4 - 5$ times with respect to the individual probe constraints. The *Importance Sampling* technique provides a slightly weaker constraining power, though totally in agreement with the other two in terms of posterior mean and standard deviation. Furthermore, the latter is found to be the fastest method, among the others, and can be applied to any pair of external cosmological probes. All the new implemented codes provide simple and flexible tools that will be soon applied to the data coming from currently available and next-generation wide-field surveys to perform powerful combined cosmological analyses.

Sommario

Le proprietà della distribuzione della materia su grande scala forniscono informazioni fondamentali sull'origine e sul contenuto di materia e di energia dell'Universo, così come sul suo tasso di espansione e sull'ampiezza delle perturbazioni di materia che osserviamo oggi. Queste informazioni sono interpretabili nel modello *standard* Λ-cold dark matter (ΛCDM), descritto da sei parametri cosmologici fondamentali. Questi possono essere efficientemente vincolati combinando diverse *probe* cosmologiche, ovvero diversi metodi per "sondare" il nostro Universo. Infatti, i vincoli ottenuti dall'analisi di singole *probe* mostrano spesso un certo livello di degenerazione, che può potenzialmente essere attenuato attraverso la tecnica di combinazione. Negli ultimi decenni, la Cosmologia è entrata in una nuova era in cui i parametri cosmologici vengono vincolati con un'accuratezza inferiore all'1%. Anche se da un lato questo ha rafforzato il modello ΛCDM, dall'altro ha fatto emergere possibili tensioni statistiche tra i vincoli ottenuti per alcuni parametri cosmologici quando vengono considerate *probe* ad alto e basso redshift.

In questo lavoro di Tesi abbiamo studiato la combinazione tra i conteggi di ammassi di galassie e di vuoti cosmici, che costituiscono gli oggetti più estremi della distribuzione di materia dell'Universo odierno, sfruttando le *Magneticum Pathfinder*, simulazioni cosmologiche idrodinamiche, a redshift $0.2 \leq z \leq 2$. Gli ammassi di galassie sono stati identificati come sovradensità sferiche da un algoritmo di ricerca interno a queste simulazioni, mentre i vuoti cosmici sono stati identificati nella distribuzione simulata di galassie impiegando un algoritmo di selezione basato su principi geometrici. Inoltre, abbiamo riscalato i cataloghi di vuoti in modo che potessero essere confrontati con il modello teorico considerato in questo lavoro per la loro distribuzione in raggio. Abbiamo poi confrontato i conteggi di ammassi di galassie con la funzione di massa teorica di aloni (Despali et al., 2016) mentre i conteggi di vuoti sono stati messi a confronto con il modello Vdn (Jennings et al., 2013), attraverso un'analisi Bayesiana che fa uso di catene di Markov Monte Carlo. Inoltre, per ottenere un confronto accurato tra dati e modelli teorici, abbiamo eseguito una calibrazione di questi ultimi, in modo tale da tenere conto della

differenza tra i tipi di traccianti di materia utilizzati in questo lavoro e quelli usati in letteratura. Abbiamo quindi estratto le distribuzioni a posteriori del parametro di densità della materia, Ω_m, e dell'ampiezza delle perturbazioni locali di materia, σ_8, marginalizzando sui restanti parametri interni ai modelli teorici. I vincoli ottenuti risultano perfettamente in accordo con il modello cosmologico adottato nelle simulazioni *Magneticum*. Abbiamo inoltre mostrato come possibili selezioni meno restrittive applicate ai traccianti di massa possano portare a vincoli cosmologici affetti da sistematiche.

Per ottenere i vincoli combinati abbiamo applicato tre nuovi algoritmi numerici, implementati nelle librerie C++/Python CosmoBolognaLib, che permettono di effettuare la combinazione di due o più *probe* cosmologiche indipendenti. L'ipotesi di indipendenza tra le due *probe* considerate è stata testata verificando che la correlazione tra i diversi set di dati utilizzati fosse trascurabile. I codici implementati sfruttano tre diversi algoritmi per la combinazione di *probe*: *Posterior Product*, *Importance Sampling* e *Posterior as Prior*. Il primo metodo fornisce la tecnica standard per attuare questo tipo di analisi, ovvero calcolando il prodotto delle singole distribuzioni a posteriori associate ai diversi set di dati. Il secondo è basato sulla tecnica di *importance sampling*, attraverso la quale una data distribuzione a posteriori viene estratta a partire da un'altra distribuzione, ovvero quella che descrive l'altra *probe* cosmologica. In base a questo metodo, un peso, calcolato attraverso il rapporto delle differenti distribuzioni a posteriori, viene associato a ogni punto dello spazio dei parametri. Nel terzo e ultimo metodo, i contorni di confidenza per la combinazione si ottengono considerando la distribuzione a posteriori di una data *probe* come la distribuzione a priori dell'altra, e viceversa.

I vincoli combinati, ottenuti sfruttando i metodi appena descritti, si sono rivelati consistenti tra loro, risultando perfettamente centrati sui parametri cosmologici usati per le simulazioni analizzate. In particolare, abbiamo mostrato come la funzione di massa degli aloni e la distribuzione in raggio dei vuoti cosmici possano essere considerate *probe* complementari per lo studio dei parametri di interesse, fornendo un'intersezione praticamente perpendicolare nello spazio dei parametri $\Omega_m - \sigma_8$. Quest'ultima è una proprietà fondamentale per ottenere la massima informazione proveniente dalla combinazione. I vincoli più stringenti sono stati ottenuti sia attraverso l'uso del metodo *Posterior Product*, sia utilizzando il *Posterior as Prior*, fornendo un miglioramento di $4-5$ volte rispetto ai vincoli delle singole *probe*. Abbiamo anche notato che l'*Importance Sampling* fornisce vincoli leggermente meno stringenti, restando comunque in accordo con gli altri metodi in termini di media e deviazione standard della distribuzione a posteriori. Inoltre, quest'ultimo metodo si è rivelato essere il più veloce rispetto agli altri, con

l'ulteriore vantaggio di poter essere applicato a una qualsiasi coppia di *probe* cosmologiche fornite esternamente.

I nuovi codici implementati forniscono strumenti semplici e flessibili per effettuare la combinazione di *probe* cosmologiche, e verranno presto applicati a dati provenienti da *wide-field surveys* attualmente disponibili e di futura generazione.

Contents

Introduction

The discovery of the accelerated expansion of the Universe (Riess et al., 1998; Perlmutter et al., 1999) has lead to the evidence of a mysterious energy component, called *dark energy*. In addition, several observations suggest that the most of the matter is in the form of yet another mysterious component, the cold dark matter (CDM), which interacts only through gravity and, thus, it is not visible. All this information is encapsulated in the Standard Cosmological Model, known as the Λ-cold dark matter (ΛCDM), where Λ is the so-called cosmological constant, responsible for the current accelerated expansion of the Universe. This model, based on the General Theory of Relativity, has been tested so far by several measurements and observations, and it describes the overall properties of our Universe with only six fundamental parameters.

According to this framework, the Universe shows an age of approximately 13.8 Gyr and its matter-energy content is composed of about 70% of dark energy and 25% of dark matter. The remaining $\sim 5\%$ is in the form of baryonic matter, that is the matter composed of protons, neutrons and electrons, constituting the visible fraction of the matter content, of which stars and galaxies are made. The present-day large-scale structures (LSS) are supposed to grow from small density perturbations in the early Universe, as the consequence of the attractive nature of gravity. As the Universe continues expanding and cooling down, these fluctuations grow over the cosmic time, giving rise to collapsed haloes of DM. Then, following the so called *bottom-up* scenario, DM haloes evolve hierarchically assembling into larger structures. Their gravitational potential wells drive the collapse of baryonic matter, which is characterised also by electro-magnetic interactions and is therefore subject to radiative cooling. The complex nonlinear physical mechanisms involved in this process strongly affect the galaxy formation and evolution. Due to this complexity, there are no exhaustive and complete models yet able to fully describe these processes, which are indeed commonly investigated with cosmological simulations. These simulations follow the evolution of billions of particles of dark and baryonic matter, and are able to reproduce the observed

distribution of tracers, like DM haloes, galaxies and clusters.

In Figure 1 we report a representation of the Magneticum Pathfinder simulations, which are the ones we will analyse in this work. As one can see, this distribution follows a foam-like pattern, usually called Cosmic Web, in which galaxies aggregate into filaments, with sizes exceeding 100 Mpc h^{-1}. This type of pattern has been confirmed by several observation from deep surveys (York et al., 2000; Peacock et al., 2001; Dark Energy Survey Collaboration et al., 2016; Alam et al., 2017; Ivezić et al., 2019). In correspondence of the filament intersections, we find groups of hundreds up to thousands of galaxies. The most massive between them - galaxy clusters - represent the largest virialised and gravitationally-bounded structures in our Universe. Complementing this fascinating pattern is the presence of vast underdense regions of space - cosmic voids - which have roundish shapes and occupy the majority of the volume of the Universe.

This Thesis work concerns the complementary study of the abundance of galaxy clusters and cosmic voids identified in cosmological simulations, at different redshifts. In particular, we focus our analyses on the combination of the cosmological constraints derived from these probes, which can be considered statistically independent, given the different aspects of Universe density field they map. Indeed, we aim at showing the orthogonality of the derived cosmological constraints and the resulting impressive power of the combination of these probes. To perform this combination we apply three newly implemented algorithms that allow us to combine independent probes. These algorithms represent a flexible and user-friendly tool to perform different techniques for probe combination and are implemented within the environment provided by the large set of *free software* C++/Python CosmoBolognaLib (Marulli et al., 2016). The usage of the new implemented codes is not obviously limited to the exploitation of cluster and void abundances presented in this work: they will be applied in the near future to a vast variety of cosmological probes, extending the analysis also to correlated probes, as already been done in e.g. Webster et al. (1998), Gawiser & Silk (1998), Bridle et al. (1999) and DES Collaboration (2019). The probe combination represents indeed a powerful tool to maximise the synergies between different probes and to possibly shed light on the tensions that are currently threatening the ΛCDM model (see e.g. Di Valentino et al., 2020a,b,c,d, for an extended review).

This Thesis is organised as follows:

- in Chapter 1 we introduce the theoretical background upon which the Standard Cosmological Model is based;

- in Chapter 2 we present the linear growth of perturbations, from which

galaxy clusters and cosmic voids form and evolve. In this Chapter we also give a description of the main statistical properties of the LSS as well as the main numerical algorithms employed to carry out cosmological simulations;

- in Chapter 3 we present how to exploit galaxy clusters and cosmic voids as cosmological probes. We present the theoretical models of the halo mass function and the void size function, with which we can predict the abundance of clusters and voids, respectively. We introduce then different finding algorithms for haloes and cosmic voids, describing the preparation of the catalogues used in this work;

- in Chapter 4 we first introduce the C++/Python libraries CosmoBolognaLib, which offer the numerical environment to carry out the entire analysis. Then we describe the fundamental concepts of Bayesian statistics required for the cosmological parameter estimation. We then present the methods and the new implemented algorithms to perform the combination of independent probes. In particular, the exploited techniques are: the *Posterior Product*, the *Importance Sampling* and the *Posterior as Prior* method;

- in Chapter 5 we apply and compare the implemented algorithms to combine the number counts of galaxy clusters and cosmic voids, identified in the Magneticum Pathfinder simulations, at $0.2 \leq z \leq 2$. We present the calibration procedures applied to the theoretical models to predict correctly the abundance of the analysed objects. By performing a Bayesian Markov Chain Monte Carlo analysis, we obtain the joint posterior distributions in the parameter space $\Omega_m - \sigma_8$ by exploiting individually the halo mass function and the void size function. Then we combine the latter with the proposed algorithms obtaining remarkably tighter constraints and showing their effective constraining power as complementary probes;

- in Chapter 6 we summarise the scientific problem on which this Thesis is focused and outline the main analyses performed. Also, we discuss about the different caveats and approximations adopted in this work, as well as possible future applications of this methodology to wide-field surveys;

- Finally, in Appendix A we test the hypothesized independence between the halo mass function and the void size function, by computing their cross-covariance matrix with a Jackknife resampling technique.
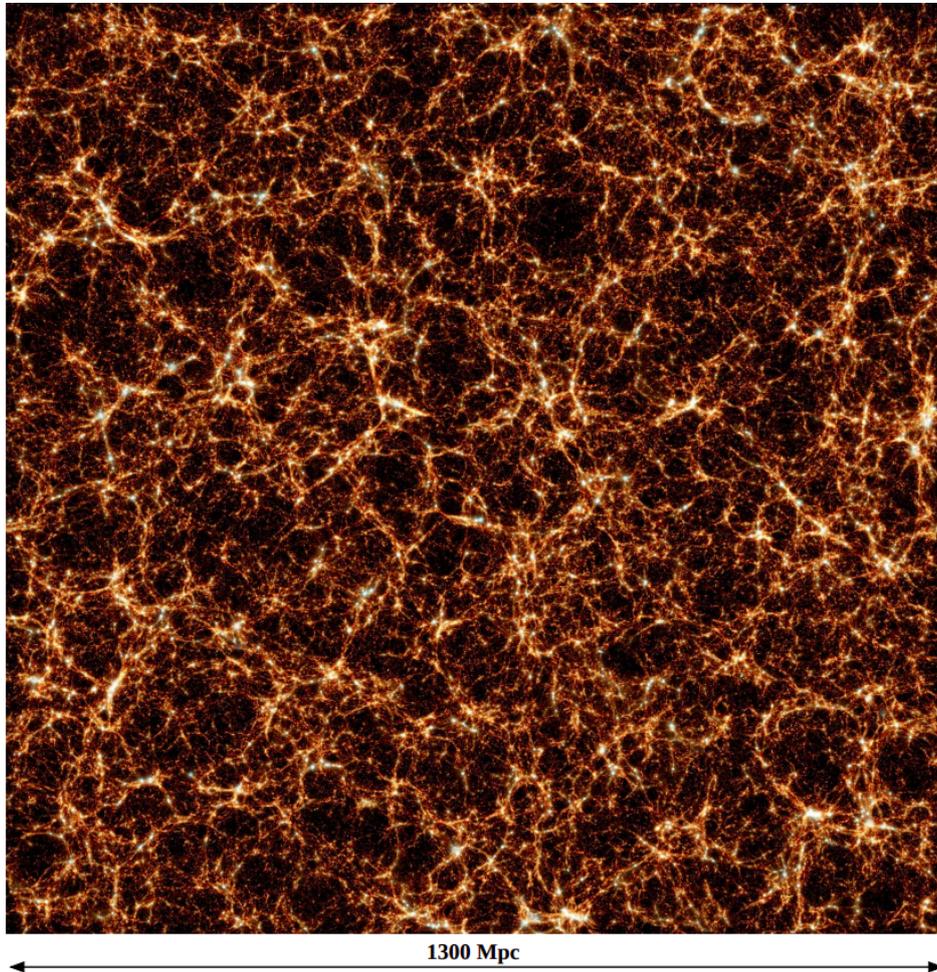
Figure 1: The spatial distribution of the baryonic matter at $z = 0$ in the simulated comoving cosmological Box1, which represents a total volume of $(896 \text{ Mpc}/h)^3$, from the Magneticum Pathfinder Simulation. The image shows a 100 Mpc h^{-1} thick and 1300 Mpc wide slice of this box. The colour, going from dark to light, indicates the gas temperature from cold to hot, while the stellar component is coloured in white. The typical features of the large-scale structure can be clearly seen: galaxy clusters, filaments and voids. Credits to Dolag et al. (2015).

Chapter 1

Cosmological framework

The study of the physical properties of our Universe is a very wide subject of research, going from planets in our Solar System, to galaxies orbiting around clusters of galaxies. Cosmology focuses on the largest scales of our Universe, considering the latter as a whole. Moreover, thanks to the cosmological model, that is the framework we are going to describe in this first Chapter, it is possible to predict how the large-scale structures evolve from an initial configuration. Until the 18th century, this field of study was only speculative, since no observational validations were available. Indeed, the first quantitative estimates of the structure of the local Universe were made by William Herschel in the late 18th century, based on the counts of stars in the sky. The introduction of the photography to record astronomical images, together with the advent of bigger and bigger telescopes led to a complete revolution in Cosmology. With these technologies it has been possible to resolve distant nebulae and stars, and to study their properties in deep details. The study of variable Cepheid stars inside globular clusters, that are gravitationally bounded groups of thousands of stars, led to the observational discovery of the period-luminosity relation of variable stars by Henrietta Leavitt (Leavitt & Pickering, 1912), from which it was possible to estimate with high precision the distances of astronomical objects. In 1926, thanks to the Leavitt's relation, Edwin Hubble was able to measure the distance of spiral nebulae (what we today call galaxies), given the first description of them as extragalactic objects (Hubble, 1925, 1926).

A decade before, Albert Einstein developed the General Theory of Relativity, which completely revolutionised our understanding of gravity as a bending of spacetime due to the presence of massive objects. This theory became the theoretical foundation of Cosmology, and the works of Friedmann, together with Einstein, de Sitter and Lemaître showed that the Einstein's equations, when applied to the study of the Universe, could predict also a

space in expansion. In 1929, Edwin Hubble discovered the distance-redshift relation, which led to the discovery of the expansion of the Universe. From that time Cosmology evolved very rapidly, and today we can analyse the property of the Universe on very large scales with high precision, both with data coming from observations of deep surveys, or with large cosmological simulations.

We start by describing the fundamental principles on which Cosmology is based. Then we describe in details the Friedmann-Lemaître-Robertson-Walker metric, which allows to obtain a general solution to the Einstein's equations, regarding a theoretical description of our Universe. Finally, we will describe the standard cosmological model, which is nowadays the most reliable model to describe the observed properties of the large-scale structures.

1.1   The cosmological principle

Nowadays we can see that our Universe looks isotropic on scales greater than, roughly, 100 Mpc (where 1 Mpc $= 10^6$pc $\approx 3.09 \cdot 10^{13}$ km). This principle tells us that on large scales there is no preferred directions or locations, and that does not exist a centre of the Universe. On smaller scales the Universe manifests itself as clumpy and particularly inhomogeneous. Indeed, let us grow a sphere of increasing radius around us. Most of the spheres of radius 3 AU (where 1 AU $\approx 1.5 \times 10^8$ km, roughly the distance from Earth to the Sun) will not contain any stars, and most of the spheres 3 Mpc across will not contain a pair of bright galaxies. Only if we consider spheres having a radius greater than 100 Mpc, then they will contain statistically the same pattern of clusters and voids. Due to the symmetries that this principle implies, we can set a cosmological time which allows us to have a reference time to study the Universe dynamics.

1.2   Fundamentals of General Relativity

On very large scales the predominant interaction between massive bodies such as galaxies and galaxy clusters is the *gravitational interaction*, briefly gravity. For more than two centuries the theory proposed by Isaac Newton has been the leading theory in describing gravitation, and even nowadays this theory is an excellent approximation of gravity in most of the classical applications. In this fashion, the gravitational force exerted on a body depends only on its mass and its distance from the gravitational source. Before Newton's

discoveries, no one imagined that the force that makes an apple fall to the ground was the same force that holds planets in orbit around the Sun.

In 1905, the so-called *Annus Mirabilis*, Albert Einstein, aware of the fact that the Newtonian mechanics was no longer enough to reconcile the laws of classical mechanics with the electromagnetic laws, developed the theory of Special Relativity. In this new theory the speed of light in the vacuum is the same for all the observers, regardless of the relative motions between them and light. Another consequence of the Einstein's theory was that no information could propagates faster that the speed of light. One particular characteristic of the gravitational interaction introduced by Newton is that the force undergone by a body is instantaneous, and that means that the Newtonian gravity was incompatible with Special Relativity. Einstein took 11 years to develop the General Theory of Relativity, which is nowadays the best theory that describes gravity, hence the fundamental theory behind cosmology.

In General Relativity space and time blend together in the concept of *spacetime*, and so there is no longer a universal clock that marks the passing of time. A generic point in spacetime, called *event*, is described by a 4-vector **x**, which in cartesian coordinates is written as

$$\mathbf{x} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \equiv \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \equiv x^\mu, \quad \text{for } \mu = 0, 1, 2, 3 , \qquad (1.1)$$

where $x^0 = ct$ it is the time coordinate of the spacetime event and t is the proper time, i.e. the time measured by a clock that is comoving with the observer. A trajectory γ of a body in spacetime is called *wordline*.

The theory of General Relativity totally upset the paradigms on gravity, in fact the latter is no longer a force, but it is a natural consequence of the curvature of spacetime. The geometry of spacetime is described by the *metric tensor* $g_{\mu\nu}$, which allows to determine the distance ds^2 between two infinitesimal close events labelled $x^j = (ct, x, y, z)$ and $x^j + dx^j = (ct + cdt, x + dx, y + dy, z + dz)$:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu . \qquad (1.2)$$

Applying the principle of least action, one can obtain the laws of motions for free-particles, called geodesics, which are no longer straight lines due to the curvature of spacetime. Therefore varying the action S on a worldline γ:

$$\delta S[\gamma] = \delta \int_\gamma ds = \delta \int_\gamma \sqrt{g_{\mu\nu}dx^\mu dx^\nu} ,$$

leads to the geodetic equations:

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}s^2} + \Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}s}\frac{\mathrm{d}x^\beta}{\mathrm{d}s} = 0 \,. \tag{1.3}$$

Here the $\Gamma^\mu_{\alpha\beta}$ coefficients are the *Christoffel* symbols, also known as the *Levi-Civita connections* related to the first partial derivative of the metric tensor components.

The geometry of the spacetime is connected to the content of energy and matter, which is described by the *stress-energy* tensor $T_{\mu\nu}$, in the famous *Einstein field equations* (EFE):

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu} \,, \tag{1.4}$$

where G represents the universal gravitational constant, c the speed of light, while $R_{\mu\nu}$ and R are the *Ricci tensor* and scalar curvature, respectively, defined starting from the *Riemann tensor*:

$$R_{\mu\nu} = R^\delta_{\mu\delta\nu} \,,$$

$$R = R^\mu_\mu = g^{\mu\nu}R_{\mu\nu} \,.$$

The Riemann tensor is related to the first derivative of the Christoffel symbols and so to the second derivative of the metric tensor components, and describe the intrinsic curvature of the spacetime.

Regarding the RHS of Eq. (1.4), the Universe can be modelled as a perfect fluid with pressure p and energy density ρc^2, so the energy momentum tensor can be expressed as:

$$T_{\mu\nu} = -pg_{\mu\nu} + (p + \rho c^2)u_\mu u_\nu \,, \tag{1.5}$$

where u_μ is the generic component of the *quadrivelocity*, defined as $\mathbf{u} = \dfrac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}$, with $\mathbf{x}$ the 4-position of the relation (1.1).

1.3 The Friedmann-Lemaître-Robertson-Walker metric

In order to describe our Universe within a physical model, we have to find a solution to the Einstein field equations that includes all the different aspects of the cosmological principle. The different components of the EFE depend only on the metric of the spacetime, and so giving a solution to these equations means finding a suitable metric for the problem.

From the CP, starting with the metric given as Eq. (1.2), the mixed terms g_{0i} need to be null, so we can write:

$$ds^2 = c^2dt^2 - g_{ij}dx^i dx^j = c^2dt^2 - dl, \quad i = 1, 2, 3. \tag{1.6}$$

In this case g_{ij} are the *spatial* metric tensor components. Any point in space-time is marked by three comoving spatial coordinates x^i for $i = 1, 2, 3$ and one comoving time coordinate t, where comoving means at rest with the expansion of the Universe, so that an observer, placed at the centre of this reference system, sees the latter expanding uniformly around him.

The only geometrical spaces that satisfy the Cosmological Principle are the *flat Euclidean space*, the *sphere* and the *hyperboloid*. Let us now consider the spherical coordinates (ρ, θ, ϕ), related to the Cartesian ones, by the transformation:

$$\begin{cases} x^1 = \rho \sin\theta \cos\phi \\ x^2 = \rho \sin\theta \sin\phi \\ x^3 = \rho \cos\theta \, , \end{cases} \tag{1.7}$$

where the ranges of values are $0 \le \rho < \infty$, $0 \le \theta < \pi$ and $0 \le \phi < 2\pi$.

For the case of a flat Euclidean space the distance between two points, hence the metric, becomes:

$$dl^2 = d\rho^2 + \rho^2 d\theta^2 + \rho^2 \sin^2\theta d\phi^2 \, . \tag{1.8}$$

Now let us define the *scale factor* a as:

$$\rho \equiv ar \, , \tag{1.9}$$

in such a way that a has the dimension of length and r has no dimension.

Therefore Eq. (1.8) reduces to:

$$dl^2 = a^2(dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2) = a^2(dr^2 + r^2 d\Omega_2^2) \, , \tag{1.10}$$

where $d\Omega_2^2 = d\theta^2 + \sin^2\theta d\phi^2$ is the metric of the unitary 2-sphere. In practice all the 3D space is made by concentric spheres having radius r.

The metric for the 3-sphere, that is a space with a positive intrinsic curvature, is of the form:

$$ds^2 = a^2(d\chi^2 + \sin^2\chi d\Omega_2^2) \, . \tag{1.11}$$

Now let us define the radial coordinate $r \equiv \sin\chi$, so that its differential becomes $dr = \cos\chi d\chi$. With this substitution, the metric is:

$$dl^2 = a^2\left(\frac{dr^2}{1 - r^2} + r^2 d\Omega_2^2\right) . \tag{1.12}$$

9

The case for the hyperboloid is very similar to the 3-sphere, unless that the trigonometric functions are hyperbolic, and now $r \equiv \sinh \chi$:

$$dl^2 = a^2(d\chi^2 + \sinh^2 \chi d\Omega_2^2) = a^2\Big(\frac{dr^2}{1+r^2} + r^2 d\Omega_2^2\Big). \qquad (1.13)$$

One can easily find the general expression that summarises the three metrics found in equations (1.10), (1.12) and (1.13):

$$dl^2 = a^2\Big(\frac{dr^2}{1-\kappa r^2} + r^2 d\Omega_2^2\Big), \qquad (1.14)$$

where we have introduced the *curvature parameter* κ, which is constant and can assume only three values, depending on what the intrinsic curvature is:

$$\kappa = 1 \quad \text{positive curvature}$$
$$\kappa = 0 \quad \text{zero curvature;}$$
$$\kappa = -1 \quad \text{negative curvature.}$$

By adding the time component of the metric, we obtain the functional form of the *Friedmann-Lemaître-Robertson-Walker* (FLRW) metric:

$$ds^2 = c^2 dt^2 - a(t)^2\Big(\frac{dr^2}{1-\kappa r^2} + r^2 d\Omega_2^2\Big), \qquad (1.15)$$

where $a = a(t)$ is the *cosmic scale factor* and has the dimensions of a length and (r, θ, ϕ) are the comoving dimensionless coordinates. This metric represents the infinitesimal distance between two events of a perfectly homogeneous and isotropic 4D-spacetime.

The scale factor is a dimensional function of time which describes how distances grow or decrease with time, while the curvature parameter is related to the *Gaussian curvature* κ_G:

$$\kappa_G = \frac{\kappa}{a^2}.$$

The assumptions of homogeneity and isotropy for the metric are very strong and powerful. Indeed in this case all we have to know about the geometry and the expansion of the Universe are the scale factor $a(t)$ and the curvature κ. This clearly represents an approximation, because the CP holds only for the largest scales in the Universe. At smaller scales, the expansion is not uniform everywhere and we observe matter lumps held together by the gravitational force, such as galaxies and clusters of galaxies.

1.4 The Hubble-Lemaître law

In Cosmology there is no longer a unique way to assign a value for the distance between two points, because the Universe's size changes with time. If we want to measure the distance between us and a given galaxy, we must specify the time at which the measurement is performed, i.e. we must know the value of the scale factor at that moment in time. The most simple definition of it is the *proper distance*, which is the length of the spatial geodesics passing between the two points, considering the scale factor fixed at $a = a(t)$. We can obtain the proper distance between the observer, placed at the origin, and a galaxy placed at comoving coordinates (r, θ, ϕ) by integrating the FLRW metric (1.15) in the case of $\mathrm{d}t = \mathrm{d}\theta = \mathrm{d}\phi = 0$:

$$d_{pr}(t) = a(t)F(r) = a(t)\int_0^r \frac{\mathrm{d}r'}{\sqrt{1 - \kappa r'}} \,. \tag{1.16}$$

Depending on what the curvature is, the value of the integral becomes:

$$F(r) = \begin{cases} \sin^{-1} r; & \kappa = +1 \\ r; & \kappa = 0 \\ \sinh^{-1} r; & \kappa = -1 \end{cases} \tag{1.17}$$

Consequently, the proper distance is a *simultaneous* measure with a chain of rulers connecting the two points at a given time t, hence it does not take into account the speed at which information travels.

We can also define a *comoving distance*, which does not depend on the expansion of the Universe, as the proper distance computed today ($t = t_0$):

$$d_c = d_{pr}(t_0) = a(t_0)F(r) = \frac{a(t_0)}{a(t)}d_{pr}(t) \,. \tag{1.18}$$

In order to estimate how fast two points drift apart as a consequence of the expansion, let us derive the proper distance with respect to the time coordinate:

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

since $\dot{F}(r) = 0$. Here the dot notation represents the derivation with respect to time.

From the definition of proper distance we obtain:

$$v = \frac{\dot{a}(t)}{a(t)}d_{pr} \equiv H(t)d_{pr} \,. \tag{1.19}$$

The latter equation is called the *Hubble-Lemaître law*, and determines the velocity with which different bodies in spacetime drift apart with respect to each other due to the expansion of the Universe. The proportionality between the velocity and the proper distance is given by the so-called *Hubble parameter*:

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)}.\qquad(1.20)$$

which has the same value for all the points in the Universe at a fixed time.
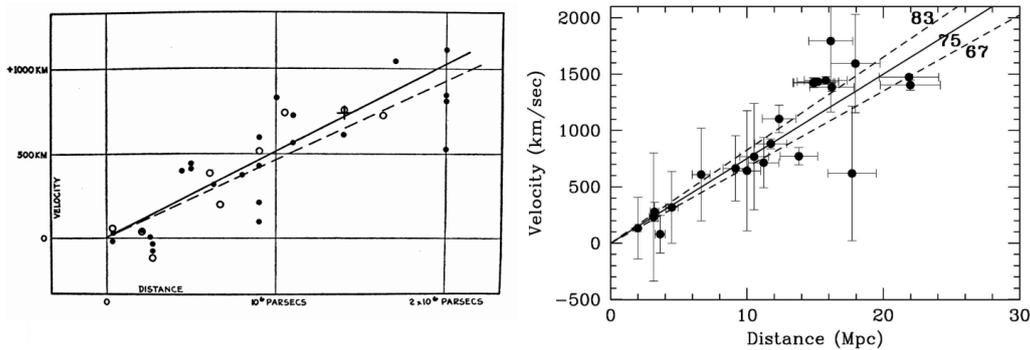
Figure 1.1: The distance-velocity relation of extragalactic sources originally obtained by Hubble in 1929 (*left panel*) and a more modern version of it obtained by Freedman et al. (2001) (*right panel*). The latter is achieved by precisely measuring the distance of Cepheid stars from observations of the Hubble Space Telescope. Credits to Hubble (1929) and Freedman et al. (2001).

Its value at the present time, indicated with H_0, is called *Hubble constant*. It is a fundamental cosmological parameter, because it sets the scales of time and distance in our Universe. The unit of measure of H_0 is s^{-1}, i.e. an inverse of the time, however it is commonly expressed in terms of km s^{-1} Mpc^{-1}, so that a galaxy 1 Mpc distant from us has a recession velocity of H_0 km s^{-1}. The value of H_0 is of fundamental importance in Cosmology, and its name derives from the pioneering works of Edwin Hubble, who in 1929 discovered the expansion of the Universe from the observations of 34 distant galaxies for which velocities had been measured, all within 2 Mpc from our Galaxy. His first estimate of this constant was $H_0 \simeq 500$ km s^{-1} Mpc^{-1}, but he was wrong by about a factor of 7. Indeed, his measurements were done with many approximations, but still he succeed to obtain the distance-velocity relation (Hubble, 1929). In Figure 1.1 it can be seen the original plot of this relation obtained by Hubble, together with a modern version of it obtained

by Freedman et al. (2001). This relation tells us that the more galaxies are distant, the more they drift apart from us with greater velocity.

Over the past 25 years many efforts have been made to estimate the value of H_0. However, the precision with which modern measurements are performed has led to the rise of tensions on the values of H_0 between different type of measurements (see Section 1.10.1). Because of these discrepancies we will express cosmological quantities as a function of the parameter $h(t)$, defined as

$$h(t) = \frac{H(t)}{100} \text{ km s}^{-1} \text{ Mpc}^{-1} \,, \tag{1.21}$$

which allows us to set a theoretical model independent of the exact value of the Hubble constant.

Lastly, the inverse of the Hubble constant H_0, known as the *Hubble time* τ_H, provides a rough estimate of the age of the Universe. It reflects the time since a linear cosmic expansion has begun (extrapolating a linear Hubble Law back to time $t = 0$). Setting $H_0 = 70 \text{ kms}^{-1} \text{ Mpc}^{-1}$, the age of the Universe is approximated as $\tau_H \simeq 14$ Gyr (where 1 Gyr = 10^9 yr).

1.5   The cosmological redshift

The global motion of objects in the Universe with respect to each other, due to its expansion, is called *Hubble flow*. One of its consequences is that the electromagnetic radiation coming to us from very distant objects is reddened, i.e. it is shifted towards longer wavelengths. This phenomenon, similar to the *Doppler effect* of sound waves, is called the *redshift* of the electromagnetic spectrum.

Let us consider an electromagnetic source that emits light at a specific monochromatic wavelength λ_{em}. With this value we indicate the value of the wavelength at rest with respect to the source. Let us indicate the shifted wavelength of the radiation which arrives to the observer as λ_{obs}. The redshift z of the electromagnetic radiation can be defined as:

$$z \equiv \frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}} = \frac{\Delta\lambda}{\lambda} \,, \tag{1.22}$$

which in principle can be less than zero (*blueshift*), indicating that the source is approaching the observer, or greater than zero (*redshift*), when the source is receding.

Now, let us consider t_{em} as the time at which the source emitted the radiation, and t_{obs} the time at which the observer, at a distance r, received it. By definition, photons move along null geodesics, which are denoted by

$ds^2 = 0$. Considering the FLRW metric (1.15) and taking $d\theta = d\phi = 0$ (radial motions only), a null geodesics satifies

$$\frac{c^2 dt^2}{a(t)^2} = \frac{dr^2}{1 - \kappa r^2} \, ,$$

(1.23)

which can be integrated along the path followed by the photon:

$$\int_{t_{em}}^{t_{obs}} \frac{cdt}{a(t)} = \int_0^r \frac{dr}{\sqrt{1 - \kappa r^2}} = F(r) \, .$$

(1.24)

Now, let us assume that a second photon is emitted by the same source at a time $\tilde{t}_{em} = t_{em} + \delta t_{em}$[1] and that it is received by the observer at the time $\tilde{t}_{obs} = t_{obs} + \delta t_{obs}$. By the fact that both the source and the observer are moving with the cosmological expansion, the integral $F(r)$ stays fixed in time because r does not change.

From Eq. (1.24) we get

$$\frac{\delta t_{obs}}{a(t_{obs})} = \frac{\delta t_{em}}{a(t_{em})} \, .$$

(1.25)

Since $\delta t = 1/\nu$ and $\lambda \nu = c$, we can write

$$\frac{\lambda_{obs}}{\lambda_{em}} = \frac{a(t_{obs})}{a(t)} \, ,$$

which leads to the relation between the redshift and the scale factor:

$$1 + z = \frac{a_0}{a_{em}} \, ,$$

(1.26)

where $a_0 = a(t_{obs})$.

Since the Universe is expanding, the scale factor grows, thus the redshift is positive and the distant galaxies move away from us. This type of redshift is called *cosmological*, in order to distinguish it from local phenomena that contribute to the shift of the electromagnetic radiation, e.g the radial proper motion of galaxies or the gravitational redshift due to the dilation of time in the potential wells of massive objects such as black holes and quasars. The observed redshift is a combination of these three types of redshifts. From the Doppler effect, we can link the redshift to the radial velocity of the source with respect to the observer:

$$z \simeq \frac{v_r}{c} \, .$$

(1.27)

---

[1]Here with δt we indicate an infinitesimal interval of time, such that $\delta t \ll t$.

The peculiarity of the cosmological redshift is that z can also be greater than unity, which is impossible for radial motions because $v_r < c$. This can happen because the expansion of the spacetime that separates the galaxies can be faster than the speed of light.

From Eq. (1.27) and considering distances $D \ll c/H_0$, the Hubble-Lemaître law results:

$$z \simeq \frac{H_0 D}{c} \, , \tag{1.28}$$

which again tells us that the redshift of distant galaxies increases with their distance.

1.6 Cosmological distances

We have shown how the comoving coordinates, i.e. the FLRW metric, are connected to the concept of proper distance. The latter is the measure of distance between events happening at the same proper time, so it is easy to realize that this measure is physically impossible to make.

Cosmology is based on observations, therefore it is useful to define a way to compute distances from the observational properties of galaxies, such as their redshift, as well as their flux or their angular diameter. One way of using measured properties to assign a distance is the *standard candle* method. A standard candle object is an object with known luminosity L. In principle, if there exist objects with the same intrinsic luminosity throughout the spacetime, one can compute their distance, called *luminosity distance*, by measuring their flux f, i.e. their luminosity per unit area:

$$d_L = \left(\frac{L}{4\pi f}\right)^{1/2} . \tag{1.29}$$

The *flux* of the source, measured by the observer placed at P_0 at time t_0, can be expressed as:

$$f = \frac{L_{obs}}{4\pi d^2} \, . \tag{1.30}$$

The denominator is the surface area of a sphere centred in P_0. This surface is dilated by the Universe expansion, so we have to write it as: $4\pi d^2 = 4\pi a^2(t_0)r^2$. Moreover, the light coming from the source is being redshifted, therefore we have to take it into account in the luminosity calculation. The emitted luminosity is defined as the rate of change of the energy of the source:

$$L_{em} = \frac{dE}{dt}. \tag{1.31}$$

As one can see from Eq. (1.25), photons emitted by the source in a small interval δt_{em} arrive to the observer in an interval $\delta t_{obs} = (a_0/a)\delta t_{em}$. Taking these two effects into account we can rewrite Eq. (1.31) as:

$$L_{em} = L_{obs}\left(\frac{a_0}{a_{em}}\right)^2 , \tag{1.32}$$

so that the flux l becomes

$$f = \frac{L_{obs}}{4\pi a_0^2 r^2} = \frac{L_{em}}{4\pi a_0^2 r^2}\left(\frac{a_{em}}{a_0}\right)^2 . \tag{1.33}$$

In order to preserve the Euclidean inverse-square law for the diminution of luminosity with distance from a point source, the luminosity distance is defined as

$$d_L \equiv \frac{a_0^2 r}{a_{em}} = a_0 r(1 + z) . \tag{1.34}$$

Another method to compute distances from observational properties is the *standard rulers* method, in which we observe objects with known intrinsic dimension $\ell \equiv D_{pr}$. Suppose the standard ruler is aligned perpendicularly to the line of sight, and indicating with $\Delta\theta$ its angular extension, then we can compute the so-called *angular-diameter* distance d_A as $d_A = \frac{D_{pr}}{\Delta\theta}$. The distance between the two ends of the ruler, at a given time t corrisponding to the emission of the radiation from the source, can be found with the FLRW metric as

$$\mathrm{d}s = a(t)r\Delta\theta . \tag{1.35}$$

From the fact that we know a-priori the length ℓ of the ruler, we can set $\ell = \mathrm{d}s$, and so the latter becomes:

$$\ell = a(t)r\Delta\theta , \tag{1.36}$$

from which it is possible to derive the following expression for the angular-diameter distance:

$$d_A \equiv \frac{\ell}{\Delta\theta} = ar . \tag{1.37}$$

By comparing the luminosity distance with the angular diameter distance, we can obtain the *duality relation*:

$$\frac{d_L}{d_A} = (1 + z)^2 . \tag{1.38}$$

Thus, if one observes an object which is both a standard ruler and a standard candle, its angular-diameter distance will be smaller than its luminosity distance.

Theoretically, this relation could be used to study the deviation of the metric from the FLRW, and hence to test the Cosmological Principle. The ideal way to test the duality relation is to work with both standard rulers and standard candles, without using a specific cosmological model (Holanda et al., 2012). A prominent approach is the observation of massive galaxy clusters, which allow to obtain both the angular diameter distance and the luminosity distance from the Sunyaev-Zel'dovich effect (SZ) [2] and their X-ray observations (Uzan et al., 2004).

1.7 The deceleration parameter

In order to study the cosmic scale factor as a function of time, we can expand it in Taylor series around $t = t_0$:

$$
\begin{aligned}
a(t) &= a_0 + \frac{\mathrm{d}a(t)}{\mathrm{d}t}\Big|_{t=t_0}(t - t_0) + \frac{1}{2}\frac{\mathrm{d}^2 a(t)}{\mathrm{d}t^2}\Big|_{t=t_0}(t - t_0)^2 + \dots \\
&= a_0\Big[1 + \frac{\dot{a}_0}{a_0}(t - t_0) + \frac{\ddot{a}_0}{2a_0}(t - t_0)^2 + \dots\Big].
\end{aligned}
\tag{1.39}
$$

To reproduce exactly the scale factor as a function of time we would need an infinite number of terms. However, if $a(t)$ does not fluctuate widely with t, then the Taylor expansion up to the second term will give a good approximation in the vicinity of $t = t_0$.

In the last expression the Hubble constant is the coefficient of the linear part of the series. Let us define the dimensionless *deceleration parameter* as:

$$
q \equiv -\frac{\ddot{a}a}{\dot{a}^2}.
\tag{1.40}
$$

Therefore the Taylor series of $a(t)$ reduces to:

$$
a(t) = a_0\Big[1 + H_0(t - t_0) - \frac{1}{2}q_0 H_0^2(t - t_0)^2 + \dots\Big],
\tag{1.41}
$$

where $q_0 = q(t_0)$.

The deceleration parameter is another crucial cosmological parameter, which expresses the second-order derivative w.r.t. time of the scale factor.

---

[2]Galaxy clusters contain a large quantity of ionised gas, which emits in the X-ray band through bremsstrahlung. The interaction between CMB photons with energetic electrons belonging to this gas can modify the spectral distribution of the CMB, causing a decrement in the CMB brightness at low frequencies and an increment at high frequencies (Sunyaev & Zeldovich, 1972).

It is defined with the minus sign, therefore it has a positive value when the Universe undergoes a decelerated expansion ($\ddot{a} < 0$).

The Taylor expansion of Eq. (1.39) is physics-free, which means that it is simply a mathematical expression of how the Universe expands at a times $t \sim t_0$, and says nothing about the forces that drive this expansion. The two cosmological parameters H_0 and q_0, therefore, are only descriptive of the kinematics of the Universe, regardless of the underlying cosmological model that describes the expansion of the Universe.

1.8 The Friedmann equations

In their well-known works, Einstein (1917) and de Sitter (1917), starting from General Relativity, obtained two possibile types of the Universe. Einstein obtains the so-called *cylindrical spacetime*, in which space has a constant, time-independent curvature, where the radius of curvature is connected to the total mass of matter present in space, while de Sitter obtained a spherical spacetime in which not only space is curved, but the whole spacetime has its own constant curvature.

It was already known from 1917 that the EFE (Eq. (1.4)) were not compatible with the hypothesis of a static Universe. Indeed, starting from them, choosing the FLRW metric and a perfect-fluid stress-energy tensor, one finds only two independent equations, called the *first* and the *second* Friedmann equations, which can be expressed as:

$$
\begin{aligned}
\text{I}: \quad & \ddot{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right)a \\
\text{II}: \quad & \dot{a}^2 + \kappa c^2 = \frac{8\pi}{3}G\rho a^2 \,.
\end{aligned}
\tag{1.42}
$$

The second Friedmann equation can be obtained from the first one the adiabatic condition, which in the cosmological context takes the meaning of an adiabatic expansion of the Universe:

$$
\mathrm{d}U = -p\mathrm{d}V \,,
\tag{1.43}
$$

where U is the internal energy of the Universe. This latter condition can be also expressed as:

$$
\mathrm{d}(\rho c^2 a^3) = -p\mathrm{d}a^3 \,.
\tag{1.44}
$$

In order to have a static Universe, one has to require that $\ddot{a} = 0$. From the first Friedmann equation, this translates into

$$
-\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right)a = 0 \,,
$$

which leads to:

$$\rho c^2 = -3p \,. \tag{1.45}$$

Therefore, to have a static, perfect-fluid, Universe described by the FLRW metric, the density of energy ρc^2 and the pressure must have discordant sign. This inevitably leads to one of the two having a negative signs, which is impossible, from the positive definition of these two physical quantities.

Considering $\rho + \frac{3p}{c^2} > 0$ what follows is a negative acceleration of the scale factor. Therefore the model based only on the classical EFE predicts an unstable Universe that collapses under its own weight, and this is in contrast with the astrophysical observations.

1.8.1   The cosmological constant

In order to recover a static Universe following from his equations, Einstein in 1917 introduced a constant Λ, called the *cosmological constant*, in Eq. (1.4):

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \,. \tag{1.46}$$

With a suitable choice of Λ, which has to be small enough not to change the planetary motion in our Solar System, one can obtain a static model for our Universe.

The effect of this additional term in the EFE is *repulsive*. From a Newtonian point of view, the cosmological constant leads to a repulsive force on a test particle proportional to its distance.

Nowadays, we know that our Universe is expanding, therefore the cosmological constant should not be necessary. However, from the observations of the flux of distant type Ia supernovae (SNIa) (Riess et al., 1998; Perlmutter et al., 1999) we know that the expansion of our Universe is now accelerated. Therefore we need again the cosmological constant in EFE, which takes different meanings depending on its position in the Einstein equations:

- LHS (left hand side): interpretation of Λ as a geometrical modification of gravity, as described by GR;

- RHS (right hand side): interpretation of Λ as an additional energy component, called *Dark Energy* (DE, hereafter), responsible for the accelerated expansion of our Universe.

The first physicist who studied the dynamical solutions for the EFE with the addition of Λ was Alexander Friedmann, which described the expansion or contraction of an isotropic homogenous Universe as a function of time (Friedman, 1922).

Considering Λ on the RHS, we can re-write Eq. (1.46) as:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}\tilde{T}_{\mu\nu}\,, \qquad (1.47)$$

where now the *effective* energy-momentum tensor is

$$\tilde{T}_{\mu\nu} = -\tilde{p}g_{\mu\nu} + (\tilde{\rho}c^2 + \tilde{p})u_\mu u_\nu\,, \qquad (1.48)$$

and the effective density and pressure are:

$$\begin{cases} \tilde{p} = p - \dfrac{\Lambda c^4}{8\pi G} = p + p_\Lambda; \\[4mm] \tilde{\rho} = \rho + \dfrac{\Lambda c^2}{8\pi G} = \rho + \rho_\Lambda\,. \end{cases} \qquad (1.49)$$

Having EFE expressed as (1.47), the Friedmann equations becomes:

$$\ddot{a} = -\frac{4\pi G}{3}(\tilde{\rho} + \frac{3\tilde{p}}{c^2})a \qquad (1.50)$$

$$\dot{a}^2 + \kappa c^2 = \frac{8\pi G}{3}\tilde{\rho}a^2\,. \qquad (1.51)$$

This is the so-called *Einstein Universe* model. Indeed, the addition of Λ permits a static Universe:

$$\begin{cases} \dot{a} = 0 \\[4mm] \ddot{a} = 0 \end{cases} \iff \Lambda = \frac{\kappa}{a^2}\,.$$

However this model puts a restriction on the curvature of the spacetime. Considering the ordinary total density ρ and considering the expression found for Λ one finds:

$$\rho = \tilde{\rho} - \rho_\Lambda = \frac{\kappa c^2}{4\pi G a^2}\,. \qquad (1.52)$$

In order to have a positive total ordinary density:

$$\rho > 0 \iff \kappa > 0 \Rightarrow \kappa = +1\,. \qquad (1.53)$$

Therefore, in the Einstein model the curvature is necessarily positive. The latter fact poses another important restriction on the value of the cosmological constant. Recovering a from Eq. (1.52) one finds:

$$\Lambda = \frac{4\pi G \rho}{c^2}\,. \qquad (1.54)$$

1.8.2 The density parameter

From the second Friedmann equation (1.42), we can derive the curvature parameter κ:

$$\frac{\kappa}{a^2} = \frac{1}{c^2} H(t)^2 \left(\frac{\rho}{\rho_c} - 1 \right). \tag{1.55}$$

Here we have defined the *critical mass density* ρ_c as:

$$\rho_c(t) \equiv \frac{3H(t)^2}{8\pi G}. \tag{1.56}$$

The latter corresponds to the value of the mass density required to have a flat Universe ($\kappa = 0$), and can be used to define the dimensionless *density parameter*:

$$\Omega(t) \equiv \frac{\rho(t)}{\rho_c(t)}. \tag{1.57}$$

From this definition, considering Eq. (1.55), we find that the curvature κ is strictly related to the density parameter, therefore the latter determines the curvature of the Universe:

$$\begin{cases} \kappa = -1 \iff \Omega(t) < 1; \\ \kappa = 0 \iff \Omega(t) = 1; \\ \kappa = +1 \iff \Omega(t) > 1. \end{cases} \tag{1.58}$$

The value of the critical density calculated today ($t = t_0$) depends on the Hubble constant H_0. The value for $\rho_{c,0} \equiv \rho_c(t_0)$ is:

$$\rho_{c,0} \simeq 1.9 \cdot 10^{-29} \, h^2 \, g \, cm^{-3}. \tag{1.59}$$

If the Universe has today a density greater than this value, which means $\Omega(t) > 1$, then it is positively curved. If $\rho_0 < \rho_{c,0}$ then the curvature is negative and the space is hyperbolic.

From the definition of Ω, the second Friedmann equation can be rewritten as

$$1 - \Omega(t) = -\frac{\kappa c^2}{a^2(t) H(t)^2}. \tag{1.60}$$

Note that the right hand side of this equation cannot change its sign during the expansion of the Universe, thus neither can the left hand side. This is a fundamental relation, indeed it follows that a Universe governed by the Friedmann equations cannot change its *geometry* during its evolution. Therefore, if we do observe a flat Universe, it is necessary that it was flat also at very high redshifts with huge precision. Surprisingly, this is our case,

in fact multiple observations show that $|\Omega_0 - 1| < 0.01$ (Komatsu et al., 2009, 2011; Planck Collaboration et al., 2020). This is the so-called *flatness problem*, which is a fine-tuning problem. Among an infinite spectrum of values for the density parameter of our Universe, it is extremely surprising to find it so "finely tuned" to the unity. The age of the Universe is so large that any tiny departure from the perfect flatness at high redshifts would have produced a huge curvature (open or closed), which would not have allowed the existence of the large-scale structures that we observe today, together with our own existence.

From the conservation of the sign of the density parameter, the conditions of Eq. (1.58) can be expressed in terms of its current value $\Omega_0 \equiv \Omega(t = t_0)$. Consequently, it is of extreme importance having increasingly precise measurements of this parameter, from which we can understand the overall curvature of our Universe.

1.9 The general Friedmann model

The Friedmann equations, along with the adiabaticity relation expressed in Eq. (1.44), describe a perfect fluid and allow to calculate the time evolution of $a(t)$, as well as $\rho(t)$ and $p(t)$. In order to solve these equations it is necessary to introduce an equation of state (EoS) for the different fluids composing the Universe. The two hypotheses we had set are the CP and the perfect fluid approximation, thus the pressure of the fluid can only be isotropic.

Perfect fluids, at rest, are completely described in terms of their energy density ρc^2 and pressure p, from which we can define a general equation of state as:

$$p = w\rho c^2 \,, \tag{1.61}$$

where w is related to the sound speed:

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

Here the subscript S indicates that we are evaluating c_s at constant entropy.

The parameter w lies in the so-called *Zel'dovich interval*:

$$0 \leq w < 1 \,, \tag{1.62}$$

in which c_s maintains a physical sense.

The value of w changes depending on the different components that fill the Universe. The latter can be divided in two families: relativistic and non-relativistic components.

The first is the *dust* component, and corresponds to a non-relativistic perfect fluid of matter. Non-relativistic means that the random thermal motions of the particles have peculiar velocities which are tiny compared to the speed of light. Such non-relativistic gas is described by the perfect fluid law:

$$p = \frac{\rho}{\mu} k_B T \,,$$

where ρ is the mass density of the gas, $k_B = 1.38 \cdot 10^{-16}$ erg K^{-1} is the Boltzmann constant, T the temperature and μ is the mean mass of the gas particles.

Let us consider a perfect gas composed of particles having mass m_p. The thermal energy associated to a single particle is $k_B T$, while its energy, associated to its rest mass, is $m_p c^2$, which is usually is much larger than the thermal energy. Thus, the pressure for a such perfect gas is

$$p = \frac{k_B T}{m_p c^2} \rho c^2 \approx 0 \,.$$

Hence it is described by an EoS with $w = 0$.

The second case is the *radiation* component, representing an ultra-relativistic non-degenerate fluid of massless particles. Although the particles are massless, they exhert a pressure due to their momentum. The equation of state of photons, or of any other relativistic gas is

$$p_r = \frac{1}{3} \rho_r c^2 \Longrightarrow w = \frac{1}{3} \,,$$

where p_r and ρ_r are, respectively, the radiative pressure and the radiative density.

We will refer to the component of the Universe which consists of non-relativistic particles (and hence has $w = 0$) as "matter", while we will refer to the component which consists of photons and other relativistic particles (so that $w = 1/3$) as "radiation".

The first Friedmann equation, considering $p = w\rho c^2$, can be rewritten as

$$\ddot{a} = -\frac{4\pi}{G} \rho (1 + 3w) a \,, \tag{1.63}$$

hence a component having $w < -1/3$ will provide a positive acceleration. A component of the Universe having $w < -1/3$ is generically called "dark energy", and may be associated, as we already mentioned, to the cosmological constant Λ. From the first of Eq. (1.49) we see that the equation of state associated with the Λ component is:

$$p_\Lambda = -\rho_\Lambda c^2 \Longrightarrow w = -1 \,, \tag{1.64}$$

23

where $\rho_\Lambda = \frac{\Lambda c^2}{8\pi G}$.

From the combination of the *adiabatic condition* and the EoS one finds:

$$\rho_w \propto a^{-3(1+w)} \propto (1+z)^{3(1+w)}, \tag{1.65}$$

which describes how the densities of the different components of the Universe vary with cosmic time.

For the dust component $w = 0$, therefore:

$$\rho_m = \rho_{0,m}(1+z)^3, \tag{1.66}$$

while for the radiation component $w = 1/3$:

$$\rho_R = \rho_{0,R}(1+z)^4. \tag{1.67}$$

The Λ-density component is constant in time, as one can clearly see by replacing $w = -1$ in (1.65):

$$\rho_\Lambda = \rho_{0,\Lambda}(1+z)^0 = \rho_{0,\Lambda}. \tag{1.68}$$

As we have just seen, the density components evolve differently with cosmic time. This implies that in the different cosmic epochs there could be a dominating component. Furthermore, by equalizing the different ρ_w components, one can compute the specific redshift at which they were balanced. As an example, at very high redshift the radiation component dominates over matter and Λ, while nowadays ρ_r is negligible and the dominant component is Λ.

Recalling the density parameter, now we can define a total density parameter for a multi-component Universe as:

$$\Omega_{TOT}(t) = \Omega_m(t) + \Omega_r(t) + \Omega_\Lambda(t), \tag{1.69}$$

where:

$$\Omega_m(t) = \frac{8\pi G}{3H^2(t)}\rho_m(t), \quad \Omega_r(t) = \frac{8\pi G}{3H^2(t)}\rho_r(t), \tag{1.70}$$

and:

$$\Omega_\Lambda(t) = \frac{8\pi G}{3H^2(t)}\rho_\Lambda(t) = \frac{\Lambda c^2}{3H^2(t)} \tag{1.71}$$

are the density parameters related to the different components.

Now 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:

$$H^2(z) = H_0^2(1+z)^2 \left(1 - \sum_i \Omega_{0,w_i} + \sum_i \Omega_{0,w_i}(1+z)^{1+3w_i} \right), \tag{1.72}$$

where $H(z)$ is the Hubble parameter at a generic redshift, $\sum_i \Omega_{0,w_i} = \Omega_{TOT,0}$ is the sum of all the different components density parameters with corresponding parameters w_i. The quantity $1 - \sum_i \Omega_{0,w_i} = \Omega_{0,\kappa}$ is related to the curvature of the Universe and it is called *curvature density parameter*.

Now, let us rewrite the first Friedmann equation, recalling the EoS $p = w\rho c^2$:

$$\ddot{a} = -\frac{4\pi G}{3}(1+3w)a. \tag{1.73}$$

This equation implies that, if w lies in the Zel'dovich interval, the acceleration of the scale factor is a negative function of time, e.g. $\ddot{a}(t) < 0$.

Additionally, we can write Eq. (1.73) for a multi-component Universe, and find an interesting relation between the deceleration parameter defined in Section 1.7 and the density parameters. Indeed, by substituting in it the critical density of Eq. (1.56) we can write

$$\frac{\ddot{a}}{a} = -\frac{1}{2}H^2 \sum_i \left(\Omega_i + 3w_i\right), \tag{1.74}$$

which can be rewritten in terms of $q \equiv -\dfrac{\ddot{a}a}{\dot{a}} \equiv \dfrac{\ddot{a}}{H^2 a}$, by dividing each side of the last equation by H^2:

$$q(z) = \frac{1}{2}\sum_i \left(\Omega_i + 3w_i\right). \tag{1.75}$$

Now, considering the values of w_i for the different components, we can make Eq. (1.75) an explicit relation between q and Ω_i:

$$q = \Omega_r(z) + \frac{1}{2}\Omega_m(z) - \Omega_\Lambda(z). \tag{1.76}$$

From the latter equation it is clear that ordinary cosmological components, i.e. matter ($w = 0$) and radiation ($w = 1/3$), can only produce a decelerated expansion of the Universe ($q(t) > 0$). Considering the last equation for $t = t_0$, we find that the Universe will currently be accelerating outward only if

$$\Omega_{0,\Lambda} > \frac{1}{2}\Omega_{0,m} + \Omega_{0,r}. \tag{1.77}$$

Finally, from the Hubble law we can conclude that $a(t)$ grows monotonically, given the positivity of $H(t)$. Furthermore, from the concavity of the second order derivative of a, this implies that there is an instant at which $a(t)$ is equal to zero, at some finite time in the past. We call this event *Big Bang* (BB). Indicating the BB with $t = 0$, we can easily see that at this instant of time the density diverges:

$$\lim_{t \to 0} \rho(t) = \lim_{t \to 0} \left(\frac{a_0}{a}\right)^{-3(1+w)} \to \infty. \tag{1.78}$$

1.9.1 The Einstein-de Sitter model

Let us now focus on the solutions of the Friedmann equations, which will depend on the curvature of the Universe. We will consider first the case of a flat mono-component Universe ($\kappa = 0$, $\Omega_0 = 1$), the so-called *Einstein-de Sitter model* (EdS). As we already said, the different evolution with time of the density of the various components is such to have entire cosmic epochs in which only one component is dominant over the others. Therefore, we can assume that the Universe in the different epochs is entirely composed by only one type of fluid. The actual epoch represents an exception, because we observe two significant matter-energy components, i.e. the DM and DE components. This assumption allows us to study the solutions of the Friedmann equations as component-dependent.

Considering a flat Universe, the second Friedmann equation (1.72) reduces to:

$$H(z) = H_0(1 + z)^{\frac{3(1+w)}{2}} \, , \tag{1.79}$$

which can be written in terms of the scale factor and it can be integrated to obtain its evolution as a function of time:

$$a(t) = a_0 \left(\frac{t}{t_0} \right)^{\frac{2}{3(1+w)}} . \tag{1.80}$$

Therefore, the expansion in a flat Universe, both for $w = 0$ and $w = 1/3$, is eternal.

We can also obtain, from the last equation, the relation between time and redshift:

$$t = t_0(1 + z)^{-\frac{3}{2}(1+w)} . \tag{1.81}$$

From Eqs. (1.79) and (1.81) we find:

$$H(t) = \frac{2}{3(1 + w)t} . \tag{1.82}$$

Computing this last equation for $t = t_0$, we can derive the relation between the *age of the Universe* and the Hubble constant:

$$t_0 = \frac{2}{3(1 + w)H_0} . \tag{1.83}$$

Thus, in an EdS model, the age of the Universe is smaller than $\tau_H \equiv 1/H_0$.

Lastly we can obtain the relation between the density and time for an EdS Universe from the previous relations:

$$\rho(t) = \rho_{0,w} \left(\frac{a_0}{a} \right)^{3(1+w)} = \frac{1}{6\pi G(1 + w)^2} \frac{1}{t^2} . \tag{1.84}$$

A relevant fact about the EdS model is that for a radiative Universe, the expansion is more decelerated because of the presence of the radiation pressure.

1.9.2 Open and closed models

Let us consider now the case of curved mono-component Universes ($\kappa \neq 0, \Omega_{0,w} \neq 1$). As we will discuss in more details in Section 3.2, the formation of galaxy clusters and cosmic voids, starting from small density perturbations, can be modelled as the evolution of closed and open Universes, respectively, growing inside a flat Einstein-de Sitter Universe. Therefore, the equations at the base of these models are crucial in order to understand the evolution of large-scale structures, on which our Thesis work is based.

Considering again the second Friedmann equation as expressed in (1.72), the curvature density parameter is constant over time while the second term is not. Therefore it exists a *critical* value a^* such that:

$$|1 - \Omega_{0,w}| = \Omega_{0,w}\left(\frac{a_0}{a^*}\right)^{1+3w} , \tag{1.85}$$

or, in terms of redshift:

$$|1 - \Omega_{0,w}| = \Omega_{0,w}\left(1 + z^*\right)^{1+3w} . \tag{1.86}$$

Given the values for w, for high redshifts $z \gg z^*$ we can neglect the curvature term, and so Eq. (1.72) reduces to:

$$H(z) = H_0 \sqrt{\Omega_{0,w}}(1 + z)^{\frac{3(1+w)}{2}} , \tag{1.87}$$

which differs from the relation found for the EdS model by the $\sqrt{\Omega_{0,w}}$ factor. From this result we can assume that our Universe is well described by an EdS model, which carries no curvature, in early cosmological epochs.

Considering $z \ll z^*$, the potential curvature of the Universe prevails, showing the underlying geometry. Moreover the sign of the curvature term $1 - \Omega_{0,w}$ depends on the values of $\Omega_{0,w}$, so we will consider open and closed models separately.

Open model for the Universe

Let us start with the open model $\kappa = -1$, $\Omega_{0,w} < 1$. Recalling Eq. (1.72), we have that the term between square brackets never vanishes. Moreover we know from observations that $\dot{a}_0 > 0$, hence $\dot{a}(t) > 0$, $\forall t$.

For $z \ll z^*$ we can neglect the second term in the square brackets of Eq. (1.72), hence we find the following expression for $H(z)$:

$$H(z) = H_0(1+z)(1-\Omega_{0,w})^{1/2}.$$

We can write this last question in terms of a:

$$\dot{a}(t) = a_0 H_0(1-\Omega_{0,w})^{1/2}.$$

By integration we can find that $a(t)$ has an asymptotically linear growth with time:

$$a(t) = a_0 H_0(1-\Omega_{0,w})^{1/2}t \Rightarrow a(t) \propto t. \tag{1.88}$$

As a consequence, the Hubble parameter decreases proportionally with the inverse of time:

$$H(t) \propto t^{-1}.$$

Closed model for the Universe

Consider now the case of a closed model ($\kappa = 1, \Omega_{0,w} > 1$). Now the curvature term $1 - \Omega_{0,w}$ can be null, hence there exists a time $t = t_{max}$ such that $\dot{a}(t_{max}) = 0$. In other words, in this type of model, the Universe reaches a maximum size $a_{max} = a(t_{max})$ given by:

$$a_{max} = a_0 \Big(\frac{\Omega_{0,w}}{\Omega_{0,w} - 1}\Big)^{1/1+3w}. \tag{1.89}$$

When $a = a_{max}$ the density reaches its minimum value:

$$\rho_{min} = \rho(a_{max}) = \rho_0 \Big(\frac{\Omega_{0,w} - 1}{\Omega_{0,w}}\Big)^{\frac{3(1+w)}{1+3w}}. \tag{1.90}$$

For $t > t_{max}$ the scale factor decreases at the same rate of its increment, therefore at a time $t_f = 2t_{max}$ the Universe will collapse in another singularity, the *Big Crunch*.

Once we have set up the models for an open or closed mono-component Universe, we can find a solution for the different values of the parameter w, that are dust and radiative curved Universes.

1.9.3 Dust Universe model

Models with $w = 0$ have an exact analytic solution, even for the case where $\Omega_w \neq 1$.

In this case Eq. (1.72) can be written as:

$$\left(\frac{\dot{a}}{a_0}\right)^2 = H_0^2\left(1 - \Omega_{0,m} + \Omega_{0,m}\frac{a_0}{a}\right). \tag{1.91}$$

Let us consider first *open* models ($\Omega_{0,m} < 1$). These models have a solution in the parametric form:

$$\begin{cases} a(\psi) = \dfrac{a_0}{2}\dfrac{\Omega_{0,m}}{1 - \Omega_{0,m}}(\cosh\psi - 1) \\ t(\psi) = \dfrac{1}{2H_0}\dfrac{\Omega_{0,m}}{(1 - \Omega_{0,m})^{3/2}}(\sinh\psi - \psi), \end{cases}$$

where $0 \le \psi < \infty$.

Recovering a_0 from the first parametric equation and substituting it in the second one, we can obtain the expression for the age of the Universe t_0:

$$t_0 = \frac{1}{2H_0}\frac{\Omega_{0,m}}{(1 - \Omega_{0,m})^{3/2}}\left[\frac{2}{\Omega_{0,m}}(1 - \Omega_{0,m})^{1/2} - \cosh\left(\frac{2}{\Omega_{0,m}} - 1\right)\right] > \frac{2}{3H_0}.$$

Therefore the age of the Universe for an open dust model is greater than for the EdS model for a fixed value of H_0.

The last equation for t_0 has the following limit for $\Omega_{0,m} \ll 1$:

$$t_0 \simeq (1 + \Omega_{0,m}\ln\Omega_{0,m})\frac{1}{H_0}.$$

Now let us consider the other case where $\Omega_{0,m} > 1$. For these models the parametric solutions are:

$$\begin{cases} a(\theta) = \dfrac{a_0}{2}\dfrac{\Omega_{0,m}}{\Omega_{0,m} - 1}(1 - \cos\theta) \\ t(\theta) = \dfrac{1}{2H_0}\dfrac{\Omega_{0,m}}{(\Omega_{0,m} - 1)^{3/2}}(\theta - \sin\theta), \end{cases} \tag{1.92}$$

where now $0 \le \theta < 2\pi$.

In this case the age of the Universe is smaller than for the EdS model for $w = 0$:

$$t_0 < \frac{2}{3H_0}.$$

Deriving the scale factor with respect to the parameter θ, one finds that $a(\theta)$ grows for $0 \le \theta < \pi$ and decreases for $\pi < \theta < 2\pi$. The maximum size reached by the Universe in this model is:

$$a_{max} = a(\pi) = a_0\frac{\Omega_{0,m}}{\Omega_{0,m} - 1}. \tag{1.93}$$

This status is reached at $t = t_{max} = t(\pi)$:

$$t_{max} = \frac{\pi}{2H_0} \frac{\Omega_{0,m}}{(\Omega_{0,m} - 1)^{3/2}} \ .$$ (1.94)
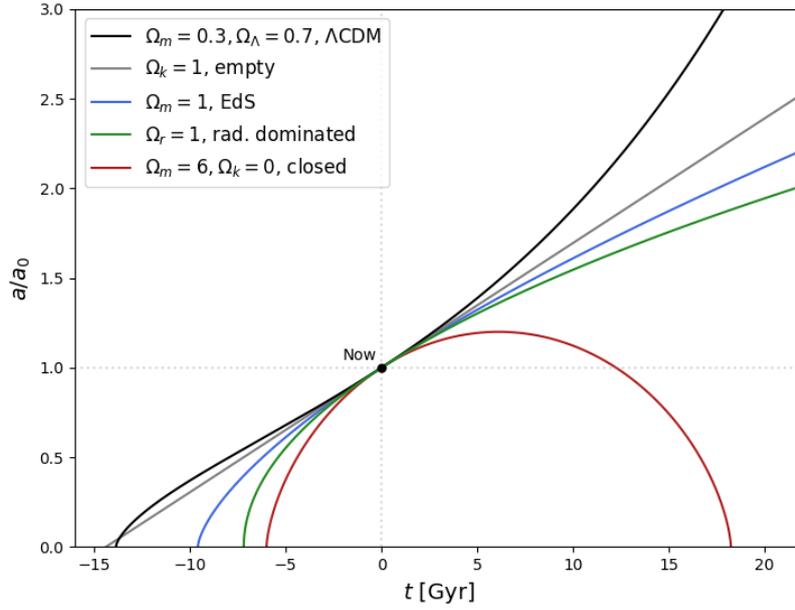
Figure 1.2: Scale factor as a function of time for different models of the Universe. The black line represents the scale factor evolution for the Standard Cosmological Model (ΛCDM), which is described in Section 1.10. The gray line represents an empty flat Universe (with $\Omega_m = 0$), the blue line shows the evolution for a flat Einstein-de Sitter Universe, while the green and the red lines represent the evolution of a radiation-dominated Universe and a closed mono-component Universe, respectively.

1.9.4 Radiative Universe model

Going back to the radiative models, for which $w = 1/3$, Eq. (1.72) has the form:

$$\left(\frac{\dot{a}}{a_0}\right)^2 = H_0^2 \left[1 - \Omega_{0,r} + \Omega_{0,r} \left(\frac{a_0}{a}\right)^2\right] ,$$ (1.95)

which has the following solution:

$$a(t) = a_0(2H_0\Omega_{0,r}^{1/2}t)^{1/2}\left(1 + \frac{1 - \Omega_{0,r}}{2\Omega_{0,r}^{1/2}}H_0t\right)^{1/2}.$$

It is possible to show that, in open models, the following relation holds for the present cosmic time t_0:

$$t_0 > \frac{1}{2H_0},$$

where $t_0 = \frac{1}{2H_0}$ is the age of the Universe for an EdS model with $w = 1/3$. Lastly, closed radiative models predict:

$$t_0 < \frac{1}{2H_0}.$$

Finally, in Figure 1.2 it is shown the evolution of the scale factor for some models of Universe described in the previous Sections. Here the different models are normalised in such a way that today they have the same value for $a(t)$ and H_0. The intersections of the different solid lines with the x-axis show how the age of the Universe t_0 changes among the different cosmological model considered.

1.10 The Standard Cosmological Model

From the beginning of the 21$^{\text{st}}$ century, the commonly accepted model describing our Universe is the so-called ΛCDM. The basic idea for it lies in the *Hot Big Bang model* (HBB), which is based on the Cosmological Principle and the Friedmann equations. The evolution of our Universe is described as a thermal history, which means that its expansion is tightly related to its temperature: going back in time the Universe was much hotter than nowadays. In particular its temperature at the present time is $T = 2.725 \pm 0.005$ K, a value that corresponds to the Cosmic Microwave Background (CMB) temperature, the relic radiation from the surface of last scattering happened only about $400,000$ years after the Big Bang. Before the last scattering, the Universe was filled by a hot plasma composed by protons and electrons fully ionised. In this context, electromagnetic radiation was continuously scattered by matter, so that the Universe was completely opaque and in thermal equilibrium. Thanks to its expansion, its temperature falls down until electrons recombine with protons, at $z \simeq 1100$. When this process completes, photons began to freely propagate into the Universe. We observe this radiation as a background redshifted to the microwave band, distributed in a spherical
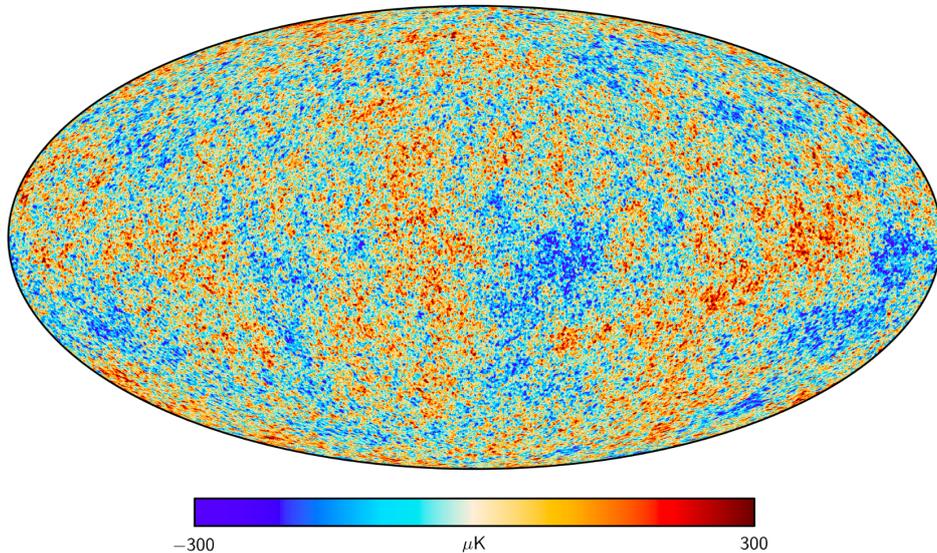
Figure 1.3: All-sky map of the CMB temperature fluctuations, as measured by the Planck satellite. The range of temperature is very small, going from -300 μK to 300 μK, highlighting a strong thermal equilibrium. Credits to Planck Collaboration 2015 (Planck Collaboration et al., 2016a).

surface around us (see Figure 1.3). The most powerful prediction of the HBB model is the existence of the CMB radiation.

The ΛCDM model establishes that our Universe is flat, and its major mass-energy components at the present time are the Cosmological Constant (Λ), necessary to explain the observed accelerated expansion of the Universe (Riess et al., 1998; Perlmutter et al., 1999), and a dust component called cold dark matter (CDM). The other mass-energy components of the Universe are the baryonic matter, which is non-relativistic, and the relativistic components such as the radiation field. The contribution of the latter is negligible nowadays and its density parameter, which can be estimated from the measure of the CMB temperature, has a value of

$$\Omega_{r,0} \approx 10^{-5} \,. \tag{1.96}$$

The substantial difference between baryonic and dark matter is their interactions. Indeed both of them interact gravitationally but only baryonic matter interacts with the observable electromagnetic radiation, hence the "dark" attribute for the CDM component. The first scientist proposing the existence of a different kind of matter was Fritz Zwicky, in order to explain the dynamics of galaxies in the Coma cluster (Zwicky, 1937). From the 1970s, astronomers

and cosmologists began to build an enthralling body of evidence for this elusive component, based on a variety of observations, including temperature anistropies of the CMB, baryonic acoustic oscillations (BAO)[3] (Anderson et al., 2014; Ross et al., 2015), SNIa, gravitational lensing of galaxy clusters and rotation curves of disk galaxies (Fuchs, 2001; Bosma, 2000). The *standard model* of particle physics contains no suitable particle to explain these observations and nowadays the proposed candidates for DM span 90 orders of magnitude in mass, ranging from ultralight bosons (often referred as "fuzzy DM") to massive primordial black holes.

The class of dark-matter candidates that has attracted the most attention over the past four decades is *weakly interacting massive particles* (WIMPs), which are hypotetical particles on the weak-interaction mass scale (between 10 GeV and 1 TeV) and MACHOs (Alcock et al. (1997); Brandt (2017)), which stands for *massive astrophysical compact halo objects*, composed of ordinary matter but which do not emits significantly light such as neutron stars, brown and white dwarfs.

There are multiple ways to evaluate the matter density parameter of the Universe, $\Omega_{m,0}$, such as measuring the radial velocities of galaxies in a galaxy cluster, or by measuring the amplitude of gravitational lensing on very distant galaxies caused by the presence of a massive objects interposed between them and the observer. Also the clustering properties of large-scale structures (LSS), such as galaxy clusters and galaxies, together with their abundances are strongly related to this parameter. All these measurements lead to the same conclusion concerning the total matter density parameter:

$$\Omega_{m,0} \approx 0.3 \,. \tag{1.97}$$

From the primordial nucleosynthesis and from the study of the CMB radiation one can estimate the density fraction of *baryonic* matter, which is approximately

$$\Omega_{b,0} \approx 0.04 - 0.05 \,. \tag{1.98}$$

Hence only a tiny fraction of the total mass is in form of baryonic matter.

The property of dark matter to be "cold" means that at the time of its decoupling was non-relativistic, and this leads to a particular structure

---

[3]BAO refers to acoustic waves propagating before the recombination, which then froze when the latter took place, leaving an imprint in the matter distribution on large scales. We observe this imprint as a peak in the two point correlation function of different mass tracers (see Section 2.4), like galaxies, which correspond to the maximum distance the acoustic waves could travel in the primordial plasma. This can be considered as a standard ruler, and thus can provide powerful constraints on the expansion rate of the Universe and allows to investigate the nature of DE.

formation scenario, called *bottom-up*, in which smaller objects formed first and then they merged forming more massive structures (see Section 2.3.1 for a detailed description).

Regarding the cosmological constant Λ, as already mentioned in Section 1.8.1, the physical component associated with it is called *Dark Energy* (DE), described by an equation of state parameter $w = -1$. Its nature is still unknown and understanding its properties and origin is one of the principal challenges in modern physics.

If the Universe is flat, then the sum of the density parameters of all the components must be equal to 1. However, if one considers a Universe composed only by matter and radiation, one gets:

$$\Omega_{m,0} + \Omega_{r,0} \approx 0.3 \,. \tag{1.99}$$

The remaining 0.7 is the density fraction of the DE component, the most dominant in the present Universe. This result comes not only by the flatness of the Universe, but also from the already cited measures (Riess et al. (1998); Perlmutter et al. (1999)). They estimated the value of the deceleration parameter using SNIa as standard candles, obtaining $q_0 \approx -0.55$, which implies an accelerated espansion for our Universe. By replacing the known values for Ω_m and q in Eq. (1.76), neglecting the contribute of Ω_r, again one obtains $\Omega_\Lambda = 0.7$.

DE leaves imprints on cosmological observations. It modifies the geometry of the Universe, increasing distances and volumes over time via the accelerated expansion, and it slows down the growth of cosmic structures. However, these effects can be mimicked by the variation of other cosmological parameters, including the DM density and curvature, or other physical models and systematics that are degenerate within a single probe. Consequently, measuring DE properties requires a *combination of cosmological probes* that are sensitive to both classes of effects to break these parameter and model degeneracies.

After the introduction of a theoretical background, it is necessary to parametrize the ΛCDM model, which is indeed described by six fundamental parameters:

- Ω_m: total matter density parameter;

- Ω_b: baryonic matter density parameter;

- A_s: normalization of the power spectrum;

- H_0: Hubble constant;

- n_s: spectral index of the primordial power spectrum;

- τ : reionization optical depth.

The most strong constraints on the latter derive from the analysis of the CMB power spectrum from the latest release of the Plack legacy (Planck Collaboration et al., 2020) in combination with lensing measurements, although other cosmological probes are becoming competitive as well. The values for these six fundamental parameters, as reported in Planck Collaboration et al. (2020), are $\Omega_m = 0.3153 \pm 0.0073$, $\Omega_b h^2 = 0.02237 \pm 0.00015$, $\ln(10^{10} A_s) = 3.044 \pm 0.014$, $H_0 = 67.36 \pm 0.54$ km s^{-1} Mpc^{-1}, $n_s = 0.9649 \pm 0.0042$ and $\tau = 0.0544 \pm 0.0073$.

1.10.1 Statistical tensions in the ΛCDM model

The ΛCDM model successfully describes a wide range of observations, going from the abundance of elements from primordial nucleosynthesis up to the present-day accelerated expansion. The astonishing precision coming from both modern observations and simulations, has led to the rise of few statistically significant tensions on cosmological parameters between different probes. The latter can be divided into two groups, *early* and *late* type probes, according to the range of redshifts to which they are applied. It is not straightforward that the same cosmological model will fit observations from widely different cosmological epochs. Although these tensions could be due to systematic errors, their confirmation during the years seems to suggest the need for new physics inside the ΛCDM framework.

Among the most effective cosmological probes we find the study of the anisotropies of CMB, made possible by the data provided by the Planck satellite (Planck Collaboration et al., 2020), successor of older missions like the Cosmic Background Explorer (COBE) and the Wilkinson Microwave Anisotropy Probe (Mather et al., 1994; Komatsu et al., 2009, 2011). Measurements from such a primordial Universe can constraint the mechanism at the base of the inflation process, and represent a prominent complementary probe of the low-redshift LSS measurements. Among the latter, an important probe is the clustering of LSS, which provides constraints on the growth of cosmic structures by looking at how much these are clustered with each other (i.e. by studying their correlation function). A powerful application of the correlation function of LSS is the measurement of position of the baryon acoustic oscillation (BAO, see e.g. Seo & Eisenstein, 2003; Blake & Glazebrook, 2003), that represents the imprint of sound waves in the pre-recombination plasma. By looking at the correlation function of cosmological objects like galaxies and clusters, it is also possible to investigate the effects

of peculiar galaxy velocities and geometrical distortions on the assumed cosmological model (see, for example Marulli et al., 2020; Alam et al., 2021). In this field, also the distortions on average cosmic void shapes have turned out to be a competitive cosmological probe (Hamaus et al., 2020; Aubert et al., 2020). We can also investigate the properties of our Universe by looking on how images of distant galaxies are distorted due to the weak lensing. It has been demonstrated its strong constraining power on both the total amount of matter in the Universe, as well on the amount of DE and amplitudes of today matter perturbations (Kaiser, 1992; Schneider, 2005; Kilbinger, 2015). Not less powerful and interesting is the study of distant of SNIa as objects of known luminosity, from which the accelerated expansion of the Universe has been inferred (Riess et al., 1998; Perlmutter et al., 1999). From the study of SNIa light-curves, characterised by a well-defined peak of luminosity, it is possible to extract fundamental constraints on the present-day expansion rate of our Universe, i.e. the Hubble constant H_0.

The most statistically significant tension is in the estimation of the latter between the analysis of the CMB power spectrum (which is an early type probe) and late-type probes, which we have just summarised. In particular, the Planck legacy found $H_0 = 67.4 \pm 0.5$ km s^{-1} Mpc^{-1} (Planck Collaboration et al., 2020), which is in tension at about 4.4σ with the 2019 SH0ES collaboration (Riess et al., 2019), which found $H_0 = 74.22 \pm 1.82$ km s^{-1} Mpc^{-1}, from the analysis of 70 Cepheids in the Large Magellanic Cloud. Other measurements, which can be seen all together in Figure 1.4, obtained slightly different values for H_0. For example, from the observation of SNIa, Dhawan et al. (2018) obtained $H_0 = 72.8 \pm 2.7$ km s^{-1} Mpc^{-1}, while the CCPH (Carnegie-Chicago Hubble Program) collaboration obtained $H_0 = 69.8 \pm 1.9$ km s^{-1} Mpc^{-1} which used Tip of the Red Giant Branch[4] (TRGB) measurements in the Large Magellanic Cloud (Freedman et al., 2019).

Another interesting tension in the ΛCDM is the discrepancy of the Planck data with weak lensing measurements and redshift surveys about the value of the derived parameter S_8, which is a combination of the total matter density parameter, Ω_m, and the amplitude of the growth of structure, σ_8, (see Section 2.4 for its formal definition), defined as:

$$S_8 = \sigma_8 \sqrt{\frac{\Omega_m}{0.3}} \, . \tag{1.100}$$

---

[4]The RGB is a stage of evolution of stars that have exhausted their centre hydrogen fuel, which turned into helium. When the helium core mass reaches 0.5 $M_\odot$, massive stars violently enhance their surface brightness (helium flash), as a consequence of the beginning of carbon production. We observe this precise stage as a tip-feature in the HR diagram, i.e. as a fixed luminosity, which can be exploited to measure distances in the local Universe.

CMB measurements, considering an underlying ΛCDM model, predicts a high value of S_8, $S_8 = 0.834\pm0.016$ (Planck Collaboration et al., 2020), which disagrees at about 2.2σ with cosmic shear measurements such as KiDS-450 ($S_8 = 0.745 \pm 0.039$) (Kuijken et al., 2015; Joudaki et al., 2017) and DES ($S_8 = 0.773^{+0.026}_{-0.020}$) (Dark Energy Survey Collaboration, 2018).

Many works have been focused on these tensions, and in recent years many theoretical solutions have been proposed, going from Modified Gravity Theories (Di Valentino et al., 2016; Planck Collaboration et al., 2016b), in which the nature of gravity changes with redshift such that the H_0 estimate from CBM can be larger, to alternative DE theories (Karwal & Kamionkowski, 2016; Poulin et al., 2019; Benevento et al., 2020).
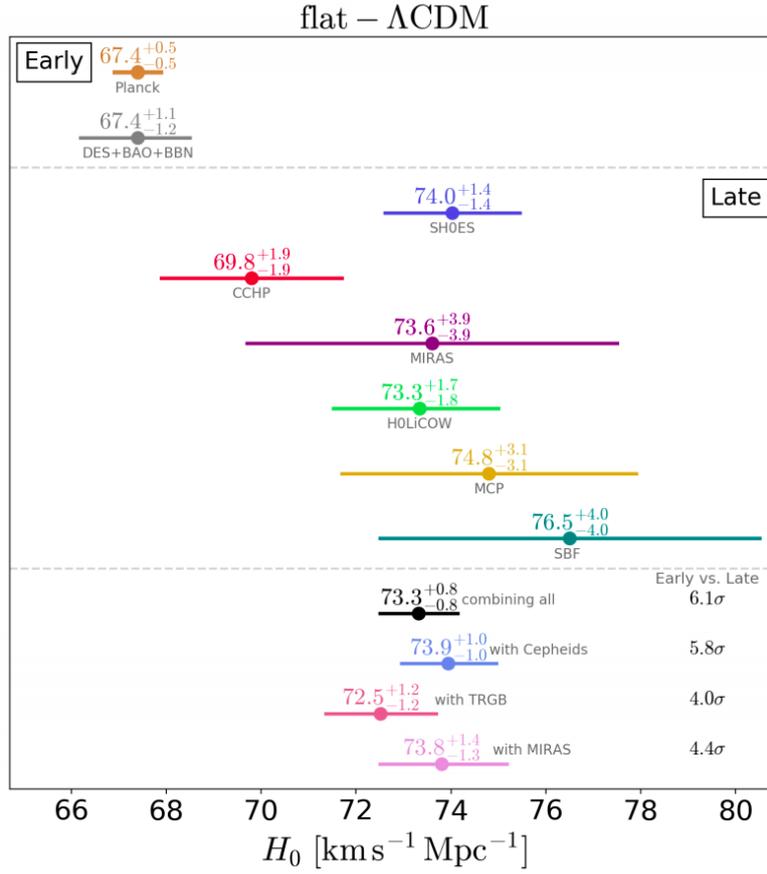
Figure 1.4: Hubble constant predictions coming from different cosmological probes. The early type measurements are represented in the top panel, while the middle panel shows the late type measurements and the bottom panel presents the combination of late type probes and the comparison with the early types. Credits to Verde et al. (2019).

Chapter 2

Formation and evolution of cosmic structures

In the standard model of large scale structure formation, small perturbations in the primordial density field evolve in time and collapse into gravitationally bounded objects, called DM haloes. In modern cosmological models like the ΛCDM, galaxy clusters as well as galaxies are supposed to grow inside these haloes (see, for example Wechsler & Tinker, 2018). One can constrain the expansion rate of the Universe, as well as the most important cosmological parameters, by observing how many collapsed structures exists at different times. Indeed the more the expansion is fast, the more the gravitational collapse will be prevented. Therefore, a theory that can predict the size of perturbations as today, starting from the very early time, is fundamental.

In this Chapter we briefly describe the origin of primordial fluctuations, which could be explained through the inflation model (Section 2.1), then we delineate the Jeans Theory, through which one can describe the evolution of density perturbations until they remain in the linear regime. We describe this theory in the classical case of a static Universe (Section 2.2.1), and also for the expanding case (Section 2.2.2). In this context, the main role of DM perturbations in the structure formation paradigm is briefly outlined in Section 2.3. In the following Section 2.4, we describe how it is possible to perform a statistical analysis of the large scale structures in the Universe, though the latter is, by definition, a single object. Finally we discuss the nonlinear evolution of perturbations in Section 2.5, with a brief description of the cosmological numerical simulations with which one can mimic the real properties of our Universe on large scales (Section 2.6).

2.1 Perturbations from Inflation

In the currently accepted models, primordial perturbations originated through the so-called *inflation* process, which occurred approximately 10^{-33} s after the Big Bang (Starobinskiǐ, 1979; Guth, 1981; Guth & Pi, 1982; Mukhanov et al., 1992). In inflation models, the Universe experienced a period of exponential expansion, reaching a size even greater than the distance that light could have traveled since the Big Bang. As a consequence, the geometry of the Universe is forced to be locally flat. This means that inflation does not change the overall geometry of the Universe, but only that the radius of curvature is magnified so much that we observe the Universe as locally flat to a high precision. This could explain our observations about the values for the density parameters, which show strong evidences for a flat Universe, and solves the *flatness problem* briefly summarised in Section 1.8.2.

At the time of the CMB the Universe shows perturbations on large scales that have already grown, and this is not possible since baryons were still coupled with the radiation field, which does not allow the formation of perturbations on such scales. The inflation process could also solve this problem, since every quantum inhomogeneities in the energy density field that were present before inflation would have exceed the size of the observable Universe (at that time), causing their amplitude to freeze.

From the Friedmann equations we know that the Universe evolves classically through a decelerated expansion. In order to have an accelerated one we need a negative pressure. Such a condition can be satisfy with assuming the presence of a simple scalar field ϕ, called *inflaton*, which would have permeated the Universe at the time of the inflation (Guth, 1981; Linde, 1983) Primordial fluctuations then translates into quantum fluctuations $\delta\phi$ of the inflaton field.

Furthermore, through inflation we are able to solve others problems that affects the Hot Big Bang model, such as the *horizon problem* and the *monopole problem*, which are beyond the scope of this Chapter (please see Liddle, 1999; Vázquez et al., 2018, for a comprehensive review of these problems).

2.2 Jeans Theory

From an observational point of view, our Universe appears to be quite inhomogeneous at scale of Mpc, and shows evidences of a highly nonlinear evolution with fluctuations of the order of 10^2 on scales of collapsed objects. We can determine the perturbation amplitude of the density field at the highest observable redshift by analysing the fluctuations of the temperature of

the CMB radiation ($z \simeq 1100$). The latter are, approximately:

$$\frac{\delta T}{\overline{T}} \simeq 10^{-5} \ , \tag{2.1}$$

showing a high level of homogeneity at that epoch, where $\overline{T}$ is the mean black body radiation of the CMB, measured to be $\overline{T} \approx 2.725 \pm 0.005$ K. These perturbations grow as a consequence of the attractive nature of the gravitational interaction. An excess of density in a given point in the space will pull the surrounding matter towards it creating larger overdense regions. As we will describe in this Chapter, these regions eventually form haloes of matter which, in turn, are able to collapse if larger than a given threshold density. The same matter falling onto these haloes evacuates from the underdense regions, which expand creating cosmic voids.

These phenomena can be described analytically adopting the Jeans theory as long as the perturbations remain in the *linear* regime, which means that the perturbed quantities are much smaller than their unperturbed counterparts. Indeed, considering a physical field f, this can be expressed as an unperturbed part f_0 combined with a small perturbation $\delta f = \delta f(\vec{x}, t)$. The linear regime is maintained until

$$\delta f \ll f_0 \ . \tag{2.2}$$

It is clear that the behaviour of collapsing objects is highly nonlinear, and for that we generally rely on numerical simulations, since the analytical solutions are not possible.

Therefore the aim of Jeans theory is to describe the evolution of these perturbations in linear regime, for non-relativistic matter on scales not exceeding the cosmological horizon. The latter is defined as:

$$R_H = a(t) \int_{t_{BB}}^{t} \frac{c \mathrm{d}t}{a(t')} \ , \tag{2.3}$$

where t_{BB} is the starting time for the expansion of the Universe at the Big Bang. This quantity represents the points which have had sufficient time to send information, in the form of photons or other relativistic particles, at a given time t. Moreover, R_H divides the spatial domain in two parts, and perturbations evolve differently depending on their scale. Indeed on scales larger than the cosmological horizon only the gravitational interaction is relevant, but it has to be treated by means of a fully relativistic theory. On the other hand, at sub-horizon scales, the Jeans approach provides a reliable theory of perturbations in an expanding Universe.

2.2.1 Jeans instability in a static Universe

Let us consider a self gravitating fluid, like a cloud of gas in a perfectly flat and static Universe. When the self-gravitational force balances its pressure, the cloud is said to be in hydrostatic equilibrium. When this balance fails, the cloud undergoes instability, leading to a collapse if the gravity prevails, or to an expansion if the prevailing contribute comes from pressure. One can show that there exists a scale, called the Jeans scale, at which there is a perfect (unstable) equilibrium between the gravitational force and the pressure, and it is

$$R_J = \frac{c_s}{\sqrt{G\rho}} \; , \tag{2.4}$$

where c_s is the sound speed, and ρ is the matter density of the self-gravitating cloud. Every perturbation larger than this scale generates a collapse.

Let us consider a homogeneous and isotropic background fluid with constant, time-independent matter density $\rho(\vec{x}, t) = const$. One can model the evolution of such fluid with the following set of hydrodynamic equations:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0 \qquad \text{Continuity equation} \tag{2.5a}$$

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla)\vec{v} = -\frac{1}{\rho}\nabla \rho - \nabla \phi \qquad \text{Euler equation} \tag{2.5b}$$

$$\nabla^2 \phi = 4\pi G\rho \qquad \text{Poisson equation} \tag{2.5c}$$

$$p = p(\rho, S) \qquad \text{Equation of state} \tag{2.5d}$$

$$\frac{dS}{dt} = 0 \qquad \text{Adiabatic condition} \tag{2.5e}$$

where $\vec{v}$ is the velocity vector of the fluid element, p the pressure, ϕ the gravitational potential and S the entropy. From the analysis of CMB anisotropies, it turned out that the primordial fluctuations can be described as adiabatic, hence we can consider the conservation of their entropy S, which leads to the barotropic equation of state $p = p(\rho)$. The Eqs. (2.5) are satisfied for a background solution (which we indicate with a subscript B):

$$\begin{cases} \rho_B = const \\ p_B = const \\ \phi_B = const \\ \vec{v}_B = 0 \end{cases} \tag{2.6}$$

Starting from this solution one can introduce a small perturbation δf where f is one of the variables that describe the fluid, obtaining a perturbed set of

equations, which describe the evolution of such perturbation. Defining the density contrast δ as

$$\delta(\vec{x}, t) \equiv \frac{\delta\rho(\vec{x}, t)}{\rho_B} \ , \tag{2.7}$$

one can consider the perturbations as superposition of Fourier modes, therefore as plane waves:

$$\delta = \delta_{\vec{k}} \exp(i\vec{k} \cdot \vec{r} + i\omega t) \ , \tag{2.8}$$

where $\vec{k}$ is the wavenumber associated with the Fourier mode, ω is the angular frequency (or *pulsation*) and $\delta_{\vec{k}} = \delta_{\vec{k}}(t)$ is the Fourier transform of the contrast $\delta(\vec{x}, t)$.

The system of perturbed equations can be solved for the density contrast, leading to the following differential equation:

$$\ddot{\delta}_{\vec{k}} + (k^2 c_s^2 - 4\pi G \rho_B)\delta_{\vec{k}} = 0 \ , \tag{2.9}$$

where $k = |\vec{k}|$ is the absolute value of the wavenumber. The latter equation has two different solutions:

$$\delta_{\vec{k}} \propto \exp(\pm i\omega(k)t) \ , \tag{2.10}$$

where

$$\omega(k) = \sqrt{k^2 c_s^2 - 4\pi G \rho_B} \ . \tag{2.11}$$

This relation is called the *dispersion relation*, and describes the behaviour of the different wavenumber components of the perturbation.

It is easy to see that the sign of the quantity inside the square root in Eq. (2.11) will determine the evolution of the density contrast fluctuation of Eq. (2.10). Indeed, ω can be real ($\omega \in \mathbb{R}$), leading to a perturbation that oscillates in time with constant phase velocity and constant amplitude (plane wave), or complex ($\omega \in \mathbb{C}$), leading to an exponential growth (positive solutions) or decrease (negative solutions). The transition value for k between these two conditions is the *Jeans wavenumber* k_J, for which $\omega(k_J) = 0$:

$$k_J = \frac{2}{c_s} \sqrt{\pi G \rho_B} \ . \tag{2.12}$$

With this quantity it is associated the aforementioned Jeans scale:

$$\lambda_J = \frac{2\pi}{k_J} = c_s \sqrt{\frac{\pi}{G \rho_B}} \ . \tag{2.13}$$

Perturbations having $\lambda > \lambda_J$ will lead to a gravitational collapse, associated to the growing solution of the dispersion relation.

Furthermore, we can define the *Jeans mass* as the mass of the perturbation contained in a sphere of radius λ_J:

$$M_J = \frac{4}{3}\pi\rho_B\lambda_J^3 \ . \tag{2.14}$$

Therefore the *Jeans instability criterion*, when expressed in terms of the mass, is:

$$M > M_J \ . \tag{2.15}$$

When this is true, the pressure of the cloud is no longer able to counterbalance the self-gravitational force, and the collapse begins.

2.2.2 Jeans instability in an Expanding Universe

In the case of an expanding Universe, the hydrodynamic equations need to be modified properly. First of all, the density field of the background is now a function of time $\rho_B = \rho_B(t)$. Also, the velocity of the unperturbed fluid is no longer null, on the contrary the fluid is subjected to the Hubble flow, which contrasts the growth of perturbations. Moreover we need to distinguish the evolution of perturbations having scale smaller or greater than the cosmological horizon $\lambda_H \equiv R_H$.

Perturbations on scales $\lambda > \lambda_H$

As previously said, this type of perturbations undergoes only the gravitational interaction, therefore we can consider the effect of the microphysics as negligible. This means that perturbations that are associated with scales larger than the horizon scale are not subjected to any kind of radiative pressure. Therefore, these perturbations can grow indefinitely.

To derive the rate of their growth, perturbations can be treated as a small closed Universe evolving in a background Einstein-de Sitter Universe. Therefore we have two equations, one for the background (B) Universe and one for the perturbed Universe (p), deriving directly from the second Friedmann equation:

$$H_B^2 = \frac{8\pi}{3}G\rho_B, \qquad H_p^2 = \frac{8\pi}{3}G\rho_p - \frac{c^2}{a^2} \ . \tag{2.16}$$

The perturbation Universe is totally enclosed in the background one, thus their scale factors are initially the same. One can synchronize the two Universes by equating their two Hubble parameters, finding the evolution law for the constrast δ:

$$\delta(t) = \frac{3c^2}{8\pi G}\frac{1}{\rho_B a^2} \propto \rho_B^{-1}a^{-2} \ . \tag{2.17}$$

As mentioned in Section 1.9, the different components of the density field evolves differently with time, letting a given component to dominate over the others. By equalising Eq. (1.66) and Eq. (1.67) one can find the redshift at which the matter and radiation components have the same density. This redshift, called *equivalence*, which we indicate as z_{eq}, is:

$$z_{eq} = \frac{\rho_{0,m}}{\rho_{0,R}} - 1 \simeq 3 \cdot 10^4 \ . \tag{2.18}$$

Therefore the background density will be different considering a time before or after the equivalence time, causing a different evolution for the perturbation. In particular:

$$\begin{aligned} \rho_B \propto a^{-4} &\Longrightarrow \delta(t) \propto a^2 \propto t, & z > z_{eq} \\ \rho_B \propto a^{-3} &\Longrightarrow \delta(t) \propto a \propto t^{2/3}, & z < z_{eq} \ . \end{aligned} \tag{2.19}$$

Hence perturbations on scale larger than the cosmological horizon will always grow.

Perturbations on scels $\lambda < \lambda_H$

Inside the cosmological horizon also the microphysics, described by the hydrodynamic equations, becomes important. The new set of equations is the following:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{u}) = 0 \tag{2.20a}$$

$$\frac{\partial \vec{u}}{\partial t} + (\vec{u} \cdot \nabla)\vec{u} = -\frac{1}{\rho}\nabla\rho - \nabla\phi \tag{2.20b}$$

$$\nabla^2 \phi = 4\pi G \rho \ , \tag{2.20c}$$

where now $\vec{u} = \frac{dr}{dt} = H\vec{r} + \vec{v}$ is the total velocity field, $\vec{v}$ the peculiar velocity and $r = ax$ the distance in physical units. Considering the perturbed system of equations, and treating the perturbations as plane waves one can recover the dispersion relation for an expanding Universe:

$$\ddot{\delta}_{\vec{k}} + 2H(t)\dot{\delta}_{\vec{k}} + (k^2 c_s^2 - 4\pi G \rho_B)\delta_{\vec{k}} = 0 \ , \tag{2.21}$$

where $2H(t)\dot{\delta}_{\vec{k}}$ is called *Hubble friction*, which opposes the growth of perturbations, while $k^2 c_s^2$ accounts for the characteristic velocity field of the fluid. As the dispersion relation found for the static case, the perturbations only grow (or decrease) for $\lambda > \lambda_J$, where now λ_J depends also on $H(t)$. Therefore

one has to take into account the cosmological model, and the specific cosmic epoch (namely before or after the equivalence).

For example considering an Einstein de Sitter model with $w = 0$ (after the equivalence), one can make the following assumption on the solution of Eq. (2.21):

$$\delta_{\vec{k}} \propto t^\alpha \ . \tag{2.22}$$

With this choice, the dispersion relation can be written in terms of α, obtaining a second order equation in this variable:

$$3\alpha^2 + \alpha + \frac{k^2 c_s^2}{4\pi G \rho_B} - 2 = 0 \ , \tag{2.23}$$

which has a solution with positive discriminant only for the case $\lambda > \lambda_J = \frac{c_s}{5}\sqrt{\frac{24\pi}{G\rho_B}}$. The growing mode of this solution can be written as:

$$\delta_+(t) \propto t^{2/3} \propto a(t) \ . \tag{2.24}$$

For a general cosmology, and considering only $z < z_{eq}$, one can use the following linear approximation for the growing solution of the dispersion relation (2.21):

$$\delta_+ \propto -H(z) \int_\infty^z \frac{(1 + z')}{a_0^2 H^3(z)} \ , \tag{2.25}$$

which shows that for an *open* Universe the growth of perturbations is slower than for an EdS model, while for a *closed* Universe is faster, as one can expect from the different rate of expansion.

Finally, during the *radiation epoch*, in which the radiation component is dominant over the others, one can show that λ_J is always greater than λ_H, because the sound velocity has a very high value and perturbations are dissipated by the radiative pressure. Consequently, during this cosmic epoch, perturbations cannot grow for $\lambda < \lambda_H$. In general, the dispersion relation that we have found is a second order differential equation, hence it has a solution that can be expressed as the combination of the growing (δ_+) and decaying (δ_-) solutions:

$$\delta(\vec{x}, t) = A(\vec{x})\delta_+(t) + B(\vec{x})\delta_-(t) \ , \tag{2.26}$$

where A and B are two functions depending on the comoving coordinates, while δ_+ and δ_- are time-dependent. It is possible to derive an approximate form for the variation of the growing solution, namely the *growth factor f*:

$$f \equiv \frac{\mathrm{d}lnD_+}{\mathrm{d}lna} \simeq \Omega_m^{0.55} + \frac{\Omega_\Lambda}{70}\left(1 + \frac{1}{2}\Omega_m\right) \ . \tag{2.27}$$

In this expression we can see how the growth of the perturbations depends mainly on the value of Ω_m, while Ω_Λ, which contributes to the Universe expansion, does not play a crucial role. The growth factor is of fundamental importance because the exponent of Ω_m, $\gamma \simeq 0.55$, comes from the General Relativity, thus the measure of f can be used to test the underlying gravity theory of our Universe.

2.3 Dark matter perturbations

The matter perturbations that we described consist of both baryons and DM, which is assumed to be collisionless. The onset of these instabilities cannot happen during the cosmological time in which matter and radiation were still coupled (before the last scattering). In fact, the strong microphysical interactions present during this era would not allow any kind of growth of the perturbation amplitudes. It is known that the decoupling of DM and radiation happened earlier than the decoupling of baryonic matter. This means that at the baryon-radiation decoupling, DM already collapsed forming the first structures, called *DM haloes*. The presence of DM potential well contributed to the growth of baryonic gravitational instabilities, an effect called *baryon catch-up*. Bound in the potential wells of DM haloes, baryons proceed to cool, condense, and form the first structures like stars and, later, galaxies. Moreover, to justify the value of density perturbations in formed structures today, it is crucial to include DM instabilities in our structure formation model. Let us briefly see why.

In order to reach at least $\delta \sim 1$ before $z = 0$, considering an EdS model, the starting overdensity field of baryonic matter should have been $\delta \sim 10^{-3}$, which is in contrast with observations of CMB radiation anisotropies (Eq. 2.1). One might try to solve this problem by considering a closed Universe, in which perturbations grow faster. However, in a Universe composed only by baryonic matter the total density parameter should be $\Omega \sim 20$ in order to justify the observed collapsed structures at the present time, which again is in contrast with the commonly accepted values $\Omega = 1$ from CMB observations (Sievers et al., 2003; Planck Collaboration et al., 2020).

2.3.1 Cold dark matter and hot dark matter

The formation of galaxies and galaxy clusters follows from the collapse of DM into haloes before the decoupling of baryonic matter and radiation. We can define two types of DM particles according to their nature at the time of their decoupling from radiation. Hot dark matter (HDM) is a kind of

DM which is still relativistic (i.e. hot) at that time, while cold dark matter (CDM) de-relativised before decoupling. Consequently, CDM particles are supposed to be more massive than HDM particles.

Critically important is the concept of Jeans mass that discriminates which perturbations are able to grow and which ones oscillate. We saw that the Jeans scale is proportional to the velocity of the perturbation. Considering the velocity dispersion instead of the velocity (for the non-collisionality property of DM), what follows is that for HDM the Jeans scale is large, leading to a greater value for the Jeans mass. The Jeans mass of HDM computed at matter-radiation equivalence is estimated to be approximately $10^{15-16}\ M_\odot$, whereas for CDM M_J is about $10^{5-6}\ M_\odot$. This huge difference implies different scenarios of structure formation. More specifically, HDM favours the so-called *top-down* scenario in which large-scale structures form first, and smaller structures are formed subsequently by fragmentation. On the other hand, in the CDM model one has that smaller structures form first, and then larger structures are formed by aggregation through merging (*bottom-up* scenario). The latter is favoured by observations, since high-z structures tend to be smaller.

2.4   Statistical properties of the Universe

The Jeans instability theory developed in this Chapter applies only to a single perturbation in the density field. In general, every single perturbation can be expressed as the superposition of multiple Fourier modes, i.e. plane waves, which evolve independently in time. From the fact that we cannot know the value of $\delta(\vec{x})$ at every position $\vec{x}$, we need a statistical approach to this problem. This implies that any comparison between theory and observations will also have to be statistical. A problem related to this is that the Universe is a unique object, not reproducible, and this could make difficult the construction of a meaningful statistical framework. Nevertheless, we can exploit the *ergodic hypothesis*, which states that the result of the mean of many realizations of the Universe is equivalent to the mean obtained considering separate, but sufficiently large, volumes in the Universe. According to this, sufficiently distant regions evolve independently with each other. The cosmological principle combined with the ergodic hypothesis forms the so called *Fair Sample* hypothesis (Peebles, 1980; Coles, 2003).

The density contrast $\delta(\vec{x})$ can be also expressed as:

$$\delta(\vec{x}) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \delta(\vec{k})e^{-i\vec{k}\cdot\vec{x}}\mathrm{d}^3\vec{k}\ , \tag{2.28}$$

i.e. as the inverse Fourier transform of $\delta(\vec{k})$, which has the dimension of a volume.

Let us consider now two perturbations on different scales $\delta(\vec{k})$ and $\delta(\vec{k'})$. One can measure the mean quadratic amplitude of the fluctuations with the so-called *power spectrum* $P(k)$, defined through the following formula:

$$\langle \delta(\vec{k})\delta^*(\vec{k}) \rangle = \langle |\delta(\vec{k})|^2 \rangle = (2\pi)^3 P(k)\delta_D^{(3)}(\vec{k} - \vec{k'}) , \qquad (2.29)$$

where $\delta^*(\vec{k}) = \delta(-\vec{k})$ is the complex conjugate of $\delta(\vec{k})$ while $\delta_D^{(3)}(\vec{k} - \vec{k'})$ is the 3-dimensional Dirac delta function. The initial shape of the matter power spectrum is determined by the fluctuations imprinted by *inflation* (Guth, 1981; Guth & Pi, 1982). 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), and that inflation does not predict any preferred scale, the initial $P(k)$ should have a power-law form:

$$P(k) = Ak^n , \qquad (2.30)$$

where the spectral index is generally assumed to be closed to unity $n = 1$, which corresponds to *white noise* gravitational, i.e. *metric*, fluctuations (Zeldovich, 1972). The normalisation A is not constrained by the inflation theory and has to be derived from observations. In fact the inflation scenario predicts a family of parallel lines of logarithmic slope n, but not a preferred one.

From the definition of the power spectrum descends the Wiener-Khinchin theorem, which states that the inverse Fourier transform of the power spectrum is the *two-point correlation function* (2PCF), its analogous quantity in real space:

$$\xi(r) = \langle \delta(\vec{x})\delta(\vec{x} + \vec{r}) \rangle = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} P(k)e^{i\vec{k}\cdot\vec{x}} \mathrm{d}^3\vec{k} . \qquad (2.31)$$

The latter measures the *auto-correlation* of the field $\delta(\vec{x})$ in positions at a distance $\vec{r}$ from $\vec{x}$, i.e. the spatial clustering of objects. This represents a double operation of mean, because the mean is evaluated for all the positions $\vec{x}$ in the Universe, and for all points at a distance r from each $\vec{x}$. The dependence of ξ on the modulus of $\vec{r}$ is a consequence of the cosmological principle, i.e. the invariance of ξ for translations and rotations.

The 2PCF can also be defined as the excess probability, with respect to a random distribution, of finding a pair of objects separated by a comoving distance r, in two different volume elements $\mathrm{d}V_1$ and $\mathrm{d}V_2$:

$$\mathrm{d}P_{12} = n^2[1 + \xi(r)]\mathrm{d}V_1\mathrm{d}V_2 . \qquad (2.32)$$

The 2PCF is often measured by employing the so-called *estimators*, which are based on the pair counts of objects at different separation r. In order to estimate the level of clustering of a given galaxy distribution, one has to compare the latter with a simulated random catalogue, having null clustering. The most used 2PCF estimator is the Landy-Szalay estimator (Landy & Szalay, 1993):

$$\hat{\xi}(r) = \frac{N_{RR}}{N_{DD}} \frac{DD(r)}{RR(r)} - 2 \frac{N_{RR}}{N_{DR}} \frac{DR(r)}{RR(r)} + 1 \; , \qquad (2.33)$$

where $DD(r)$, $RR(r)$ and $DR(r)$ are the number of pairs of data-data, random-random and data-random objects with distance $r + \Delta r$, while N_{RR}, N_{DD} and N_{DR} are the total number of data-data, random-random and data-random pairs in the catalogues, respectively.

In the real space, the assumption of stochastical generation of fluctuations from the inflation process implies that their amplitudes have a Gaussian distribution (Bardeen et al., 1986). A very relevant fact about this property is that the ergodic hyphotesis becomes a *theorem* for Gaussian distributed fields.

The variance σ^2 of $\delta(\vec{x})$ is defined by:

$$\sigma^2 = \langle |\delta^2(\vec{x})| \rangle = \sum_{\vec{k}} \langle |\delta_{\vec{k}}|^2 \rangle = \frac{1}{V_u} \sum_{\vec{k}} \delta_{\vec{k}}^2 \; , \qquad (2.34)$$

where the average is taken over an ensemble of realisations of volume V_u. By considering the limit $V_u \to \infty$ and assuming the validity of the CP, we obtain:

$$\sigma^2 \to \frac{1}{2\pi^2} \int_0^\infty P(k) k^2 \mathrm{d}k \; . \qquad (2.35)$$

This relation tells us that the variance of fluctuations is an integrated information of the power spectrum.

From an observational point of view, we cannot measure a punctual value for δ, from the fact that in reality the distribution of mass tracers is discrete, and there is the necessity of averaging on a scale. Therefore it is convenient to describe the fluctuation field as a function of some resolution scale R. We can define a density fluctuation from a discrete tracers distribution as:

$$\delta_M = \frac{M - \langle M \rangle}{\langle M \rangle} \; , \qquad (2.36)$$

where $\langle M \rangle$ is the average of mass found in a spherical region of radius R. From this definition one can recover the mass variance using (2.34):

$$\sigma_M^2 = \langle |\delta_M^2| \rangle = \frac{\langle (M - \langle M \rangle)^2 \rangle}{\langle M \rangle^2} \; . \qquad (2.37)$$

The fluctuation δ_M is related to the convolution of the punctual density with a *window function* (or *filter*) W of radius R:

$$\delta_M(\vec{x}) = \delta(\vec{x}) \otimes W(\vec{x}, R) \ . \tag{2.38}$$

For each filter, one can define a mass $M = \gamma_f \bar{\rho} R^3$, where γ_f is some constant that depends on the shape of the filter. Usually $W(\vec{x}, R)$ is a *top-hat* filter:

$$W(\vec{x}, R) = \frac{3}{4\pi R^3} \Theta\left(1 - \frac{|\vec{x} - \vec{x}'|}{R}\right) \ , \tag{2.39}$$

where Θ is the Heaviside step function.

From the last two equations and using the convolution theorem, it is possible to express the mass variance, passing in the limit expressed by Eq. (2.35), as

$$\sigma_M^2 = \frac{1}{2\pi^3} \int \mathrm{d}^3\vec{k} P(\vec{k}) \hat{W}^2(\vec{k}, R) \ , \tag{2.40}$$

where $\hat{W}$ is the Fourier-transform of the window function.

Though the normalisation of the power spectrum is not predicted by inflation theory, it is often fixed by using the value at $z = 0$ of the mass variance computed at 8 Mpch^{-1}, indicated as σ_8:

$$\sigma_8^2 = \frac{1}{2\pi^2} \int P(k) k^2 W^2(R = 8 \ h^{-1}\mathrm{Mpc}) \mathrm{d}k \ . \tag{2.41}$$

This is a key cosmological parameter, useful to predict the phenomenology of the low-redshift Universe. It has been measured with a variety of cosmological probes such as X-ray cluster counts (Pierpaoli et al., 2003), weak lensing by large scale structures (*cosmic shear*) (Refregier et al., 2002), CMB temperature (Komatsu et al., 2009, 2011; Planck Collaboration et al., 2020) wich have constrained it to be in the range $\sigma_8 \in (0.7-0.9)$. Also the numer counts of cosmic voids can provide reliable constraints on this parameter. Though it has not been exploited as stand-alone probe yet, void number counts will be soon applied to the upcoming wide-field surveys to derive cosmological constraints. Contarini et al. (2019) have indeed demonstrated the constraining power of this statistics by analysing the abundance of voids identified in DM haloes mock catalogues. An analogous method will be tested in Chapter 5 using voids identified in simulated galaxy distribution.

From the properties of the window function, one can see that the mass variance tends to its punctual value when $R \to 0$, and tends to zero when filtering on large scales ($R \to \infty$), because the perturbations are more and more smoothed. 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 $\lambda \sim k^{-1} > R$. Note that the *Zel'dovich spectrum* of Eq. (2.30), where $n = 1$, is a growing function of k, so larger wavelenghts contribute less. Moreover, if $\delta(\vec{x})$ is a Gaussian random field, then so it is $\delta(\vec{x}, R)$, from the fact that the convolution operation preserves Gaussianity. In particular, the distribution of values for the smoothed contrast becomes:

$$\mathcal{P}(\delta_M)\mathrm{d}\delta_M = \frac{1}{\sqrt{2\pi}\sigma_M} \exp\left[-\frac{\delta_M^2}{2\sigma_M^2}\right]\mathrm{d}\delta_M \ , \qquad (2.42)$$

where $\mathcal{P}(\delta_M)\mathrm{d}\delta_M$ is defined as the probability of having a contrast δ_M between δ_M and $\delta_M + \mathrm{d}\delta_M$.

2.4.1 Evolution of the power spectrum

When perturbations enter inside the horizon in an epoch before the equivalence, the density perturbations are damped by an effect called *stagnation*, or *Mészáros effect* (Mészáros, 1974). This effect is simply a manifestation of the fact that the Hubble drag term during the radiation dominated era is larger than during the matter dominated era. Indeed if one compares the free-fall time, i.e. the characteristic time that would take a perturbation to collapse under its own gravitational force, and the Hubble time, i.e. the characteristic time for the expansion of the Universe, one finds:

$$\begin{aligned}
\frac{\tau_H}{\tau_{ff}} &\propto (\rho_m/\rho_R)^{1/2} \gg 1, \quad t < t_{eq} \\
\frac{\tau_H}{\tau_{ff}} &\sim 1, \quad t > t_{eq} \ .
\end{aligned} \qquad (2.43)$$

Perturbations stay approximately constant in this regime, and this is the main effect that changes the primordial shape of the perturbations power spectrum.

Let us consider a perturbation of DM having $\lambda > \lambda_H$. Since $\lambda_H \equiv R_H$ expands with time, larger perturbations will enter the cosmological horizon at later times, hence they will undergo less stagnation (or zero stagnation, if they do not enter in λ_H before t_{eq}). On the other hand, the perturbations on scales bigger than the horizon continue to grow at the same rate, as seen in Eq. (2.19), independently of the scale or wavenumber. As a consequence, the power spectrum at the moment of the equivalence has a peak in correspondence of $k_{H,eq}$, that is the wavenumber associated with the cosmological horizon at the equivalence time. The overall evolution of the primordial power spectrum $P(k)$ can be summarised defining a *transfer function $T(k)$*. This function gives the fraction of the power spectrum that is not affected

by the microphysical effects inside the horizon. Considering a generic cosmological time represented by a_i and the equivalence a_{eq}, the transfer function is defined as follows:

$$P(k, t_{eq}) = P(k, a_i)T^2(k) \ . \tag{2.44}$$

For the CDM/bottom-up scenario, one has:

$$T(k) = \begin{cases} k^0, & k < k_{H,eq}, \\ k^{-2}, & k > k_{H,eq} \ . \end{cases} \tag{2.45}$$

This function therefore acts as a filter that blocks larger wavenumbers. The position of $k_{H,eq}$ corresponds to a physical scale determined by the matter ($\Omega_m h^2$) and radiation densities ($\Omega_r h^2$). Moreover, the shape of the observed power spectrum $P(k)$ depends on the amount and the nature of the matter in the Universe, providing powerful constraints for Cosmology. For example, if the DM were totally hot, then the matter power spectrum would fall off sharply to zero to the right of the peak, as it can be seen in Figure 2.1.
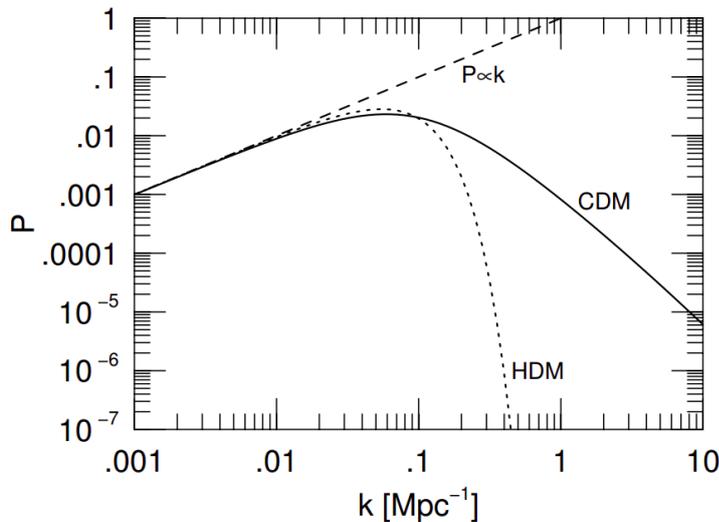

Figure 2.1: The power spectrum at the equivalence time for CDM (*solid line*) and HDM (*dotted line*). The *dashed line* represents the primordial Zel'dovich power spectrum. Credits to Ryden (2003).

Figure 2.2 shows the matter power spectrum as measured using different tracers at the present time. According to the theory, the power spectrum on large scales still has its primordial power-law shape and galaxy clusters trace
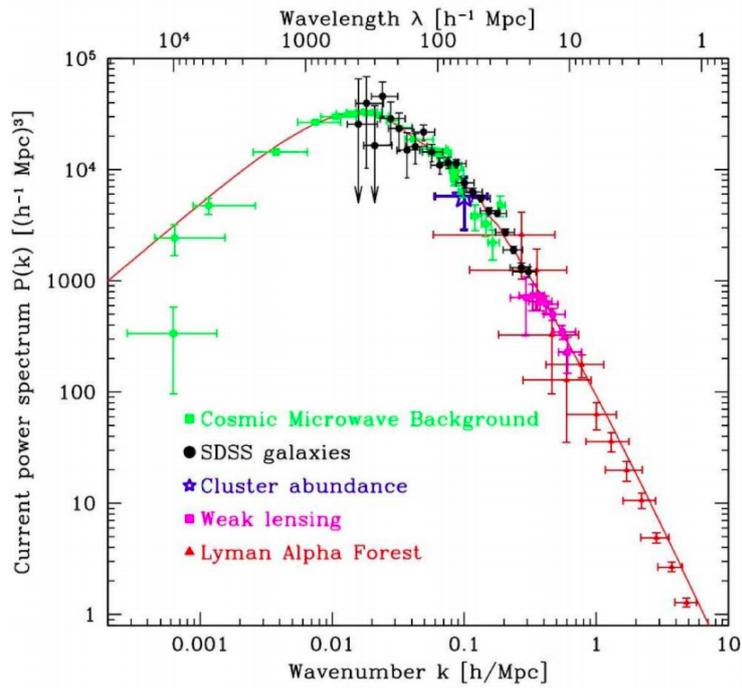
Figure 2.2: The matter power spectrum $P(k)$ as a function of wave-number k at the present time. Note the turnover at $k_{H,eq} \sim 0.01h^{-1}Mpc$. The plot combines data from different scales: CMB, large galaxy surveys, weak lensing and Lyα forest, in order of decreasing co-moving wavelength. In addition, there is a single data point for galaxy clusters. Figure from Tegmark et al. (2004).

the linear domain of the matter fluctuations. As shown in the same figure, there are various probes that can be used to constrain the power spectrum shape. CMB temperature fluctuation observations provides constraints on large scales, while large galaxy surveys predict the shape at the turnover up to 0.1 $h^{-1}Mpc$. Cosmic shear measurements can also be used for this aim, as shown e.g. in Lewis & Challinor (2006). Finally, the absorption features of the Lyman-α emission line in the spectra of high-redshift quasars are widely used as biased tracers of the density fluctuations of a photo-ionised warm intergalactic medium, and are collectively known as the Lyman-α forest (Croft et al., 2002; Viel et al., 2004).

2.4.2 Bias parameter

A fundamental problem in Cosmology is to understand how the spatial distribution of tracers is related to that of the underlying mass. If one counts galaxies in a volume V, one can still define an overdensity field as

$$\delta_{gal} = \frac{N_{gal}(V) - \bar{N}_{gal}(V)}{\bar{N}_{gal}(V)} , \qquad (2.46)$$

where N_{gal} is the number counts and $\bar{N}_{gal}$ the mean number counts insed a volume V.

However, one can not know in principle how the galaxy distribution reflects the distribution of the total matter in the Universe. The simplest model to parametrise the relation between δ_{gal} and δ_M is the linear, local, non-stochastic bias model:

$$\delta_M = b\delta_{gal} , \qquad (2.47)$$

where b is the linear bias factor, which depends on Cosmology, and on galaxy properties such as luminosity, morphology, colour and redshift. The Eq. (2.47) was proposed by Kaiser (1984) considering galaxy clusters, which are still in a linear regime, as discussed in the previous Section. This does not hold for smaller scales, where the relation becomes very complex due to the nonlinear evolution of galaxies. The bias does not only affect galaxies. Also the distribution of DM haloes is biased, because they form from the total matter distribution via collapse, only when a specific contrast threshold is exceeded (see Section 3.2.1).

One can use the analytical bias parameter proposed by Mo & White (1996) for the DM halo bias, which was found applying the so-called *Excursion-set* (ES) formalism (see Section 3.4 for details):

$$b(M, z) = 1 + \frac{1}{\delta_c}\Big(\frac{\delta_c^2}{\sigma_M^2 \delta_+^2(z)} - 1\Big) . \qquad (2.48)$$

Eq. (2.48) implies that massive haloes are positively biased, as also confirmed by simulations (e.g. Hu & Kravtsov, 2003). Moreover the bias factor grows with redshift.

One has to take into account the bias effect also for the observations of the 2PCF of galaxies. It can be inferred by computing the square root of the ratio of the galaxy 2PCF relative to that of DM particles:

$$b = \sqrt{\frac{\xi_{gal}}{\xi_{DM}}} \ . \tag{2.49}$$

The same can be defined for the ratio between DM haloes and DM particles, thus massive haloes are more clustered. Similarly, also the observed power spectrum allows to constrain both the amplitude and the scale dependence of the galaxy bias (Padmanabhan et al., 2007).

2.5 Nonlinear evolution of perturbations

Up to now we described the evolution of perturbations and their power spectrum in the linear regime, i.e. as long as $\delta \ll 1$. As the density contrast δ approaches unity, the evolution of the density fluctuations becomes nonlinear. During the nonlinear evolution, overdensities contract, causing matter to flow from larger to smaller scales causing the power spectrum to deform. When the variance of the perturbation distribution reaches the value $\sigma^2 \simeq 1$, the structures can be considered as formed and the nonlinear regime is established.

Consider the mass variance as expressed in Eq. (2.40) and a power-law shape for $P(k)$. Neglecting the effect of the window function one can find that

$$\sigma_M^2 \propto \delta_+^2(t) k^{n+3} \ , \tag{2.50}$$

where we have used $\delta_M(t) = \delta_{M,in} \delta_+(t)$. Moreover, the mass is dimensionally proportional to the cube of R, that implies:

$$\sigma_M^2 \propto \delta_+^2(t) M^{-\frac{n+3}{3}}. \tag{2.51}$$

Let us associate the formation of a structure of mass M_* to $\sigma_M^2 = 1$, which happens at a given time $t = t_*$. One has that

$$M_* \propto \delta_+^{6/(n+3)}. \tag{2.52}$$

For an EdS Universe, which is a good approximation for our Universe when it began to be nonlinear, we know that $\delta_+(t) \propto a \propto t^{2/3}$. Thus we can express the time t_* as

$$t_* \propto (1+z)^{-3/2} \propto M_*^{(3+n)/4}. \tag{2.53}$$

In a hierarchical model (bottom-up), the formation time must be smaller for lighter objects, therefore $n > -3$. In addition the energetic of the systems should increase with the mass, so that $n < 1$. Therefore the hierarchical scenario predicts a range for the spectral index n, which is:

$$-3 < n < 1 . \tag{2.54}$$

Data from Planck 2018 release, put a constraint of $n \simeq 0.96$ (Planck Collaboration et al., 2020).

As one can see from Eq. (2.51), the mass variance grows as the mass decreases. Therefore, at any fixed moment in time, a perturbation having mass M_1 will have grown more with respect to a perturbation of mass $M_2 > M_1$. This latter fact is true only for $n > -3$, and tells us that structures with lower masses reach nonlinearity earlier. As a consequence, the probability distribution of the density contrasts, δ, on small scales becomes larger with time first, until the shape of the distribution deviates completely from a Gaussian. From the definition of density contrast we find that the latter is limited in its negative part:

$$\delta = \frac{\rho - \bar{\rho}}{\bar{\rho}} = \frac{\rho}{\bar{\rho}} - 1 > -1 , \tag{2.55}$$

which implies that $\mathcal{P}(\delta < -1) = 0$. This causes a non-Gaussian evolution for $\mathcal{P}(\delta_M)$, hence it will become asymmetrical, tending to privilege the formation of underdense regions and leading to the formation of the *cosmic web* pattern observed in large galaxy surveys (see Figure 2.3 for an example). This is due both to the fact that it meets the barrier at $\delta_M = -1$ and to the fact that the area subtended by the curve $\mathcal{P}(\delta_M)$ must remain constant during its evolution, being a normalised probability.

Even though the linear perturbation theory fails for $|\delta| \geq 1$, the onset of nonlinear evolution can still be described analytically with approximate models such as the *spherical collapse* which leads to the Press-Schechter formalism (see Section 3.2.1). In general, however, it is necessary to use numerical methods in order to follow the nonlinear evolution of the structures in the Universe. Therefore, the exploitation of N-body simulations is mandatory, and typically the initial conditions are set through the use of the Zel'dovich approximation.

2.5.1 The Zel'dovich approximation

The transition for the density contrast between linear and nonlinear regime can be described analitically by the Zel'dovich approximation (Zel'Dovich,
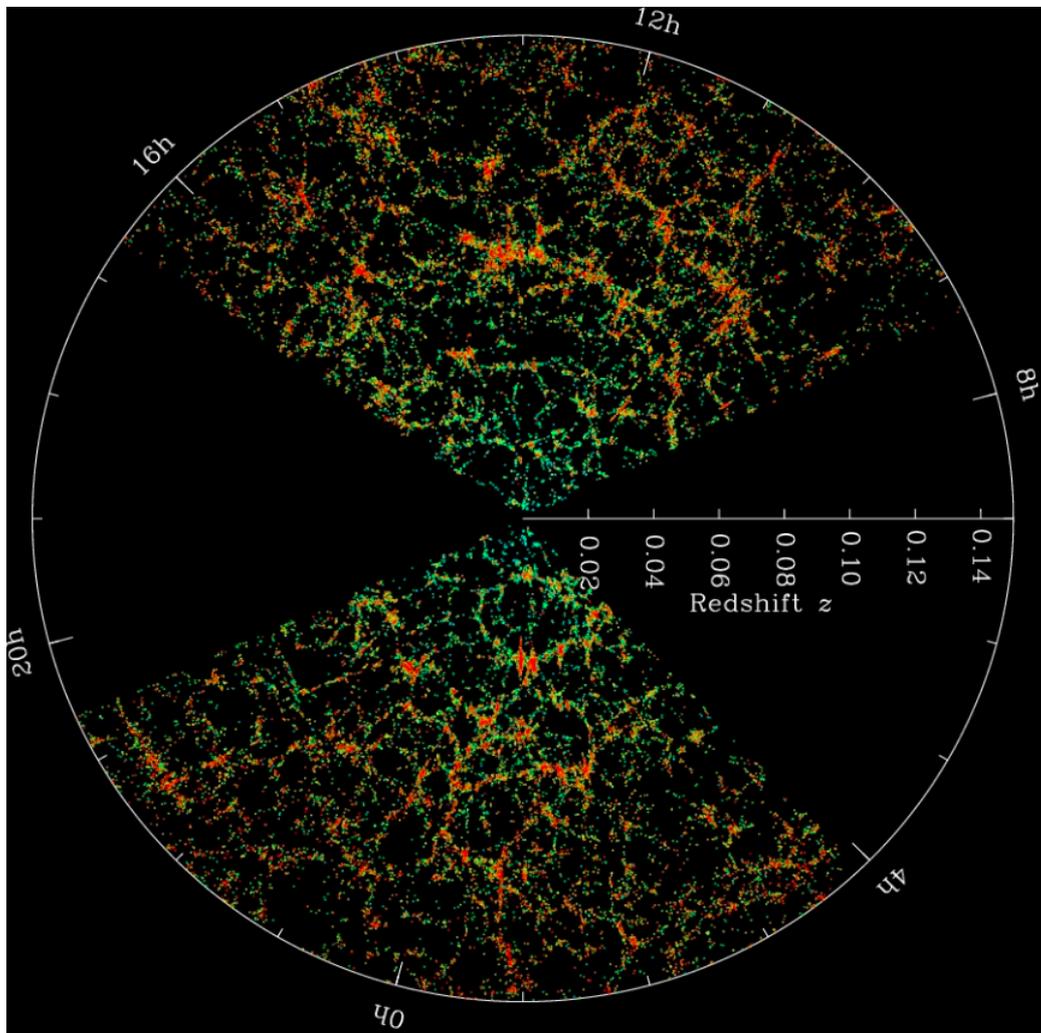
Figure 2.3: The distribution of galaxies in two slices of the Sloan Digital Sky Survey (SDSS). Every dot is a galaxy containing about 100 billion stars. Redder colours represent galaxies made of older, more clustered stars. Our own Galaxy is in the centre of this image, and the galaxies at the outer circle are approximately 600 Mpc far away from us. Credits to Mr. Blanton and the Sloan Digital Sky Survey (http://www.sdss3.org/science/).

1970) (see the review by Shandarin & Zeldovich (1989), for an exhaustive description). The Zel'dovich approximation is particularly simple in comoving coordinates $\vec{r} = \vec{x}/a(t)$, where $a(t)$ is the scale factor and $\vec{x}$ are physical coordinates. It relates the initial Lagrangian coordinates $\vec{q}$ at $t \to 0$ and $\vec{r}$ at time t by an explicit relation:

$$\vec{r}(\vec{q}, t) = \vec{q} + \delta_+(t)\vec{s}(\vec{q}) \; , \tag{2.56}$$

where the vector field $\vec{s}(\vec{q})$ is called the *initial displacement field* and it is determined by the initial density perturbations, while $\delta_+(t)$ is the linear density growth factor fully specified by the cosmological parameters. The Zel'dovich approximation assumes that $\vec{s}(\vec{q})$ is a potential vector field:

$$\vec{s}(\vec{q}) = -\nabla_q \Psi(\vec{q}) \; . \tag{2.57}$$

The main limit of this approach is that the particles feel only the initial force, which causes their displacement, but at later times there are no additional interactions. This implies that two particles can cross each other without any interaction, giving rise to the so-called *shell-crossing problem*. However these problems affect mainly the modelling of the small scales, where nonlinearity develops first.

An additionally important aspect of the Zel'dovich approximation, reflecting its Lagrangian nature, is the deformation of mass elements. The deformation is specified by the following tensor:

$$d_{ij} = -\frac{\partial^2 \Psi}{\partial q_i \partial q_j} \; . \tag{2.58}$$

From Eq. (2.56) one may easily infer an explicit expression for the density as a function of Lagrangian coordinates and time. If we consider the conservation of mass in differential form:

$$\rho(\vec{r}, t)\mathrm{d}^3\vec{r} = \bar{\rho}(t)\mathrm{d}^3\vec{q} \; , \tag{2.59}$$

the density evolution directly follows from:

$$\begin{aligned}
\rho(\vec{x}, t) &= \bar{\rho}\Big[\delta_{ij} + \delta_+(t)d_{ij}\Big] \\
&= \bar{\rho}[1 - \delta_+(t)\lambda_1]^{-1}[1 - \delta_+(t)\lambda_2]^{-1}[1 - \delta_+(t)\lambda_3]^{-1} \; ,
\end{aligned} \tag{2.60}$$

where $\lambda_1 > \lambda_2 > \lambda_3$ are the eigenvalues of the symmetric deformation tensor d_{ij}. On the basis of the second equation of (2.60), we may immediately infer two key features of structure formation described by the Zel'dovich

approximation. The first one is that the density becomes infinite as soon as $\delta_+(t)\lambda_i = 1$. In addition, this collapse is nearly always anisotropic as in general the eigenvalues $\lambda_1 \neq \lambda_2 \neq \lambda_3$.

The general idea is that the bracket containing the largest λ_i tends to zero first, and this means that the minor axis collapses earlier. On the other hand, if the eigenvalues are negative, the brackets can not be null at any time, and a dilatation takes place, instead of a collapse. Therefore different combinations of positive and negative eigenvalues lead to different evolution of large scale structures.

2.6 Numerical simulations in Cosmology

So far we have seen how the structure formation can be treated analytically both in the linear and in the mild nonlinear regime. However, galaxies form in a strong nonlinear regime, which has very complex dynamics. Therefore, numerical simulations are required in order to properly model the main properties of the large scale structure properties of our Universe and their cosmological evolution. The first studies in which numerical simulations were employed are Aarseth (1963), Peebles (1970) and Press & Schechter (1974), in which simple N-body problems have been solved using N-body codes with few hundred particles. Nowadays we can make reliable predictions about a very large range of phenomena, using simulations having billions of particles. Especially, the last years have seen enormous progress on two fronts: large volume simulations modeling large samples of galaxies, and zoomed simulations with refined galaxy formation models that resolve the physical processes in more detail. Moreover, cosmological simulations are typically performed with periodic boundary conditions in order to mimic the large scale homogeneity and isotropy of the matter distribution in the Universe, i.e. the cosmological principle.

There exist two main kinds of numerical simulations, depending on the type of study we want to carry out:

- *N-body simulations*: they are used to simulate the DM distribution through particles from small scales to large scales. Only the gravitational interaction is involved;

- *Hydrodynamic simulations*: they involve also hydrodynamics, and so they can be used to describe the baryonic matter density field. Among others, these simulations describe the cooling of gas, the formation of stars, and the energy and momentum injection caused by supermassive black holes and supernovae.

In this Thesis work, in order to test the algorithms we implemented on the combination of multiple probes, we will use a set of cosmological and hydrodynamical simulations. In particular, these simulations are based on a specific type of algorithm, the *tree-particle-mesh* combined with a *smoothed-particle-hydrodynamics* code (see later), but we provide as well, in the following sections, a general description of the main numerical algorithms that are commonly used in Cosmology.

2.6.1 N-body simulations

Having a set of N particles with mass m_i, for $i = 1, \ldots, N$, the simplest kind of N-body simulations, which considers only gravitational effects, solves the Newtonian system of dynamical equations:

$$\begin{cases} \vec{F}_i = G m_i \sum_{i \neq j} \frac{m_j}{r_{ij}^2} \hat{r}_{ij}^2 \\ \ddot{\vec{x}}_i = \frac{\mathrm{d}\vec{v}_i}{\mathrm{d}t} = \frac{\vec{F}_i}{m_i} \\ \dot{\vec{x}}_i = \frac{\mathrm{d}\vec{x}_i}{\mathrm{d}t} = \vec{v}_i \; , \end{cases} \tag{2.61}$$

where $\vec{F}_i$ is the gravitational force acting on the i-th particle, $\vec{x}_i$ and $\vec{v}_i$ are respectively the comoving coordinates and the velocity of the i-th particle, r_{ij} is the comoving distance between i-th and j-th particles, and $\hat{r}_{ij}$ is the related unitary vector. The main challenge of the N-body simulations is to efficiently calculate the gravitational force that governs the motion of the sample of particles. Given the system (2.61), the fundamental equations that has to be considered for the gravitational interactions are, in comoving coordinates, the Euler equation:

$$\frac{\mathrm{d}\vec{v}_i}{\mathrm{d}t} + 2 \frac{\dot{a}}{a} \vec{v}_i = -\frac{1}{a^2} \nabla \phi = -\frac{G}{a^3} \sum_{i \neq j} m_j \frac{\vec{x}_i - \vec{x}_j}{|\vec{x}_i - \vec{x}_j|^3} = \frac{\vec{F}_i}{a^3} \tag{2.62}$$

and the Poisson equation:

$$\nabla^2 \phi = 4\pi G \bar{\rho} a^2 \delta = \frac{3}{2} H_0^2 \Omega_0 \frac{\delta}{a} \; , \tag{2.63}$$

where we used the second Friedmann equation (1.42). The Poisson equation can be written also in its integral form:

$$\phi(r) = -G \int \frac{\rho(\vec{x})}{|\vec{x}' - \vec{x}|} \mathrm{d}\vec{x}' \tag{2.64}$$

The simplest method to calculate the gravitational force acting on the i-th particle is the *Particle-Particle method* (PP), which is also the most accurate. This method consists in the evaluation of $\vec{F}_i$ acting on each particle at any time step. However it is the most expensive in terms of computational time, because for each time step it requires the computation of the $N(N-1)/2$ distances between the particles, i.e. the number of operations scales as $\mathcal{O}(N^2)$[1]. Moreover, the nature of the gravitational force leads to the fact that in the limit $r \to 0$ the force, and so the potential, would grow to infinity. Therefore it is necessary to introduce a *softening length*, below which the gravitational force is set to zero.

Other methods to calculate gravitational forces of the N-body system, that are more efficients than PP, are described below:

- **Hierarchical Tree (HT)**. This method consists in the division of the simulated volume of the Universe in cells, characterised by a proper centre of mass, whose sizes follow a hierarchy based on the spatial distribution of the particles. The volume of the simulation (*root*) is iteratively divided into smaller cells (*branches*) until each cell contains at most one particle (*leaves*). Given a particle P, then the gravitational force acting on it is computed with the PP method if the particles with which it interacts are sufficiently close. On the other hand, for distant clumps of particles it is sufficient to compute the force produced by the biggest cell containing the clump, with total mass equal to the sum of the clump's masses and position equal to their centre of mass. This is the most common method used to accelerate the PP method, and the number of operations involved with this method scales as $\mathcal{O}(N \log N)$. (Barnes & Hut, 1986)

- **Particle-Mesh (PM)**. While PP and tree method solve the integral form of the Poisson equation, mesh-based methods aim at solving the differential form of Poisson's equation (2.63). This can be done efficiently through Fast Fourier Transform-based methods (FFT), with Poisson's equation written in Fourier space: $\vec{k}^2 \hat{\phi}(\vec{k}) = -4\pi G \hat{\rho}(\vec{k})$, where $\hat{\phi}$ and $\hat{\rho}$ are, respectively, the Fourier transforms of the gravitational potential and density field, leading the so called particle-mesh method (PM) (Hockney & Eastwood, 1981). Here the computation of the forces is performed by implementing a regular grid, from which the density field is computed by interpolation. With FFT methods one

---

[1]This notation, known as the asymptotic notation, is used in computer science to classify algorithms according to how their run time or requirements of space grow as the input size grows.

can compute the FT of the density field and then calculate $\hat{\phi}$. Consequently the forces are obtained from the differentiation of the potential and then brought back to real space by applying an inverse FT. This method is faster than PP, indeed it scales as $\mathcal{O}(N \log N)$, where here N is the number of cells in the grid. However, this method is less accurate, because the resolution depends on the size of the cells.

- **Hybrid methods**. A variety of schemes combine direct summation-based techniques (such as PP), for short range forces, with FFT methods, for long range forces. The most basic example of this is the particle-particle plus particle-mesh method (P^{3}M) (Efstathiou et al., 1985). This algorithm is more accurate than the PM: particles directly interact with each other in presence of zones with strong clustering. However, it implies an increasing of computational cost, which locally tends to be proportional to $\mathcal{O}(N^2)$. Assuming a scale R_k that discriminates these two regimes, with k a factor that depends on that scale, the number of operation scales as $\mathcal{O}(kN \log N)$. Another common hybrid scheme is the tree-particle-mesh (TPM) method, which combines the PM method on large scales with a tree code to handle particle-particle interactions at small separations.

Once the method used for the computation of the gravitational force is chosen, the system (2.61) has to be updated in order to obtain the time evolution, hence the orbits, of the particles. Given the time interval Δt, for each time step t_j one can update the positions and the velocities of the particles by using the so-called *leapfrog method*, in which spatial coordinates are computed at half time step ($t_{j+1/2}$) in order to ensure a better accuracy for the velocity computation at t_{j+1}. Consequently, one has to update the spatial coordinates at t_{j+1} using the latter value of velocity:

$$\begin{cases} \vec{x}(t_{j+1/2}) = \vec{x}(t_j) + \dfrac{1}{2}\vec{v}(t_j)\Delta t \\ \vec{v}(t_{j+1}) = \vec{v}(t_j) + \vec{g}[\vec{x}(t_{j+1/2})]\Delta t \\ \vec{x}(t_{j+1}) = \vec{x}(t_{j+1/2}) + \dfrac{1}{2}\vec{v}(t_{j+1})\Delta t \ , \end{cases} \tag{2.65}$$

where $\vec{g} = -\nabla\phi(x)$ is the gravitational vector field.

There exist other methods for the time integration, such as the Runge-Kutta methods or the Euler method, however the leapfrog presents several desirable properties: it is accurate at the second order, it is time reversible, it preserves phase-space density (i.e. it is symplectic) and for a spherically-symmetric potential, it conserves angular momentum exactly.

2.6.2 Hydrodynamic simulations

DM and DE dominate the energy budget of the Universe, but the visible component of galaxies consists of baryons. The latter can be heated or cooled in a variety of ways, and the modelling of these processes is a very complex task, because of both the computational limitations of present day machines, and of our lack of knowledge of the physics involved. Moreover, baryonic interactions are important only on relatively small scales, hence the study of structure formation at scales larger than few Mpc can be considered sufficiently accurate even without accounting for the baryonic interactions. However, when studying smaller scales, the radiative processes related to baryons are fundamental in order to properly model the formation and evolution of galaxies, as well as to being able to link the matter distribution to the light distribution that we actually observe in the Universe.

The physics of this component can be included by making use of the exact solution of the set of Eqs. (2.5). Also, the hydrodynamic equations have to be complemented by various astrophysical processes that shape the galaxy population. Among these there are: the cooling of gas, the impact of the interstellar medium, the star formation process, stellar radiative feedbacks, supermassive black hole feedbacks and AGN feedback.

The hydrodynamical equations can be discretized in different ways employing methods that roughly fall into two classes: Lagrangian and Eulerian methods.

- **Eulerian methods**. These are the traditional methods to solve the system of hyperbolic partial differential equations that constitute ideal hydrodynamics. In these methods the mean values of the hydrodynamic variables are calculated at the nodes of one (or more) spatial grids, that could be fixed or adaptive. The latter have the advantage of being able to be refined in the regions where we want to study the phenomenon in greater detail. The advantage of using eulerian methods is that they allow to describe in more details shock events, i.e. phenomena that cause a steepening in the physical variables gradients. The major problem of these methods is the resolution, problem which arises from the fact that they deal with grids.

- **Lagrangian methods**. In these methods the fluid is modelled as an ensemble of particles, in analogy with the N-body simulations. The reference system moves with the flow of variables to be obtained (e.g. gas density, temperature, ..) and the hydrodynamic quantities are calculated at the position of particles. The advantage in using these methods lies in the fact that we are able to follow the evolution of the fluid de-

pending on the density. The code considers spatial accuracy greater in overdense regions than in underdense regions. However one needs to introduce an artificial viscosity in order to deal with steep gradients. The most used lagrangian algorithm is the *Smoothed-Particle-Hydrodynamics* (SPH). In this type of algorithm the fluid's properties are computed, at every positions, as the weighted average of the variables around that position. The fluid is discretised in a finite number of particles, in order to perform integrals as summations, and the hydrodynamical interactions are computed between a restricted group of closest particles, hence the total ensemble of particles is not considered.

The choice of the most efficient method depends on the type of problem to solve. Typically, if we want to simulate objects having high density (galaxies, clusters of galaxies) Lagrangian methods are recommended, since they do not have spatial resolution problems. On the other hand, if we want to do cosmological simulations, in which we simulate both overdense regions as well as underdense regions, Eulerian methods are better.

As for N-body simulations, also hydrodinamic simulations could be built to cover large scales or to cover only "small" environments like a single galaxy + interstellar medium around it. The results of hydrodynamical simulations can be directly compared with observational data providing important tests for galaxy formation models. Despite the differences in the methods and implementations of the galaxy formation, these simulations have now converged on a wide range of predictions for the evolution of galaxies. It therefore seems that the basic physical mechanisms that shape the galaxy population have been identified, and that their current modeling is sufficient to produce realistic galaxy populations. However, there is also a wide range of predictions that diverge among different simulations. For example, the characteristics of gas around galaxies are very sensitive to the feedback implementations used in the different galaxy formation models. This can lead to rather different outcomes for the thermodynamic structure of gas around galaxies.

Chapter 3

Statistical properties of galaxy clusters and cosmic voids

When we observe the properties of our Universe on large scales, what we see is a filamentary pattern traced by matter. This structure is composed by galaxies, which attract each other due to gravitational interactions and group together in galaxy clusters in the most overdense zones. Departing from these filaments and high-density nodes, luminous objects become increasingly difficult to find. Cosmic voids, the most wide and underdense regions of our Universe, fill the remaining volume of this intricate web of matter. Statistical properties of galaxy clusters and cosmic voids, like number counts and density profiles, can be exploited to test the underlying cosmological model and represent powerful probes to bind the fundamental cosmological parameters.

In this Thesis work, we implement new algorithms for probe combination, and investigate th constraining power of a joint analysis of the mass function of galaxy clusters and the size function of cosmic voids, considering these two as independent probes (see Chapter 4). In this Chapter we introduce the theoretical background for the evolution of both over- and under-dense perturbations of matter, starting from the spherical collapse model, that describes the formation and evolution of both DM haloes and cosmic voids. From this theoretical modelling we will derive an analytic expression for the mass function of DM haloes, inside which galaxy clusters form, following the Press & Schechter formalism (Press & Schechter, 1974). Then, we will present an extension of this formalism, called *excursion-set*, to derive also the size function of cosmic voids, as proposed by Sheth & van de Weygaert (2004). We will describe in details the models we will use in the next two Chapters, together with the prescriptions to properly obtain the abundance of galaxy clusters and voids from a catalogue of galaxies.

3.1 Galaxy clusters and cosmic voids

The evolution of the DM density fluctuations gives rise to the formation of deep potential wells inside which DM haloes form. The baryonic matter falls into these haloes, cools down, and collapses forming stars and galaxies, which will then cluster inside the most massive haloes. The way in which galaxy clusters build up from the interactions of baryonic matter inside DM haloes is a very complex and fascinating subject of study, and nowadays is mostly analysed using cosmological hydrodynamical simulations. Galaxy clusters represent the principal tracers of the matter distribution, being placed at the highest density peaks of the DM distribution.

From an observational point of view, galaxy clusters have masses between 10^{13} $M_\odot$ and 10^{15} $M_\odot$, and are typically composed by $10^2 - 10^3$ galaxies, having velocity dispersion around 10^3 km s^{-1}. Though their mass budget is dominated by DM ($\sim 80\% - 90\%$), the majority of the baryonic mass in these systems is represented by the hot plasma, called *IntraCluster Medium* (ICM), which is heated up to $T_{gas} = 10^8$ K, and is observable in the X-ray band. The strong X-ray emission of the ICM makes galaxy clusters one of the brightest class of objects that we observe in this spectroscopic band. At different wavelengths, one can study very disparate phenomena associated to galaxy clusters. For example, in the visible part of the light spectrum we observe the light coming from the star component, while at radio frequencies we can study the magnetic fields within the ICM, together with shocks caused by the interactions of galaxies with the hot gas. Moreover, the CMB photons interact with energetic electrons of the ICM and gain energy, leading to a distortion of the observed CMB spectrum in correspondence of clusters, the SZ effect. Galaxy clusters are not randomly distributed, but they group in even larger superclusters, having scales of hundreds of Mpc. Moreover, they are not isolated structures. Indeed, they are connected into a complex network by filaments of galaxies in the so-called supercluster-voids network (Einasto et al., 1980; Zeldovich et al., 1982; Einasto et al., 2011). In this network, galaxies and galaxy clusters constitute the high-density boundaries of cosmic voids. The latter can span a wide range of sizes, going from radii of few Mpc (*mini-voids*) to hundreds of Mpc (*super-voids*) (Tikhonov & Karachentsev, 2006; Granett et al., 2008). Thanks to the advancement of spectroscopic and photometric surveys, in the recent years cosmic voids have become a powerful probe to constrain cosmological parameters (see Pisani et al., 2019, for a comprehensive review). Indeed, their abundance and their internal density profiles and shapes, as well as their clustering properties, are strongly dependent on the underlying cosmological scenario. Moreover, being voids large regions nearly devoid of matter, they represent excellent

laboratories to test alternative gravity theories, as well as to study in details the effects of massive neutrinos (Villaescusa-Navarro et al., 2013; Schuster et al., 2019; Contarini et al., 2020) and *non-standard* DE models (Bos et al., 2012; Pisani et al., 2015; Verza et al., 2019).

According to the ΛCDM model, galaxies formed in voids should have different observational properties than those found in denser environments. For example, from cosmological simulations we see the evidence for lower mass haloes inside cosmic voids, with masses not high enough to trigger the star formation (Tikhonov & Klypin, 2009; Einasto et al., 2011). In general, it has been found that void galaxies have lower stellar mass (Croton et al., 2005), are more gas-rich (Kreckel et al., 2012) and are bluer than galaxies in denser environments (Hoyle et al., 2012).

One problem related to the observation and the analysis of galaxy clusters is the effect of redshift-space distorsions (RSD) caused by the peculiar motions of galaxies around and inside clusters. The latter cause a shift in galaxy observed positions along the line-of-sight, leading to a distortion in their spatial distribution. Since galaxies tend to fall inwards towards the cluster centre of gravity, we observe a fictitious squashing of the cluster shape. Modelling RSD is a very complex task, in particular on small scales, where perturbative models cannot precisely reproduce the nonlinear behaviour of the growth of cosmic structures ($R < 20$ Mpc h^{-1}). The effect of RSD is present also in cosmic voids, however their shapes and statistics are easier to model since voids remain in a mildly nonlinear regime during their evolution. Moreover, the galaxies inside them show coherent single-stream motions (Shandarin, 2011) from the centre towards the edge of these objects. We expect therefore an opposite effect with respect to the one observed for galaxy clusters: in this case the projection of the galaxy peculiar motions will produce an elongation of the void shape along the line-of-sight. The study of the phenomenon of the RSD in both clusters and voids is left as future improvement of this work, since these effects are not included in the analysed data-sets.

3.2   Spherical evolution

We are going now to describe the models at the base of DM halo and void formation and evolution from the growth of the density fluctuations.

Though numerical simulations are needed to study in details the growth of cosmic structures, we can learn about the evolution of perturbations in the nonlinear regime by making use of some assumptions. In the following, we consider a simple analytical model to describe the *isolated* formation of collapsed objects from overdensities, the so-called *top-hat model*, also known

as the *spherical evolution model* (Gunn & Gott, 1972), which has also been applied in theoretical studies of voids, which evolve from underdensity regions (Gunn & Gott, 1972; Peebles, 1980; Sheth & van de Weygaert, 2004). This model considers the nonlinear evolution of an initially spherical perturbation, which can be positive or negative, described as a closed or open Universe that evolves in an EdS background. We consider an initial time $t_i > t_{eq}$, where t_{eq} is the matter-radiation equivalence time, thus we study the evolution of perturbations in the matter-dominated cosmic epoch, but at redshifts high enough for assuming an EdS model for the background.

3.2.1 Overdensities

Dealing with overdensities, we expect that, if a perturbation has a sufficiently high density contrast, it will decouple from the Hubble flow and collapse. We know that the growing and decaying mode of perturbations in a dust EdS Universe are $\delta_+ \propto t^{2/3}$ and $\delta_- \propto t^{-1}$, respectively. Therefore the density contrast can be expressed as the combination of these two modes:

$$\delta = \delta_+(t_i)\left(\frac{t}{t_i}\right)^{2/3} + \delta_-(t_i)\left(\frac{t}{t_i}\right)^{-1}. \tag{3.1}$$

Assuming a null initial velocity of perturbations, we can derive with respect to the time the latter relation for δ, considering $t = t_i$, finding:

$$\frac{2}{3}\delta_+(t_i) - \delta_-(t_i) = 0 \implies \delta_-(t_i) = \frac{2}{3}\delta_+(t_i). \tag{3.2}$$

Therefore, Eq. (3.1) can be written, for $t = t_i$, as:

$$\delta_i = \frac{5}{3}\delta_+(t_i). \tag{3.3}$$

Hence 3/5 of the initial perturbation is represented by the growing mode, while the remaining 2/5 decays with time, tending to become negligible.

Now, let us consider the density parameter of the perturbation Universe Ω_p. The perturbation is described in terms of a closed Universe, which we know, from Section 1.9.3, that will undergo to a collapse. Consequently, we have $\Omega_p > 1$, since:

$$\Omega_p(t_i) \equiv \frac{\rho_p(t_i)}{\rho_{crit}(t_i)} = \frac{\rho_b(t_i)(1+\delta_i)}{\rho_{crit}(t_i)} = \Omega(t_i)(1+\delta_i) > 1, \tag{3.4}$$

where $\Omega(t_i)$ is the initial density parameter of the background Universe, and ρ_{crit} is the critical density. Thus, we find that for a closed Universe, it

is necessary that $(1 + \delta_i) > \Omega(t_i)^{-1}$. By considering a mono-component Universe with $w = 0$, we can use the equation for the evolution of the density parameter Eq. (1.60) in order to find the threshold for $\delta_+(t_i)$ that will lead to the collapse:

$$\delta_+(t_i) = \frac{3}{5}\delta_i > \frac{1 - \Omega_{0,b}}{\Omega_{0,b}(1 + z)} \,. \tag{3.5}$$

Here we clearly see that for closed and flat background Universes (i.e. for $\Omega_{0,b} \geq 0$) the collapse is achieved for any positive values of the initial density contrast, while for open Universes (i.e. for $\Omega_{0,b} < 0$) the expansion inhibits the collapse if δ_i is not sufficiently greater.

In Section 1.9.3 we obtained the parametric solution for a dust Universe, and we saw that this kind of Universe reaches its maximum size for $\theta = \pi$ at $t = t_{max}$, a status called *turn around*, after which it collapses under its own gravity. It is possible to show that the density of the perturbation at the turn around (i.e. for $t = t_{max}$) is:

$$\rho_p(t_{max}) = \frac{3\pi}{32Gt_{max}^2} \,. \tag{3.6}$$

We can calculate the density contrast at the turn around by computing the background density at t_{max} from Eq. (1.84). We obtain:

$$\delta_+(t_{max}) \simeq \frac{\rho_p(t_{max})}{\rho_b(t_{max})} - 1 = \left(\frac{3\pi}{4}\right)^2 4.6 - 1 \simeq 4.6 \,, \tag{3.7}$$

therefore the perturbation is already in the nonlinear regime before the collapse. The same quantity obtained with the linear theory, described in Section 2.2.2, would be instead:

$$\delta_+(t_{max}) = \delta_+(t_i)\left(\frac{t_{max}}{t_i}\right)^{2/3} \simeq 1.062 \,. \tag{3.8}$$

After the turn-around, the physical scale of the perturbation decreases until $t = 2t_{max}$, time at which the full collapse would be reached, leading to the formation of a black hole. However, it is much more likely that collapse stops, forming bounded objects in a way we have summarised in Section 2.3. The temperature of the gaseous baryonic matter increases until the internal pressure gradients become sufficient to balance the gravitational force. The cold DM component, during collapse, fragments into sub-units. Then, through the process of violent relaxation, these regions come to a dynamical equilibrium under the influence of large-scale gravitational potential gradients. In particular, from numerical simulations we know that the virialisation is

reached at a time $t_{vir} = 3t_{max}$, at which the size of the perturbation becomes stable, at the virialisation radius R_{vir}.

Let us assume that the perturbation system has kinetic energy $\mathcal{T}$ (or internal thermal energy, associated with motions of particles) and gravitational potential energy $\mathcal{V}$. The final result is a system which satisfies the *virial theorem*, which states that:

$$2\mathcal{T} + \mathcal{V} = 0\,. \tag{3.9}$$

Considering the potential energy of a self-gravitating sphere of mass M

$$\mathcal{V} = -\frac{3}{5}\frac{GM^2}{R}\,, \tag{3.10}$$

the total energy of the system becomes:

$$E = \mathcal{T} + \mathcal{V} = \frac{1}{2}\mathcal{V} = -\frac{3}{10}\frac{GM^2}{R}\,. \tag{3.11}$$

Let us assume the absence of any mass or energy losses since the turn-around, therefore:

$$E(t_{vir}) = E(t_{max}) \tag{3.12}$$

which leads to $2R_{vir} = R_{max}$. From the fact that $\rho_p(t_{vir}) \propto R_{vir}^{-3}$, it follows that

$$\rho_p(t_{vir}) = 8\rho_p(t_{max}). \tag{3.13}$$

Therefore we can now compute the density contrast at $t_{coll} = 2t_{max}$ and $t_{vir} = 3t_{max}$:

$$\delta_+(t_{coll}) = \frac{8\rho_p(t_{max})}{\rho_b(t_{max})}\left(\frac{t_{coll}}{t_{max}}\right)^2 \simeq 180\,,$$

$$\delta_+(t_{vir}) = \frac{8\rho_p(t_{max})}{\rho_b(t_{max})}\left(\frac{t_{vir}}{t_{max}}\right)^2 \simeq 400\,. \tag{3.14}$$

While the same quantities extrapolated from the linear theory are:

$$\delta_+(t_{coll}) = \frac{8\rho_p(t_{max})}{\rho_b(t_{max})}\left(\frac{t_{coll}}{t_{max}}\right)^{2/3} \simeq 1.686\,,$$

$$\delta_+(t_{vir}) = \frac{8\rho_p(t_{max})}{\rho_b(t_{max})}\left(\frac{t_{vir}}{t_{max}}\right)^{2/3} \simeq 2.2\,. \tag{3.15}$$

We stress the fact that these values represent the initial thresholds. Therefore one has to take into account also the *growth factor* in order to compute the desired collapse barrier at different redshifts, accordingly to the linear theory.

The quantities in Eq. (3.14) depend strongly on the cosmological model assumed for the background Universe, hence on its curvature. On the other hand, though the linearly extrapolated values for δ_+ are very different with respect to their nonlinear counterpart, the dependence of these quantities on the given cosmological model is much weaker (Jenkins et al., 2001; Kitayama & Suto, 1996). For example, in the ΛCDM model these become $\delta_{max} = 1.303$ and $\delta_c = 1.674$, for turn-around and collapse at $z = 0$. This kind of independence for δ_c of the Cosmology will be useful in the derivation of the Press-Schechter formalism (Section 3.4).

3.2.2 Underdensities

We will now turn to the description of the evolution of underdense perturbations, following the same underlying spherical model used for overdensities.

Let us consider an inverse top-hat spherically symmetric underdense perturbation as a set of concentric shells, uniform and without substructures. The evolution of each radius r is determined by the total mass M contained within R via the acceleration equation in the Newtonian regime (i.e. $\dot{r} \ll c$ and $r \ll r_c \sim c/H$):

$$\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = -\frac{GM}{r} = -\frac{4\pi G}{3}\rho_b(1+\Delta)r\,, \qquad (3.16)$$

where, at the initial time

$$
\begin{aligned}
M &= \frac{4\pi}{3}\rho_b r_i^3 (1 + \Delta_i)\,, \\
\Delta_i &= \frac{3}{r_i^3}\int_0^{r_i} \delta_i(r) r^2 \mathrm{d}r\,.
\end{aligned}
\qquad (3.17)
$$

The same equation, applied to an unperturbed region, is the first Friedmann equation, which yields the expansion history of the Universe. The Eq. (3.16) can be solved analytically and the solution for the size of the radius as a function of time takes the following parametric form when applied to voids (Gunn & Gott, 1972; Lilje & Lahav, 1991; Sheth & van de Weygaert, 2004):

$$
\begin{aligned}
R &= A(\cosh\theta - 1)\,, \\
t + T &= B(\theta - \sinh\theta)\,, \\
A^3 &= GMB^2\,,
\end{aligned}
\qquad (3.18)
$$

which, not surprisingly, are the same kind of solutions found in Section 1.9.3 for the parametric evolution of open dust Universes. Here A, B and T are constants that can be fixed once the initial conditions have been fixed.

In the case of an EdS model for the background Universe, the density deficit evolves as

$$1 + \Delta(r, t) = \frac{\rho(r, t)}{\rho_b(r, t)} = \frac{9}{2} \frac{(\sinh\theta - \theta)^2}{(\cosh\theta - 1)^3}, \qquad (3.19)$$

The constants in Eq. (3.18) are:

$$\begin{aligned} A &= \frac{r_i}{2\Delta_i}, \\ B &= \frac{3}{4} \frac{t_i}{\Delta_i^{3/2}}, \\ T &= 0, \end{aligned} \qquad (3.20)$$

where r_i and t_i represent the initial coordinates of the perturbation at (Blumenthal et al., 1992). One can show that the linear initial density deficit is:

$$\Delta_i^L(\theta) = -\left(\frac{3}{4}\right)^{2/3} \frac{3}{5}(\sinh\theta - \theta)^{2/3}. \qquad (3.21)$$

The low-density environment expands faster than the Hubble flow, thus more rapidly with respect to the background Universe. As matter streams out of the voids, the value of density decreases asymptotically to $\delta = -1$. Since the density gradually decreases going towards the centre of a void, the matter near the centre moves outward faster than matter in proximity of the external boundaries. Shells that were initially close to the centre will ultimately catch up the shells further outside, until they eventually pass them. This phenomenon is called *shell-crossing*, and brings to the tendency of astrophysical objects to accumulate around voids, leading to the formation of sheets and filaments. From the shell-crossing, the evolution of the void can be described by a self-similar outward moving shell (Suto et al., 1984).

The solutions in Eq. (3.18) represent a family of trajectories labeled by r_i and parametrized by θ. We can find out when and where shell-crossing phenomenon first occurs by differentiating the parametrized solutions with respect to r and θ, and requiring that dr and dt vanish:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} dr/r \\ d\theta \end{bmatrix} = 0, \qquad (3.22)$$

where A_{ij} are functions of Δ_i and θ. To obtain non-zero solutions from this homogeneous system of linear equations, we must have $\det A = 0$ (Jennings et al., 2013). Thus we derive the shell-crossing condition:

$$\frac{\sinh\theta_{sc}(\sinh\theta_{sc} - \theta_{sc})}{(\theta_{sc} - 1)^2} = \frac{8}{9}. \qquad (3.23)$$

From the latter equation we obtain the angle θ_{sc} of the shell-crossing, which is:

$$\theta_{sc} \simeq 3.49 \ . \tag{3.24}$$

Moreover, we can show that, at shell-crossing event, the void has a precisely determined excess Hubble expansion rate (Sheth & van de Weygaert, 2004):

$$H_{sc} = \frac{4}{3} H(t_{sc}) \ , \tag{3.25}$$

where $H(t_{sc})$ is the Hubble parameter of the background Universe. Now, substituting θ_{sc} in Eq. (3.21) we find that at the shell-crossing event, the void interior has a relative density

$$1 + \delta_v^{NL} \simeq 0.2047 \ , \tag{3.26}$$

which implies that the void has expanded by a factor of $(1+\delta_v^{NL})^{-1/3} = 1.697$ in comoving radius. Note that these numbers do not depend on the size of the void. Moreover, from the latter relation we see that voids are only nearly nonlinear objects, since $|\Delta_{sc}| \simeq 0.8 < 1$.

The linear extrapolated quantity of Eq. (3.26) is:

$$\delta_v^L \simeq -2.71 \ . \tag{3.27}$$

This is the underdense counterpart of the critical density contrast δ_c found in Eq. (3.15). We can conclude that in the evolution of spherical voids, an expansion occurs, in contrast with the collapse for an overdensity. During this evolution, void borders become denser and the central parts reach lower density contrasts. Moreover, the outward expansion makes voids evolve towards a spherical geometry. As their nature to be emptier than the background Universe, voids experience super-Hubble expansion. The shell-crossing phenomenon is usually associated to the formation of a void and marks the transition from a quasi-linear towards a mildly nonlinear stage. All of these properties for the evolution of a spherical void are shown in Figure 3.1, which shows the evolution of a top-hat void density profile.

3.3 Press-Schechter formalism

The statistical description of the distribution of structures in the Universe has been developed by Press & Schechter (1974) (PS, hereafter), which considers DM haloes as spherically symmetric collapsed objects. This work has been extended then by Sheth & Tormen (2002) and Sheth et al. (2001), which
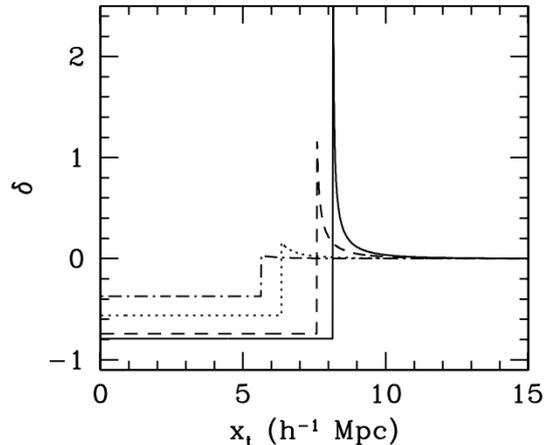
Figure 3.1: Spherical evolution of a void up to the shell-crossing time (*solid line*) in the case of a pure "top-hat" depression. Note how the void becomes emptier as it evolves. Credits to Sheth & van de Weygaert (2004).

included also the modelling of ellipsoidal DM haloes. Also, Lacey & Cole (1993) provided an analytical description, derived from PS formalism, for the merging rate of virialized haloes.

According to the spherical collapse model, regions with linear density contrast $\delta(\vec{x}, t) \geq \delta_c \simeq 1.686$ will collapse to produce DM haloes by the time t. Considering $\delta(\vec{x}, t) = \delta_{in}(\vec{x})\delta_+(t)$, we can also interpret this concept differently: regions with $\delta_{in}(\vec{x}) \geq \delta_c/\delta_+(t)$ will collapse to produce DM haloes by the time t. In the latter case we can consider the density field to be static (fixing its value to the one linearly extrapolated to our reference time), while the *collapse barrier* evolves with time. The aim is now to associate a mass to those haloes, and to use the statistics of the linear density field to infer the halo mass function, i.e. the (comoving) number density of haloes as a function of halo mass.

A possible strategy to achieve this goal is to assess that the number density of haloes with mass greater than some value M is equal to the comoving number density of peaks in the smoothed density field, thus:

$$n_h(> M) = n_{pk}(\delta_M) , \qquad (3.28)$$

where $n_{pk}(\delta_M)$ represents the number density of peaks above δ_c in the density field, smoothed on mass scale M.

This idea, firstly proposed by Doroshkevich (1970), was explored in Bardeen et al. (1986), a seminal paper known as "BBKS". Using elegant, clever math-

75

ematics they were able to compute the number density, clustering properties, shapes and density profiles of peaks in a smoothed Gaussian random field, all as a function of the peak height. However, this formalism leads to the so-called *peak-in-peak* problem. In few words, peaks that are part of higher peaks when filtering the density field with a greater radius are associated with those higher peaks, loosing the corresponding peak (i.e. halo). Because of this problem, the peak formalism of BBKS has largely been abandoned in favor of the less rigorous, but more successful PS formalism. The latter is based on the following postulate (*PS ansatz*):

The probability that $\delta_M > \delta_c(t)$ is the same as the mass fraction that at time t is contained in haloes with mass greater than M.

We will indicate this probability as $F(> M)$. Considering Eq. (2.42) as the Gaussian probability distribution for the filtered density contrast δ_M, which in linear theory remains Gaussian as it evolves, we can express the PS ansatz as:

$$
\begin{aligned}
F(> M) \equiv \mathcal{P}_{>\delta_c}(M) &= \int_{\delta_c}^{\infty} \mathcal{P}(\delta_M)\mathrm{d}\delta_M \\
&= \frac{1}{\sqrt{2\pi}\sigma_M} \int_{\delta_c}^{\infty} \exp\Big[-\frac{\delta_M^2}{2\sigma_M^2}\Big]\mathrm{d}\delta_M = \frac{1}{2}\mathrm{erfc}\Big[\frac{\delta_c}{\sqrt{2}\sigma_M}\Big] \,,
\end{aligned}
\tag{3.29}
$$

which directly depends on the mass M defined by the filter and also on the redshift, since σ_M is time-dependent. Here erfc is the complementary error function $\mathrm{erfc}(x) = 1 - \mathrm{erf}(x)$.

Since $\lim_{M\to 0} \sigma_M = \infty$ and $\mathrm{erfc}(0) = 1$, we see that the PS ansatz predicts that only $1/2$ of all matter in the Universe is locked-up in collapsed haloes. This may seem logical because only the regions that are initially overdense end up in collapsed objects. However, underdense regions can be enclosed within larger overdense regions, giving them a finite probability of being included in some larger collapsed object. This is the so-called *cloud-in-cloud* problem, and PS accounted for it by introducing an ad-hoc factor of 2, representing the accretion from underdense regions onto the overdensities:

$$
F(> M) = 2\mathcal{P}_{>\delta_c}(M)
\tag{3.30}
$$

Though Press & Schechter did not give a rigorous proof, it is possible to obtain analytically this forcing factor with the use of the *Extended PS* formalism, also known as *Excursion set* formalism (Bond et al., 1991), which we will describe in details in the Section 3.4.

3.3.1 The PS halo mass function

We are now ready to define the *halo mass function* (HMF, hereafter) in the PS formalism. Let us define it as the number of haloes with mass between M and $M + \mathrm{d}M$ per unit of comoving volume, which can be written as:

$$n(M,t)\mathrm{d}M = \frac{\bar{\rho}}{M}\frac{\partial F(>M)}{\partial M}\mathrm{d}M \ , \tag{3.31}$$

where $\frac{\partial F(>M)}{\partial M}\mathrm{d}M$ is equal to the fraction of mass that is locked up in haloes with masses in the range $[M, M + \mathrm{d}M]$. Therefore we obtain the desired number density by multiplying it for $\bar{\rho}/M$.

Using the PS ansatz and accounting for the additional factor given by the cloud-in-cloud problem we obtain the expression for the PS halo mass function:

$$\begin{aligned}
n(M)\mathrm{d}M &= 2\frac{\bar{\rho}}{M}\frac{\partial P_{>\delta_c}(M)}{\partial M}\mathrm{d}M \\
&= \sqrt{\frac{2}{\pi}}\frac{\bar{\rho}}{M^2}\frac{\delta_c}{\sigma_M}\left|\frac{\mathrm{d}\ln\sigma_M}{\mathrm{d}\ln M}\right|\exp\left(-\frac{\delta_c^2}{2\sigma_M^2}\right)\mathrm{d}M \ .
\end{aligned} \tag{3.32}$$

We stress again the fact that this expression for the HMF holds only for structures formed through a spherical collapse.

A more general and compact way to write the HMF is through its logarithmic differential:

$$\frac{\mathrm{d}n(M,z)}{\mathrm{d}\ln M} = \frac{\bar{\rho}}{M}f_{\ln\sigma}(\sigma_M, z)\frac{\mathrm{d}\ln\sigma_M^{-1}}{\mathrm{d}\ln M} \ . \tag{3.33}$$

where $f_{\ln\sigma}(\sigma_M, z)$ is the so-called *multiplicity function*, which depends on the assumptions of a given mass function model. For the PS mass function it is defined as:

$$f_{\ln\sigma}(\sigma_M, z) = \frac{\mathrm{d}\sigma_M}{\mathrm{d}\ln M} = \sqrt{\frac{2}{\pi}}\frac{\delta_c}{\sigma_M}\exp\left(-\frac{\delta_c^2}{2\sigma_M^2}\right) \ . \tag{3.34}$$

Let us consider now a power-law power spectrum, written as Eq. (2.30). As we already saw, this yelds to Eq. (2.51) for the mass variance and to Eq. (2.52) for the mass M_* that is formed when nonlinear regime is reached. Writing Eq. (2.52) in terms of $\delta_+(t)$ it is possible to obtain a new form for the mass variance:

$$\sigma_M \propto \left(\frac{M_*}{M}\right)^{(3+n)/6} = \left(\frac{M_*}{M}\right)^{\alpha} \ , \tag{3.35}$$

where n is the spectral index of the power spectrum, and $\alpha \equiv (3+n)/6$. The latter relation for σ_M simplifies the logarithmic derivative in Eq. (3.32), which becomes:

$$\left| \frac{\mathrm{d}\ln\sigma_M}{\mathrm{d}\ln M} \right| = \alpha \ . \qquad (3.36)$$

With few simple mathematical steps, it is possible to show that the PS halo mass function becomes:

$$n(M)\mathrm{d}M = \alpha\sqrt{\frac{2}{\pi}}\frac{\bar{\rho}}{M_c^2}\Big(\frac{M}{M_c}\Big)^{\alpha-2} \exp\Big[\Big(-\frac{M}{M_c}\Big)^{2\alpha}\Big]\mathrm{d}M \ , \qquad (3.37)$$

where we defined the *characteristic* mass as $M_c \equiv M_*\Big(\frac{2}{\delta_c^2}\Big)^{1/2\alpha}$. This mass represents the transition between two different trends of the HMF. Indeed for $M \ll M_c$ the mass function scales as $M^{\alpha-2}$, hence it shows a linear proportionality with the logarithm of this quantity. On the other hand, the MF possesses an exponential cut-off for $M \gg M_c$.

3.4 Excursion-set formalism

Considering the explicit dependence on the filter scale R in the smoothed density contrast $\delta_M(\vec{x}) \equiv \delta(\vec{x}, R)$, we can see the main weakness of the PS formalism: $\delta(\vec{x}, R)$ can be smaller or larger than the collapse barrier δ_c according to the considered smoothing scale R. Therefore the more diffuse patches (those that arise when filtering with smaller radii) may disappear due to the collapse of larger volumes. This is the already mentioned cloud-in-cloud problem, and can be solved analytically with the following excursion-set (ES) formalism (Bond et al., 1991).

Let us consider again the evaluation of the smoothed field $\delta(\vec{x}, R)$ at different filtering scales R, at the same point $\vec{x}$. To emphasize the fact that the variation is only considered on R, let suppress for now the dependence on the point $\vec{x}$, therefore $\delta = \delta(R)$. The idea behind the ES formalism is to find the largest scale R_{max}, for which no region crosses the collapse barrier, and then gradually to decrease the filtering radius until the highest peak of δ touches the barrier. To the latter we can associate a mass enclosed to that specific filter scale. Next, we reduce again the size of the filter until the next peak reaches the threshold, but only if it crosses it *for the first time*, so that the associated mass has not already been counted.

It is common to relate the filter scale R to the variance of the contrast density field, which is expressed as in Eq. (2.40). Let us adopt the following notation

$$S \equiv \sigma^2(R) = \frac{1}{2\pi^3}\int \mathrm{d}^3\vec{k}P(\vec{k})\hat{W}^2(\vec{k}, R) \ , \qquad (3.38)$$

as our radius variable, which can be associated also with the filtering mass M using the relation between mass and radius ($M \propto R^3$). For a hierarchical cosmological model such as the bottom-up scenario, we have that S is a monotonically decreasing function of R, so there is a clear, one-to-one relation between S and R. Also, decreasing R corresponds to increasing S.

The problem is to compute the probability that the first up-crossing of the barrier δ_c occurs on a specific scale R, i.e. the probability that the first up-crossing occurs between a value S and $S + \mathrm{d}S$. For each value of the filtering size R, i.e. for each value of S, the smoothed contrast will have a different value, when computed at the same point $\vec{x}$. Therefore to each point $\vec{x}$ corresponds a *trajectory* in the $(S, \delta(S))$ plane. Now, for $R \to \infty$, the variance, and hence S, tends to zero, thus all trajectories start from the same point $(S, \delta(S)) = (0, 0)$, the origin. It can be shown that these trajectories are purely *markovian* when a k-space top-hat filter is used. This means that the next step in the trajectory does not depends on the previous one, leading to a *random walk*.

As a result of their markovian nature, each trajectories has a mirrored version with respect to the collapse barrier, which is equally likely. Therefore, by double counting each trajectories, we can account naturally for the missing 2 factor of the PS formalism, solving the cloud-in-cloud problem. From these markovian trajectories we can infer the PS halo mass function, following the ES formalism ansatz:

> *The fraction of trajectories that crosses for the first time the collapse barrier $\delta_S = \delta_c$ at $S > S' = \sigma^2(R')$, is equal to the mass fraction that at time t resides in haloes with masses $M < M'$.*

3.5   Modern mass function models

The HMF provides the link between cosmological model and the number density of clusters. It essentially bundles the complicated nonlinear physics of halo formation into a simple analytical formula involving only the linear power spectrum and other linear-theory quantities. There are numerous mass function models in the literature, both analytical and semi-analytical, hence calibrated by means of cosmological simulations. In general, analytical models modify the PS mass function in order to make it more accurate by improving it over its approximations. Indeed the main approximation of the PS formalism is the spherical assumption for the halo collapse. This also leads to fewer high-mass objects and more low-mass objects with respect to those seen in numerical simulations. As already mentioned, Sheth et al. (2001) and Sheth & Tormen (2002) considered an ellipsoidal collapse, which

is thought to be closer to the reality. Indeed, DM haloes are modelled to take into account also tidal forces to which they are initially subject, according to the Zel'dovich approximation. In this case the critical (linear) overdensity for collapse can be computed by solving:

$$\frac{\delta_{ec}}{\delta_c} \approx 1 + 0.47\Big[5(e^2 \pm p^2)\frac{\delta_{ec}^2}{\delta_c^2}\Big]^{0.615} \;, \tag{3.39}$$

where δ_{ec} is the (linear) overdensity for an ellipsoidal collapse, $\delta_c \simeq 1.686$ is the spherical collapse threshold, while the parameters e and p characterise the asymmetry of the initial tidal field, and are functions of the tensor deformation's eigenvalues λ_i (see Section 2.5.1).

Other models consider more realistic profiles for the initial overdensity perturbation (*secondary infall models*), or a non-Gaussian distribution for the initial perturbations (Achitouv & Corasaniti, 2012), or a spherical collapse with non-radial orbits, hence by introducing an angular momentum in the collapse phase (White & Zaritsky, 1992).

3.5.1 Halo finder methods

The majority of HMF models are obtained through cosmological simulations, given the complexity of finding an analytical rigorous theory that describes the formation of nonlinear objects. In this latter case, a method to find bound gravitational structures in a system of particles is necessary in order to obtain a DM halo sample. This goal is achieved by means of specifically designed codes called *halo finders*, and the resulting halo catalogues may vary depending on the chosen method used to identify these structures (Knebe et al., 2013). The two standard techniques for finding haloes are the *Spherical Overdensity* (SO) method, firstly proposed by Press & Schechter (1974) and the *Friend-of-Friend* (FoF) method (Davis et al., 1985). Though other types of codes have been developed in the last decade, these two remain the foundation of nearly every finding code. The former is based on the definition of a spherical overdensity around density peaks, which are found by sorting particles by density. Considering a density peak, we can find haloes by growing a sphere around it, and stop when the mean density within the sphere is equal to a specific value of Δ. Δ indicates the overdensity within a sphere of radius R_Δ with respect to the mean density of the Universe at a given epoch:

$$\Delta = \frac{M_\Delta}{(4/3)\pi R_\Delta^3 \bar{\rho}_m} \;. \tag{3.40}$$

The overdensity threshold could be given also in terms of ρ_b or ρ_c, and typical values considered in the literature are $\Delta = 200, 500, 1000$.

On the other hand, in FoF method particles are "linked" together if their distance is smaller than a given value, called *linking length b*, typically expressed in units of the mean separation between particles. The value usually used for the latter is $b = 0.2$ (Frenk et al., 1988).

Moreover, since DM haloes formed through a hierarchical scenario, we expect a lot of substructures within them, called *subhaloes*. These structures can be examined in high-resolution N-body simulations. Finding these substructures is not trivial, mostly for the FoF algorithm. A related question is whether or not the particles belonging to a subhalo should also be associated to the host halo. This essentially leads to a decision on whether or not any single particle can be in more than one object at the same time.

3.5.2 Universality of the halo mass function

Although many works parametrise the HMF in terms of the mass variance σ, as we saw in Eq. (3.33), Sheth & Tormen (1999) and Courtin et al. (2011) showed that the HMF, for any time and for every cosmological model, assumes a universal functional form when mass and variance are parametrised by the use of the following scaled variable ν:

$$\nu = \frac{\delta_c(z)^2}{\sigma_M^2(M)} \ , \tag{3.41}$$

where $\delta_c(z)$ is the critical overdensity threshold computed in the linear theory, divided by the growth factor. With this parameterisation the HMF can be written as:

$$\nu f(\nu) = \frac{M^2}{\bar{\rho}} \frac{\mathrm{d}n}{\mathrm{d}M} \frac{\mathrm{d}\ln M}{\mathrm{d}\ln \nu} \ , \tag{3.42}$$

where:

$$\nu f(\nu) = A \left(1 + \frac{1}{\nu'^p} \right) \left(\frac{\nu'}{2\pi} \right)^{1/2} e^{-\nu'/2} \ , \tag{3.43}$$

with $\nu' = a\nu$. The parameters of this model are (a, p, A), calibrated using cosmological simulations, and they define the mass function cut-off at high masses, its form in the low mass range and its normalisation, respectively.

With the massive increase of available cosmological simulations, it has become possible to test the universality of these parameterisations (Tinker et al., 2008; Crocce et al., 2010; Watson et al., 2013). In particular, Despali et al. (2016) investigated whether the different halo definitions, i.e. the methods used to identify DM haloes, could lead to different calibrations of the mass function parameters. They used a halo finder based on the SO method,
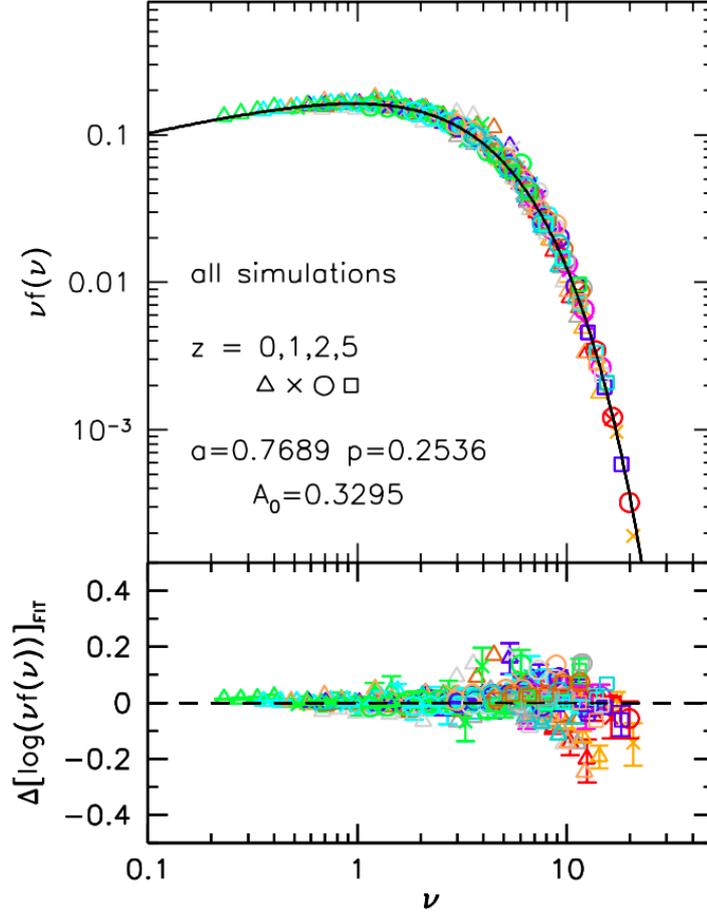
Figure 3.2: Universality of the halo mass function, when parametrised in terms of the variable ν and using a virial spherical overdensity Δ_{vir} to identify the DM haloes. The symbols represent the data coming from multiple simulations, evolved at different redshifts ($z = 0, 1, 2, 5$) and with different cosmologies (WMAP7, Planck13). The lower sub-panel show the residuals between these data and the best-fit model (*black solid line*), which is calibrated using all the sets of simulations having $z \leq 1.25$. Credits to Despali et al. (2016).

considering different density thresholds for the definition of haloes and studying the HMF from a set of 6 DM-only cosmological simulations (called Le SBARBINE), having different mass and spatial resolutions. These simulations are evolved with a Planck13 Cosmology (Planck Collaboration et al., 2014), though Despali et al. (2016) tested also the universality of the calibrated mass function parameters with different cosmological models. They made use of another set of simulations, ran assuming the WMAP7 cosmological parameters and characterised by lower resolution. They found that only if the threshold is set to be equal to the *virial overdensity* Δ_{vir} (considering $t_{vir} = 2t_{max}$, see Section 3.2.1) the universality is maintained, i.e. the independence of the model parameters of the simulation redshift and of the assumed cosmological model. This is no longer the case when the haloes are found with other overdensity thresholds. The best-fit values for parameters at $z = 0$ succeeded to model also higher redshifts data. Moreover, they provide best-fit values to the combination of data from Planck13 Cosmology up to $z = 1.25$. They showed that the latter are able to fit to an even greater level of accuracy also the mass function at high-redshift. The relation found for Planck13 Cosmology holds also for WMAP7 Cosmology, highlighting the HMF universality when considering different cosmological models. Finally, they obtained the best-fit values by fitting all the catalogues obtained from the different available cosmological simulations, with different redshifts and different cosmologies. In Figure 3.2 it is shown the mass function for the latter case, in which all the simulations are fitted simultaneously with a single best-fit model.

As already said, these universal properties for the HMF are obtained only when considering Δ_{vir} as the spherical overdensity threshold set in the halo finder algorithm. However, Despali et al. (2016) showed also an innovative method that allows to derive the HMF of halo catalogues build with different thresholds by rescaling the virial HMF. They provided fitting functions for the HMF parameters as a function of the overdensity Δ used, in order to predict the halo abundance for a wide range of halo definitions. In particular, they fitted the model parameters as a function of x, where the latter is defined as

$$x = \log(\Delta(z)/\Delta_{vir}(z)) , \tag{3.44}$$

obtaining the following relations for (a, p, A):

$$
\begin{aligned}
a &= 0.4332x^2 + 0.2263x + 0.7665 \\
p &= -0.1151x^2 + 0.2554x + 0.2488 \\
A &= -0.1362x + 0.3292 .
\end{aligned}
\tag{3.45}
$$

These relations can be used therefore to predict the HMFs at overdensity

thresholds different from Δ_{vir}.

3.5.3 The halo mass function as a cosmological probe

The HMF universality highlighted by Despali et al. (2016) is a fundamental feature to exploit in cosmological analyses. Indeed, its universal shape can be used to derive unbiased cosmological constraints by modelling data at any redshift and for any cosmological scenario.

The HMF is nowadays a powerful probe of Cosmology, notably for the high constraining power on the growth of cosmic structures. Large multi-wavelength surveys have been already exploited indeed to derive cluster number counts (or number density) and obtain constraints on the different cosmological parameters (Vikhlinin et al., 2009; Planck Collaboration et al., 2016c; Bocquet et al., 2019; Costanzi et al., 2019, 2021; Lesci et al., 2020), like Ω_m and σ_8. The dependence on Ω_m comes directly by the presence of $\bar{\rho}_m$ in the general form of the HMF. Moreover, the latter depends on the filtered mass variance σ_M^2, which is an integral of the power spectrum $P(k) = \delta_+^2 k^n T^2(k)$. The presence of δ_+, i.e. the *growth factor* (Eq. 2.27), introduces another dependence on Ω_m. Furthermore, the transfer function $T(k)$ and the spectral index n of the primordial fluctuation power-law, are integrated over the mass variance, hence the HMF does not depend strongly on them. On the other hand, the power spectrum normalisation, which is parametrised by σ_8 as explained in Section 2.4, and the growing mode δ_+ are not integrated, causing a strong dependency on them. Also, the two cosmological parameters Ω_m and σ_8 are degenerate, meaning that their variations cause similar effects to the HMF, and they will be the main focus of the analysis described in Chapter 5. Figure 3.3 shows the cosmological dependence of the HMF. In this case we considered, as an example, the most general fitting case of the Despali model, as a function of Ω_m and σ_8, at $z = 0$.

The expansion rate of the Universe could be also constrained, i.e. the value of the Hubble constant H_0, as well as the properties of the DE, such as its density parameter Ω_Λ as well as its equation of state w_{DE}. Indeed there are different models for the dark energy EoS, which could be constant in time as it is in the ΛCDM model, or could vary (*quintessence models*) (Chevallier & Polarski, 2001; Linder, 2003; Yoo & Watanabe, 2012). Having a Λ-density component that varies with redshift means that also the Ω_Λ varies, contributing to the inhibition of the structure growth. However, the variation of Ω_Λ could be due also to the variation of the critical density, which depends on H_0. Moreover, if H_0 increases, also the angular diameter distances increase and the number of galaxy clusters at a given redshift changes, causing other degeneracies between these parameters. Furthermore, the HMF shape is also

affected by the neutrino mass (Brandbyge et al., 2010; Marulli et al., 2011). One way to disentangle this set of degeneracies is by combining different probes, that are sensitive to the same parameters in different ways.
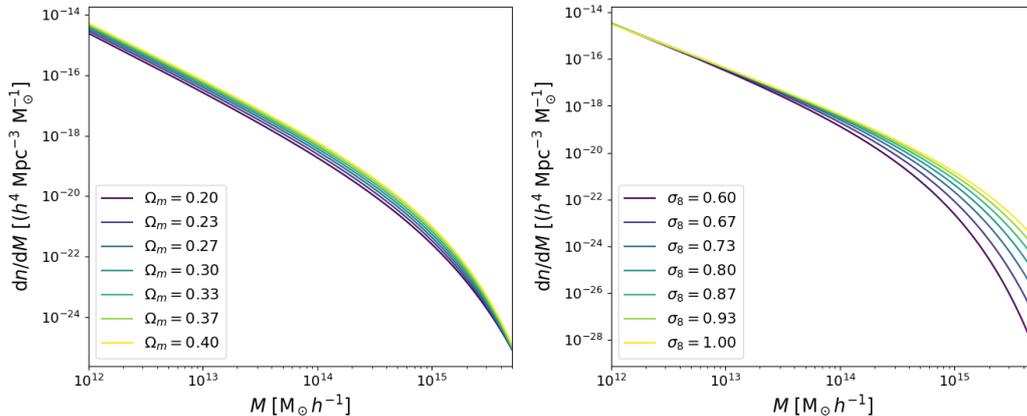

Figure 3.3: Dependence of the HMF model on the cosmological parameters Ω_m and σ_8 at $z = 0$. The parameters are considered in the range $\Omega_m \in [0.2, 0.4]$ and $\sigma_8 \in [0.6, 1]$.

3.6 The void size function

In this section we will show how the same excursion-set theory described in Section 3.4 can be applied to the underdensity case, demonstrating that it is possible to obtain the size function of voids identified in the DM distribution (Sheth & van de Weygaert, 2004; Jennings et al., 2013), that is the comoving number density of voids as a function of their size.

For the overdensity case, the HMF is obtained by determining the fraction of trajectories that cross the collapse barrier for the first time. We can extend the excursion-set formalism to underdensities by considering a *two-barrier problem*: one barrier is necessary to take into account the void formation and merging ($\delta_v \approx -2.7$), and the other one the void collapse ($\delta_c \approx 1.69$). This formalism, which for the overdensity case solved the cloud-in-cloud problem, aims in this case at solving the *void-in-void* problem, in which underdensities identified with a certain filtering radius may be embedded in other underdensities on larger scales. However, other phenomena may occur, such as the *void-in-cloud* scenario. According to the latter, a void may vanish due to the collapse of an ovederdense region on larger scales. In

other words, for the formation of a void it is necessary not only to reach a density contrast below a specific underdensity threshold, crossing the barrier δ_v, but also to avoid being extinguished by a larger collapsing overdensity, crossing the barrier δ_c. From now on we will refer to this model, proposed by Sheth & van de Weygaert (2004), as the SVdW model.

By the analogy between overdensities and underdensities described in this chapter, the void size function (VSF, hereafter) in linear theory can be written as Eq. (3.33) by replacing $\bar{\rho}/M = 1/V(r_L)$:

$$\left(\frac{\mathrm{d}n}{\mathrm{d}\ln r}\right)_{\mathrm{L}} = \frac{f_{\ln\sigma}}{V(r_L)}\frac{\mathrm{d}\ln\sigma^{-1}}{\mathrm{d}\ln r_L}, \tag{3.46}$$

where $V(r_L) = \frac{4}{3}\pi r_L^3$ is the volume of a spherical region of radius r_L, while the subscript L indicates that the quantities are obtained in linear theory. We can show that $f_{\ln\sigma}$ gives the fraction of random walks that cross δ_v, the threshold for the void formation, and have never crossed δ_c, the barrier for the overdensity collapse, at any scale larger than R. Following the SVdW model, the latter results in the following infinite series

$$f_{\ln\sigma} = 2\sum_{j=1}^{\infty} j\pi x^2 \sin(j\pi\mathcal{D})\exp\left[-\frac{(j\pi x)^2}{2}\right], \tag{3.47}$$

where

$$\mathcal{D} \equiv \frac{|\delta_v|}{\delta_c + |\delta_v|} \quad \text{and} \quad x \equiv \frac{\mathcal{D}}{|\delta_v|}\sigma . \tag{3.48}$$

Here we can see how the SVdW model takes into account the void-in-cloud problem by expressing $\mathcal{D}$ as a function of both δ_c and δ_v.

Jennings et al. (2013) provided an approximation for the infinite series $f_{\ln\sigma}$, accurate at the 0.2% level:

$$f_{\ln\sigma}(\sigma) = \begin{cases} \sqrt{\frac{2}{\pi}}\frac{|\delta_v|}{\sigma}\exp\left(-\frac{\delta_v^2}{2\sigma^2}\right), & x \leq 0.276 \\ 2\sum_{j=1}^{4} j\pi x^2 \sin(j\pi\mathcal{D})\exp\left[-\frac{(j\pi x)^2}{2}\right], & x > 0.276 . \end{cases} \tag{3.49}$$

From Eq. (3.26), we know that the void has expanded by a factor of ~ 1.7 when the shell-crossing occurs, therefore the void abundance becomes:

$$\left(\frac{\mathrm{d}n}{\mathrm{d}\ln r}\right)_{\mathrm{SVdW}} = \left.\frac{\mathrm{d}n_L}{\mathrm{d}\ln r_L}\right|_{r_L=r/1.7} . \tag{3.50}$$

This translates into a shift of the linear model towards larger radii, without any change in amplitude, which means that in the SVdW model the comoving

number density of voids is conserved by passing to the nonlinear regime. This property is derived directly from the spherical evolution model, in which it is assumed that voids evolve as isolated objects.

3.6.1 The volume-conserving model

The assumption that voids only evolve in radius keeping their comoving number density fixed is invalid considering large voids. In contrast to haloes, which tend to collapse towards a confined configuration, voids tend to expand and overlap with each other, changing their abundance with time. We can compute the cumulative fraction $\mathcal{F}(R)$ of volume occupied by voids having a radius greater than R as:

$$\mathcal{F}(R) = \int_R^\infty \frac{\mathrm{d}r}{r} V(r) \frac{\mathrm{d}n}{\mathrm{d}\ln r} . \tag{3.51}$$

This quantity exceeds unity for the radii of interest, meaning that the fraction of volume occupied by voids in nonlinear theory exceeds the total volume of the Universe, which is clearly an unphysical condition.

In order to account for this problem, Jennings et al. (2013) considered instead the conservation of both the volume fraction and the shape of the void abundance function in the transition from the linear to the nonlinear regime. This can be achieved by considering Eq. (3.51) in the linear theory and imposing it to be equal to the same quantity computed in the nonlinear regime:

$$V(r)\mathrm{d}n = V(r_L)\mathrm{d}r_L(r) . \tag{3.52}$$

As result, the nonlinear VSF obtained is the following:

$$\left(\frac{\mathrm{d}n}{\mathrm{d}\ln r} \right)_{V\mathrm{d}n} = \frac{V(r_L)}{V(r)} \frac{\mathrm{d}n_L}{\mathrm{d}\ln r_L} \frac{\mathrm{d}\ln r_L}{\mathrm{d}\ln r} . \tag{3.53}$$

The latter is the so-called *volume-conserving model* (Vdn, hereafter), in which the void abundance varies in the transition between r_L to r as they blend together expanding towards larger scales, conserving their total volume instead of the number density.

As already mentioned, the model depends on two thresholds δ_v and δ_c. The former can be fixed to the shell-crossing value while the latter is usually left to vary between the turn-around and the collapse values $1.06 \leq \delta_c \leq 1.69$, since both can be considered acceptable assumptions for the void-in-cloud scenario. Figure 3.4 shows the comparison between the VSF models described in these last two sections, with the explicit dependency on the collapse barrier δ_c, which affects significantly the VSF only for $r \lesssim 1$ Mpc/h.
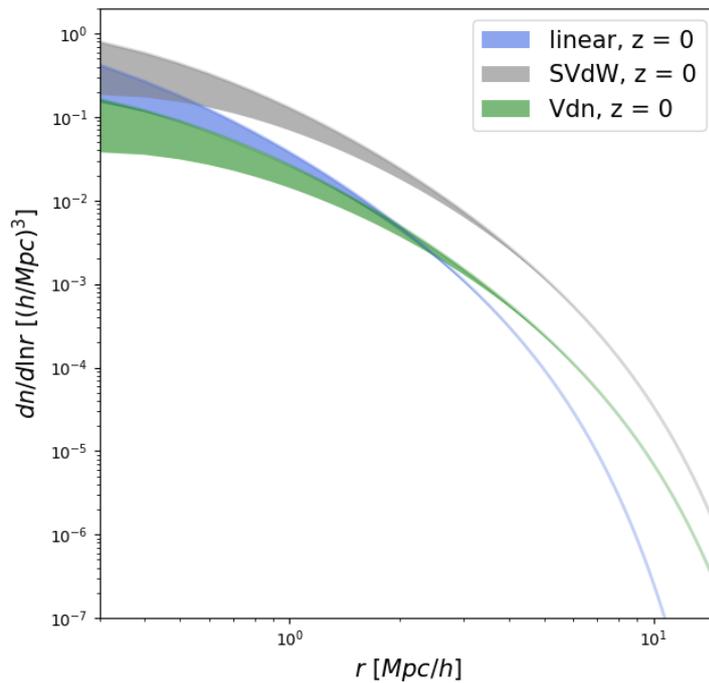
Figure 3.4: Different models for the VSF, computed at $z = 0$. The shaded area are obtained by considering the dependency on the linear collapse threshold, which is allowed to vary in the range $1.06 \leq \delta_c \leq 1.69$. The cosmological parameters used in this plot are those of WMAP7 Cosmology.

These scales that are commonly excluded from the analysis of the void abundance and profiles, given the extremely high spatial resolution required to model accurately voids in this radius range. Therefore the dependency of the VSF model on the overdensity threshold δ_c can be neglected when considering voids traced by biased tracers like galaxies or clusters, which are indeed characterised by larger sizes.

In Figure 3.5 we compare the VSF models previously described with void abundance measured in the Coupled Dark Energy Cosmological Simulations (CoDECS) (Baldi et al., 2010; Baldi, 2012), at redshift $z = 0$. The simulation box analysed has a side length of 10^3 Mpc$/h$ and is evolved with a ΛCDM Cosmology. Note how the SVdW model predicts an over-abundance of voids for all the void radii considered. A possible strategy to obtain a better agreement with the measured abundances using the SVDW model is to consider δ_v as a free parameter, fine-tuning its value by means of cosmological simulations. However, this severally affects the possibility of using the VSF as a cosmological probe. On the contrary, Jennings et al. (2013) showed that the Vdn model does not require this fine-tuning, as long as the void sample is prepared through a proper selection and re-shaping of the underdensities identified by means of the void finding algorithm.

Finally, the VSF models shown in this chapter are effective to predict the abundance of voids detected in the distribution of simulated DM particles. In order to analyse more realistic void samples, we have to consider voids identified in the distribution of biased tracers such as DM haloes and galaxies. In this case we have to take into account the bias of the mass tracers, parametrising its effects in the theoretical model (see Chapter 3.6.4).

3.6.2 Void finders

Given a model that can accurately predict the abundances of voids identified in the matter distribution, this has to be compared to the measured void number counts coming from cosmological simulations or observed data catalogues. However, the detection of cosmic voids is not trivial, for two main reasons. First, being voids essentially regions devoid of matter, their position and shape have to be reconstructed by using luminous tracers, that in turn are difficult to find in the most underdense zones. Second, since there is not a unique way to define what a cosmic void is, different algorithms of void finding have been proposed during the years, without following a common prescription to assign the centre and the radius of each void.

Therefore, a variety of different void finders have been implemented in the last decade, following different definitions of void (Colberg et al., 2008). Following the classification proposed by Lavaux & Wandelt (2010), void find-
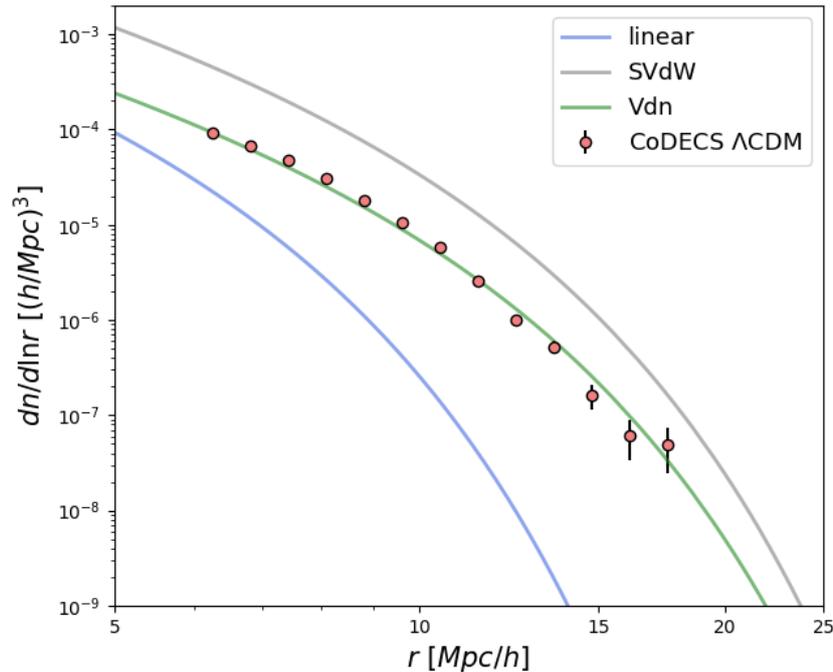
Figure 3.5: Comparison between the measured void abundance (red dots) and different theoretical VSF models (solid lines). The data are obtained extracting cosmic voids from the Coupled Dark Energy Cosmological Simulations (CoDECS) (Baldi, 2012), considering a ΛCDM model at redshift $z = 0$. The parameters of the models are set to $\delta_v = -0.795$ and $\delta_c = 1.686$. We note how only the Vdn model, represented with a green line, can accurately predict the size distribution of the detected void sample, while the *linear* and the *SVdW* models, in blue and grey respectively, are not in agreement with the measured abundances.

ers can be divided into three main categories, based on *density, geometrical* and *dynamical* criteria. The first includes void finder algorithms that identify cosmic voids as underdense regions with central density lower than a given fraction of the mean cosmic density, $\bar{\rho}$. Such algorithms are extensively tested in the past, for example in Elyiv et al. (2013) and Micheletti et al. (2014), but are currently rarely used. The second class considers cosmic voids as composed of underdense geometrical patches of similar shapes, such as spheres or polyhedra, detected in the 3D distribution of tracers (Shandarin et al., 2006; Neyrinck, 2008; Platen et al., 2007; Sutter et al., 2015). The third class of void finders, instead, identifies voids dynamically by finding the gravitationally unstable points in the distribution of tracers. In this case, matter tracers are not used to reconstruct the underlying distribution of matter, but as test particles of the cosmic velocity field (Forero-Romero et al., 2009; Lavaux & Wandelt, 2010; Elyiv et al., 2015). The first two classes of void finders suffer from large *shot-noise* errors, caused by the few objects used to identify voids as underdense volumes. The third class of void finders, instead, has the advantage that tracers are defined in Lagrangian coordinates, which remarkably reduces the shot noise.

In order to extract void catalogues from the distribution of galaxies, we make use of the publicly available Void IDentification and Examination toolkit (VIDE) code (Sutter et al., 2015), a void finder based on geometrical criteria, which implements an enhanced version of the ZOnes Bordering On Voidness (ZOBOV) algorithm (Neyrinck, 2008). The latter finds voids in a 3D distribution of particles, without introducing any free parameter or assumptions about the shape of voids. It consists of three main steps, which are also illustrated in Figure 3.6:

1. *Density field estimation.* In this first step, ZOBOV associates to every particle in the tracer distribution a *Voronoi cell*, consisting of all the points in the space closest to that particle than to any other. The whole set of cells is called *Voronoi tasselation.* The value of the density in a specific point can be computed as the inverse of the cell's volume $V(i)$ surrounding the i-th particle;

2. *Partition of particles into zones.* The Voronoi tasselation gives a natural set of neighbours for each particle. In this step, density minima are found, that are given by the Voronoi cells whose density is lower than all the respective neighbours. Consecutively, starting from density minima, cells are merged together until a cell with lower density with respect to the previous merged cells is encountered;

3. *Merging of zones into voids.* Although the zones found with the previ-

ous steps could be considered already as voids, ZOBOV also performs the so-called *watershed technique* (Platen et al., 2007), to construct a hierarchy of voids and sub-voids. Following this technique, a void is defined in analogy to a catchment basin in the density field. The water level of a given basin is left to rise until a saddle point is encountered, eventually causing the water to flows into deeper basins. 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.

Since voids are not perfectly spherical by nature, there is not a unique definition of their centre. In practice, there are two main options: to define the centre as the lowest density minimum point inside the void, or through the barycentre obtained from the distribution of tracers present within the void. VIDE follows the latter option, by associating to each of them a centre defined by means of the following volume-weighted barycentre:

$$\vec{X}_c = \frac{\sum_{i=1}^{N} \vec{x}_i V_i}{\sum_{i=1}^{N} V_i} \ . \tag{3.54}$$

Here the summation is computed over the N cells that define each void and $\vec{x}_i$ are the coordinates of the i-th tracer inside its corresponding Voronoi cell. In addition, VIDE associates an effective radius to every void as the radius of a sphere having as volume the summation of each Voronoi-cell composing the void:

$$R_v = \left(\frac{3}{4\pi} \sum_{i=1}^{N} V_i \right)^{1/3} \ . \tag{3.55}$$

The void catalogues we will consider in this work are then composed of the 3D comoving coordinates of the centres of voids, computed as Eq. (3.54), along with their effective radius.

3.6.3 Cleaning algorithm

The void size functions described in the previous sections model the size distribution of voids considering them as underdense, spherical, non-overlapped regions that have gone through shell-crossing (or, more in general, with a specific internal density contrast). It is therefore necessary to adopt the same void definition when building the void catalogue or to clean properly the void sample detected with whatever method of void finding, as firstly proposed by Jennings et al. (2013). Following the latter approach, Ronconi & Marulli (2017) implemented a cleaning algorithm in the CosmoBolognaLib (Section
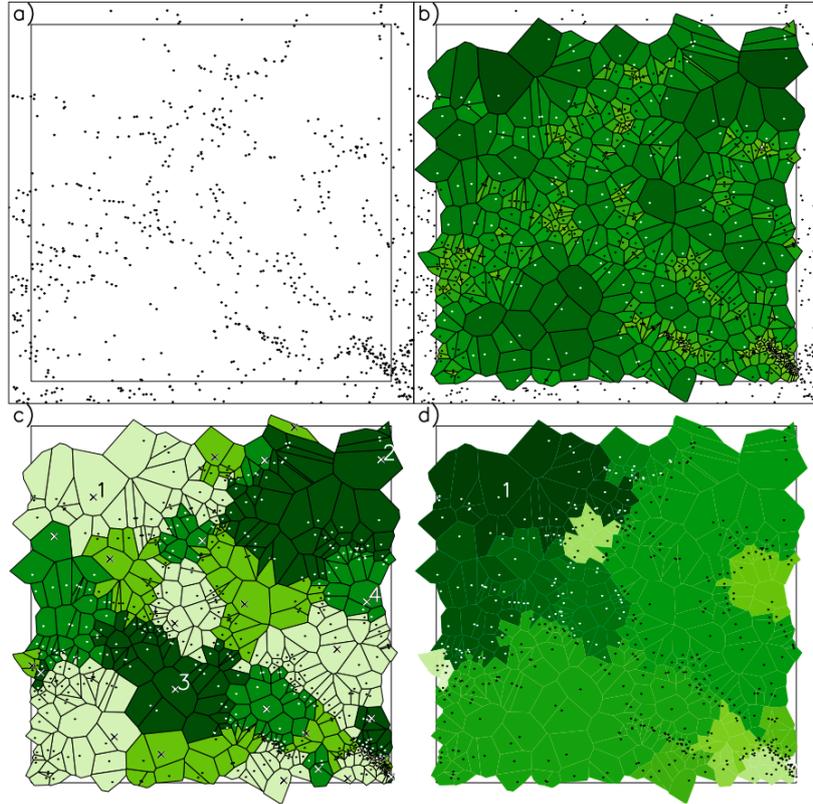
Figure 3.6: Representation of ZOBOV's steps applied to a distribution of galaxies in a 2D slice of the Millennium simulation (Springel et al., 2005). The *upper-left* panel shows the spatial 2D distribution of galaxies, represented with black dots. The *upper-right* panel shows the Voronoi tessellation, with colours shaded accordingly to the area of the Voronoi cells. The *lower-left* panel represents the second step of ZOBOV, which groups cells together, forming assembled zones. Here the density minima are marked with black crosses. Lastly, in the *lower-right* panel it is shown the result of the watershed technique for the void marked with "1", with zone colours going from darkest to lightest according to the step at which they are added to the void. Credits to Neyrinck (2008).

[4.1](#)) that selects and rescales the detected underdense regions in order to ensure a proper comparison with the model predictions. The algorithm is divided in three main steps:

1. *Removing non-relevant objects.* The algorithm starts by removing those underdensities that have an effective radius outside a range $[r_{min}, r_{max}]$, selected by the user. Also, underdensities having a central contrast greater than $(1 + \delta_v^{NL})\bar{\rho}$ can be removed from the catalogue, where δ_v^{NL} is a nonlinear underdensity threshold given as input and $\bar{\rho}$ is the mean density of the tracer catalogue;

2. *Rescaling voids.* The second step is aimed at rescaling the effective radius of the underdensities to make them embedding a specific density contrast, thus the same threshold considered in the Vdn model (see Section [3.6.1](#)). By treating voids as growing spheres, the effective radius obtained after this step corresponds to the largest distance from the void centre enclosing an underdensity contrast of δ_v^{NL}, selected by the user;

3. *Checking for overlaps.* The rescaled voids obtained with the previous step are then scanned one-by-one, checking for overlaps. When two voids do overlap, i.e. when the distance between their centres is less then the sum of their effective radii, the one with highest central density is rejected.

The impact of the cleaning procedure on the void abundances can be seen in Figure [3.7](#). Here the data are obtained from a ΛCDM N-body simulation with 256^3 DM particles and a boxside length of 128 Mpc/h and are compared to the Vdn model ([Ronconi et al., 2019](#)). This procedure has been tested systematically in [Ronconi et al. (2019)](#), which furthermore applied it to a set of cosmological simulations having different spatial resolutions and boxside lengths, considering also different redshifts. The same cleaning procedure will be applied in this work to the void catalogues obtained from VIDE, built upon the galaxy distribution given by the Magneticum simulation.

3.6.4 The bias inside cosmic voids

As already mentioned, the theoretical VSF models described up to now consider only the abundance of voids, as a function of their radius, detected in an unbiased tracer distribution, namely using DM particles catalogues. However, dealing with real data catalogues coming from redshift surveys we rarely have access to the total matter distribution of the Universe. We can in general study voids traced by the distribution of biased tracers, like galaxies and
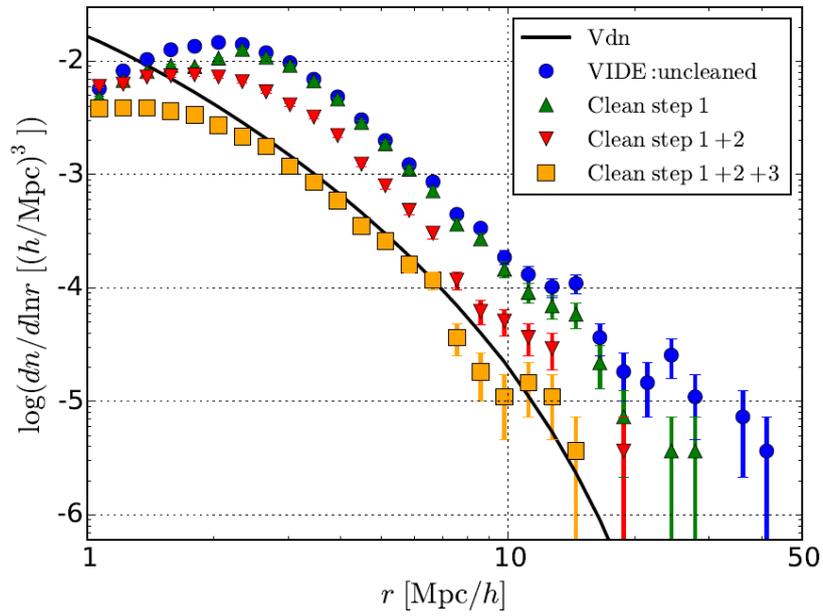
Figure 3.7: Cleaning procedure effects on the measured size function of voids identified in a ΛCDM N-body simulation. The different symbols show the VSF after each step of the procedure, as indicated by the labels, while the Vdn model is represented with a solid black line. Credits to Ronconi et al. (2019).

galaxy clusters. Understanding the link between the statistical properties of voids identified in the DM particle field and those found in a biased tracer field is therefore of fundamental importance to extract reliable cosmological constraints.

In order to measure the VSF from a biased tracer distribution it is necessary to rescale the threshold $\delta_{v,\text{DM}}^{NL}$, by means of the bias measured inside voids. We follow the same procedure described in Contarini et al. (2019, 2020), which measured the size function for voids identified in the distribution of DM haloes, which are biased tracers of the DM particle field, using a set of cosmological simulations. They first made the assumption that voids identified in the biased tracer distribution have the same centre with respect to those identified in the total density field. From the definition of bias described in Section 2.4.2, we can deduce then that voids identified in the biased distribution have an embedded density contrast higher (i.e. closer to positive values) with respect to the contrast of those found in the unbiased one. It follows then that the radii of the spherical voids predicted by the Vdn model can be rescaled in order to match the radii of voids identified in the biased tracer distribution. As a consequence, the characteristic threshold of the Vdn model can be modified to encapsulate the expected variation of the density contrast inside voids that, in nonlinear theory is simply expressed by:

$$\delta_{v,\text{tr}}^{NL} = b \ \delta_{v,\text{DM}}^{NL} \ , \qquad (3.56)$$

where $\delta_{v,\text{DM}}^{NL}$ represents the density contrast embedded by voids identified in the DM particle field, while $\delta_{v,\text{tr}}^{NL}$ is the density contrast inside voids in the biased tracer field and b is the bias characterising the tracers located within these voids. With this prescription, it is evident that considering the theoretical shell-crossing event, thus fixing $\delta_{v,\text{DM}}^{NL} = -0.795$, and values $b > 1$ for the bias factor, the corresponding threshold $\delta_{v,\text{tr}}^{NL}$ could easily becomes less than -1. This leads to the fact that the shell-crossing may not even occur for voids traced by biased objects. Following Contarini et al. (2019) and Contarini et al. (2020) we adopted therefore the approach to fix the value of $\delta_{v,\text{tr}}^{NL}$ to an arbitrary negative density contrast and to compute consistently the value of $\delta_{v,\text{DM}}^{NL}$, that will be inserted in the Vdn model. In particular, we rescale the voids found in tracer catalogues by means of the cleaning algorithm (see Section 3.7), such that the spherically-averaged contrast that they contain is $\delta_{v,\text{tr}}^{NL} = -0.7$. This value is chosen to be not too negative, in order to avoid rescaling too many underdensities to radii below the spatial resolution of the tracer catalogue and, at the same time, not too close to 0 so that the identified underdense regions can be still classified as voids. Consequently, we re-parametrise the Vdn model using a different value for the threshold $\delta_{v,\text{DM}}^{NL}$, computed as $-0.7/b$. However, we recall that the theoretical

VSF requires this threshold to be expressed in linear theory. This can be obtained with the following fitting formula provided by Bernardeau (1994):

$$\delta^L_{v,\text{DM}} = \mathcal{C}[1 - (1 + \delta^{NL}_{v,\text{DM}})^{-1/\mathcal{C}}] \ , \tag{3.57}$$

with $\mathcal{C} = 1.594$, which is exact for cosmologies with $\Lambda = 0$, and very precise also for other values of Λ, especially applied to the negative density contrasts. The resulting outcome of this procedure is equivalent to expand the radii of voids predicted by the Vdn model, so identified in the unbiased distribution, in order to make them match the same radius of the ones identified in the tracer field (all embedding the same density contrast -0.7 by construction). With this re-parameterisation of the threshold δ_v we are in practice modifying the Vdn model to predict the VSF of voids identified in the distribution of tracers having whatever bias factor.

Contarini et al. (2019) showed that the bias b, characterising the tracers inside cosmic voids, does not coincide with the one computed on large scales b_{eff}. The latter is the linear effective bias and can be inferred, for example, by measuring the 2PCF. The former, instead, can be estimated, as done by these authors, as the ratio between $\delta^{NL}_{v,\text{tr}}$ and $\delta^{NL}_{v,\text{DM}}$:

$$b_{\text{punct}} = \left\langle \frac{\delta^{NL}_{v,\text{tr}}(R = R_{\text{eff}})}{\delta^{NL}_{v,\text{DM}}(R = R_{\text{eff}})} \right\rangle \ , \tag{3.58}$$

whose name is related to its peculiarity of being computed *punctually*, at a distance $R = R_{\text{eff}}$ from the void centres. Since we have generally not access to the total matter density field when analysing real data catalogues, the value of b_{punct} is directly measurable only using cosmological simulations. Therefore it is useful to search for a relation between the latter and the effective bias on large scales, b_{eff}. This relation, which we will denote as $\mathcal{F}(b_{\text{eff}})$, has been calibrated by means of ΛCDM simulations at different redshifts in Contarini et al. (2019) and Contarini et al. (2020) using different types of mass tracers: FoF DM haloes (identified with a FoF halo finder) in the former, and considering $200c$ and $500c$ haloes (identified with a SO halo finder, fixing the overdensity to the values $\Delta = 200\rho_c$ and $500\rho_c$, respectively) in the latter. The results of their analysis are shown in Figure 3.8, and the resulting linear relations are:

$$\begin{aligned}
\mathcal{F}(b_{\text{eff}}) &= (0.85 \pm 0.01)\, b_{\text{eff}} + (0.42 \pm 0.01), \quad \text{for FoF haloes} \\
\mathcal{F}(b_{\text{eff}}) &= (0.87 \pm 0.02)\, b_{\text{eff}} + (0.36 \pm 0.03), \quad \text{for 200c haloes} \\
\mathcal{F}(b_{\text{eff}}) &= (0.82 \pm 0.02)\, b_{\text{eff}} + (0.37 \pm 0.02), \quad \text{for 500c haloes} \ .
\end{aligned} \tag{3.59}$$

Since in this work we will analyse cosmic voids traced by different objects (i.e. galaxies and not DM haloes) we will perform a new calibration of the
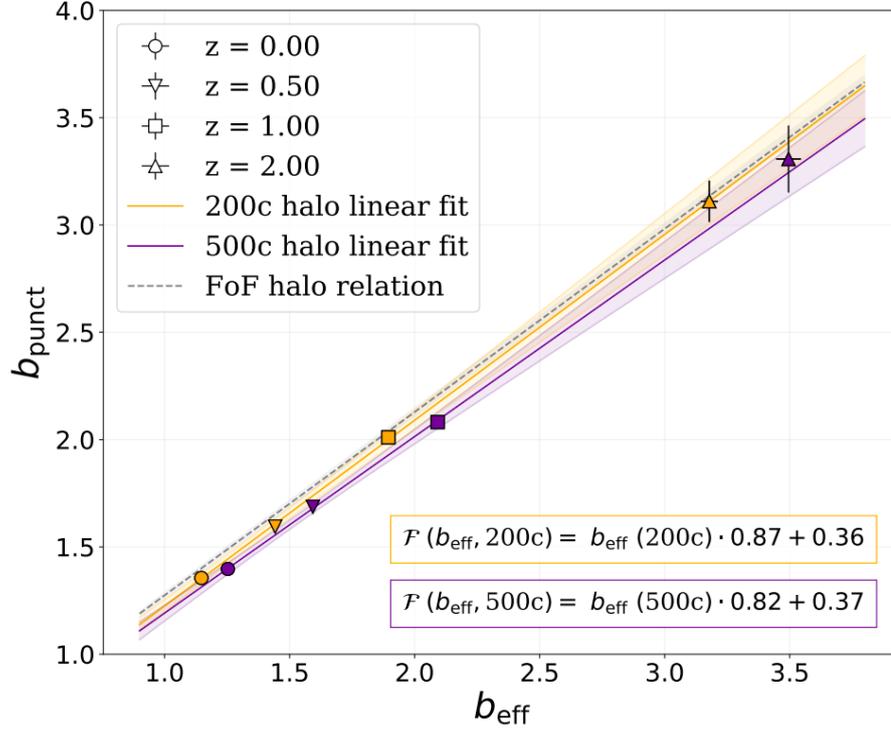
Figure 3.8: Linear relations between the effective bias (b_{eff}) and the punctual bias computed inside voids (b_{punct}), at different redshifts, calibrated using DM halo catalogues built with different halo finders. The orange and violet markers represent the data obtained in Contarini et al. (2020) for $200c$ and $500c$ haloes, respectively, and the solid lines represent the associated relations found with a linear fit of these data. The dashed gray line represents the relation calibrated in Contarini et al. (2019) using haloes identified by means of a FoF method. Credits to Contarini et al. (2020).

relation $\mathcal{F}(b_{\mathrm{eff}})$. Contarini et al. (2020) also demonstrated the negligible dependence of this calibration on the considered cosmological model, that is a fundamental prerogative to exploit the VSF as cosmological probe.

3.6.5 The void size function as a cosmological probe

Cosmic voids, the most large underdense regions of space in the Universe, are promising laboratories to extract cosmological constraints. Their unique low density interiors and very large sizes make them powerful tools to investigate elusive components as massive neutrinos (Villaescusa-Navarro et al., 2013; Massara et al., 2015; Schuster et al., 2019; Kreisch et al., 2019; Contarini et al., 2020), which are predominant in these regions, or to study the properties of the DE (Bos et al., 2012; Pisani et al., 2015; Verza et al., 2019) and the effects of modified gravity theories (Spolyar et al., 2013; Barreira et al., 2015; Contarini et al., 2020), to which they are very sensitive, as well as to test primordial non-Gaussianity (Chan et al., 2019) and physics beyond the standard model (Peebles, 2001; Yang et al., 2015; Baldi & Villaescusa-Navarro, 2016). In the previous sections we saw how the VSF, as well as the HMF, is defined through the mass variance σ_M. For the same arguments described in Section 3.5.3, we find that the VSF is particularly sensitive to the total matter content of the Universe, Ω_m, and the amplitude of the power spectrum, σ_8. These parameters are characterised by strong degeneracy, and their variation rules the growth of cosmic structures, causing the damping or the enhancement of the evolution of cosmic voids. In this work we will indeed exploit the VSF to derive constraints on these cosmological parameters, showcasing its constraining power in combination with the HMF, with the aim of providing a powerful tool to test the cosmological scenario. The dependence of the Vdn size function model on Ω_m and σ_8, at $z = 0$ is shown in Figure 3.9.
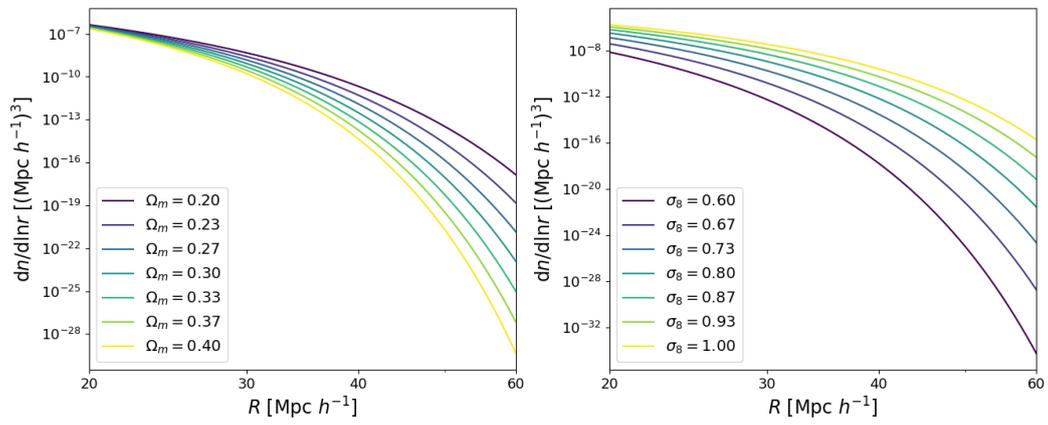
Figure 3.9: Dependence of the VSF model on the cosmological parameters Ω_m and σ_8 at $z = 0$. The considered parameters are in the range $\Omega_m \in [0.2, 0.4]$ and $\sigma_8 \in [0.6, 1]$.

Chapter 4

Algorithms for independent probe combination

In the previous chapter we described how the HMF and the VSF can provide constraints on the cosmological parameters. In general, different probes constrain the cosmological model differently, and a simultaneous analysis may help breaking the degeneracies between cosmological parameters constraints. Therefore, by properly combining different cosmological probes we can obtain more precise and accurate constraints, which is of fundamental importance for solving the current tensions in the ΛCDM framework (see Section 1.10.1). Previous works focused on the combination of different cosmological data-sets considering, for example, observations of CMB anisotropies with LSS surveys (Webster et al., 1998; Gawiser & Silk, 1998; Bridle et al., 1999), SN Ia with Cepheids distance measurements and CMB (Lahav, 2001), as well as SN Ia, BAO feature, weak gravitational lensing and galaxy clustering (see e.g. DES Collaboration, 2019).

In this chapter we present the implementation of new codes in the CosmoBolognaLib C++/Python libraries (see Section 4.1), aimed at providing new numerical tools to perform a joint cosmological analysis on different data-sets, i.e the *probe combination*. The cosmological parameter estimation that will be performed applying the presented codes is based on Bayesian statistics, which will be described in Section 4.2. In particular, we implemented three different algorithms for the probe combination, described in Sections 4.3.1, 4.3.2 and 4.3.3. These new implemented codes will be applied for the study of the mass function of galaxy clusters and the size function of cosmic voids, that will be presented in the next chapter.

4.1 CosmoBolognaLib

The CosmoBolognaLib (CBL, hereafter) is a large set of free software numerical libraries, specifically designed for cosmological calculations (Marulli et al., 2016). They are based on a object-oriented programming and are totally written in C++, but they offer also a wrapped Python version. The CBL are well suited for working with extragalactic source catalogues, both real and simulated, as well as for performing statistical analyses and extracting cosmological constraints. They provide a useful tools to compute all the HMF and all the VSF models described in Section 3.6. All of these models are implemented as public members of the internal class `cbl::Cosmology`, where `cbl` is the general namespace. With this class, the user can easily define a cosmological model, setting the values for each cosmological parameter, with which it is possible to compute the theoretical number densities for DM haloes and voids. Furthermore, in the CBL are implemented different Bayesian inference methods (see Section 4.2), as well as methods to perform the Markov Chain Monte Carlo (MCMC) posterior sampling (Section 4.2.2).

The CBL is a living project, constantly extended and optimised, and has already been used by several astrophysicists and cosmologists around the world. One of the main goal of this Thesis work is to improve these libraries by adding new codes for probe combination, providing a publicly available and user-friendly tool to perform joint cosmological analyses.

4.2 Bayesian parameter estimation

One of the main goals of observational Cosmology is to use astronomical information, i.e. measurements of physical properties of our Universe, to derive precise constraints on the parameters of the cosmological model.

Since every measure of the statistical proprieties our Universe (see Section 2.4) is necessarily referred to a single realisation of the Universe, every prediction in Cosmology is essentially based on the statistical inference. The statistical "world" is indeed divided in *frequentists*, which conceive probability as the frequency of occurrence of some events, and *Bayesians* that define probability as the degree of belief in a certain hypothesis.

Let us assume that we collect some data, $\mathcal{D}$, and we want to interpret them in terms of a model $\mathcal{M}$, dependent on a set of parameters $\vec{\theta}$, which we want to estimate. Through the Bayesian inference approach we can determine, given the observation of the data, the whole underlying probability distribution of $\vec{\theta}$: the so-called posterior distribution, $\mathcal{P}(\vec{\theta}|\mathcal{D})$, which represents our degree of belief about the value of the parameters given a specific

data-set. We can express it by means of the *Bayes Theorem*:

$$\mathcal{P}(\vec{\theta}|\mathcal{D}) = \frac{\mathcal{L}(\mathcal{D}|\vec{\theta})\mathcal{P}r(\vec{\theta})}{\mathcal{P}(\mathcal{D})} \,. \tag{4.1}$$

This theorem relates the posterior distribution of $\vec{\theta}$ to the likelihood distribution, $\mathcal{L}(\mathcal{D}|\vec{\theta})$, the prior distribution, $\mathcal{P}r(\vec{\theta})$, and finally the evidence, $\mathcal{P}(\mathcal{D})$. The likelihood represents the probability to obtain the observed data, given a certain value of the parameters, while the prior is our initial degree of belief in the value of $\vec{\theta}$, without having information on the data. The prior is often set to be flat, thus *non-informative*, by specifying a physical-motivated range of values for the parameters that we want to determine. In these cases the posterior distribution will be directly proportional to the likelihood. Lastly, the evidence is a normalising constant that ensures that the posterior distribution is normalised to unity:

$$\mathcal{P}(\mathcal{D}) = \int \mathcal{L}(\mathcal{D}|\vec{\theta})\mathcal{P}r(\vec{\theta})\mathrm{d}\vec{\theta} \,. \tag{4.2}$$

The evidence plays an important role in *model selection* when more than one underlying theoretical model is being considered and one wants to choose which model gives the most reasonable representation of the data. The data will be much more likely to be explained by a model having a higher value of the evidence, with respect to a different one. For example, one can consider, in addition to the ΛCDM framework, the free-variation of the DE equation of state parameter w_{DE} (a model indicated as wCDM), and then consider if this inclusion represent the data more accurately. In this fashion, the evidence incorporates the Occam's razor concept, i.e. that a simpler theory having a more compact parameter space will generally have a larger evidence with respect to a more complicated theory, unless the latter is significantly better in explaining the data. However, in this Thesis work we are not interested in the comparison of different cosmological models on the same data-set, but on the joint analysis of different data-sets with distinct models within the same cosmological framework. For this reason, we can neglect the evidence in Eq. (4.1), and consider the non-normalised posterior distribution.

In a general case, we have a vector of parameters $\vec{\theta} = \{\theta_1, \dots, \theta_N\}$ for a given model, hence the posterior distribution expressed as in Eq. (4.1) is a *multivariate probability distribution*, i.e. the *joint* distribution of two or more parameters. It is possible to derive the probability distribution of a single parameter θ_1, regardless of the value of the others, by integrating the posterior distribution over them:

$$p(\theta_1|\mathcal{D}) = \int p(\vec{\theta}|\mathcal{D})\mathrm{d}\theta_2 \dots \mathrm{d}\theta_N \,, \tag{4.3}$$

where $p(\theta_1|\mathcal{D})$ is called the *marginal posterior* and the process is called *marginalisation*. The latter is equivalent to the projection of the posterior distribution in the parameter space onto the direction of θ_1, thus the parameter about which we want to know the probability distribution. This process is frequently applied to the so-called *nuisance* parameters, which are parameters that we must take into account for the statistical analysis, but of which we are not directly interested. Indeed we can described the overall properties of our Universe with a very compact set of cosmological parameters, though we might need to include a number of nuisance parameters (such as the bias of galaxies inside cosmic voids, as well as internal parameters calibrated for a given physical model) which could be highly correlated with the cosmological ones.

4.2.1 The Gaussian likelihood

The parameter estimation that we described so far is presented in a general form, which can then be applied to a wide range of cosmological problems. However, it is common to consider a Gaussian form for the likelihood function $\mathcal{L}(\mathcal{D}|\vec{\theta})$. In this sense, the underlying process with which the data are generated it is considered as random. Having a data-set $\mathcal{D} = \{d_1, \ldots, d_N\}$ and a set of parameters of interest $\vec{\theta}$, the likelihood can be written as a multivariate Gaussian distribution, denoted as $\mathcal{G}(\mathcal{D}|\mu, \Sigma)$, which is the probability that the first element of the data-set has the value d_1 *and* the second has the value d_2 and so on:

$$\mathcal{L}(\mathcal{D}|\vec{\theta}) \equiv \mathcal{G}(\mathcal{D}|\mu, \Sigma) = \frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}\chi^2\right) , \qquad (4.4)$$

where

$$\chi^2 = (\mathcal{D} - \mu)^{\mathrm{T}}\Sigma^{-1}(\mathcal{D} - \mu) . \qquad (4.5)$$

In Eq. (4.4), n is the number of items in the data-set, $\mu \equiv \mathbb{E}[\mathcal{D}]$ is their expectation value and Σ is their covariance matrix. The latter reduces to the variance $\sigma^2 \equiv \mathbb{E}[(d_i - \mu)^2]$ for the one-dimensional case ($n = 1$). Indeed, a generic element of the matrix Σ is defined as

$$\Sigma_{ij} = \mathbb{E}\left[(d_i - \mu)(d_j - \mu)^{\mathrm{T}}\right] . \qquad (4.6)$$

The covariance matrix is a fundamental tool that has to take into account the correlation between the items in the data-set. In general, both μ and Σ may depend on the parameters of interest $\vec{\theta}$, and this dependence could be very complex to model.

4.2.2 Markov Chain Monte Carlo

We note that the likelihood function can be represented by a non-trivial mathematical expression, therefore the posterior distribution cannot always be computed analytically. The latter can be indeed very difficult to obtain in some cosmological applications, since the parameter space is typically multi-dimensional and the likelihood function has a complex form. For these reasons, in the last decades the parameter estimation has generally been computed by means of numerical algorithms. To construct the posterior it is necessary to map the likelihood distribution in the parameter space, sampling with great accuracy especially the most peaked regions of the distribution. An algorithm specifically designed to accomplish this task relies in the Markov Chain Monte Carlo method (MCMC, hereafter).

With the MCMC method we aim at building a map of the posterior distribution drawing a sequence (or *chain*) of points (or *samples*) in the parameter space. There are many algorithms that can be implemented to compute the chains, but the general idea is that the sequence has to follow a *Markovian* trajectory in parameter space, such that each point in the chain $X(t) = \vec{\theta}^{(t)}$ depends only on the previous one. In other words, the probability of the t–th element in the chain only depends on the value of $(t-1)$–th element. A fundamental property for these chains is the convergence to a stationary state (i.e. stable with the variation of t).

The probability of moving from $\vec{\theta}^{(t)}$ to $\vec{\theta}^{(t+1)}$ is described by a *transition probability* $T(\vec{\theta}^{(t)}, \vec{\theta}^{(t+1)})$, which has to satisfy the following *detailed balance condition*:

$$\frac{T(\vec{\theta}^{(t)}, \vec{\theta}^{(t+1)})}{T(\vec{\theta}^{(t+1)}, \vec{\theta}^{(t)})} = \frac{p(\vec{\theta}^{(t+1)}|\mathcal{D})}{p(\vec{\theta}^{(t)}|\mathcal{D})}, \tag{4.7}$$

i.e. the ratio of the transition probabilities has to be inversely proportional to the ratio of the posterior probabilities at the two points.

As already mentioned, the CBL provide methods to perform the MCMC posterior sampling. In particular two algorithms have been implemented:

- The *Metropolis-Hastings* algorithm (Hastings, 1970). This is the simplest and widely used MCMC algorithm. During the parameter-space sampling, a candidate point $\vec{\theta}^{(t)}$ is drawn from a *proposal distribution* $q(\vec{\theta}^{(t-1)}|\vec{\theta}^{(t)})$, that could be for example Gaussian, centred on the current point and having a fixed variance. In this algorithm, the proposal distribution is considered symmetric. Then the point is accepted with probability

$$\alpha = \min\left\{\frac{p_t q(\vec{\theta}^{(t)}|\vec{\theta}^{(t-1)})}{p_{t-1} q(\vec{\theta}^{(t-1)}|\vec{\theta}^{(t)})}, 1\right\} = \min\left\{\frac{p_t}{p_{t-1}}, 1\right\}, \tag{4.8}$$

where $p_t = p(\vec{\theta}^{(t)}|\mathcal{D})$ is the posterior distribution computed at the candidate point. The second equality derives from the symmetric property of q. If the candidate point is accepted, then the algorithm will move at that position adding it to the chain, else it will stay on the old point, double counting it. The process is iteratively repeated, by drawing another candidate point until a convergence is reached.

- The *Stretch-Move* algorithm (Goodman & Weare, 2010; Foreman-Mackey et al., 2013). This algorithm follows the simultaneous evolution of an ensemble of K walkers $S = \{X_k\}$. The position X_k of a given walker k is updated to Y by drawing randomly another walker $X_j \in S_{[k]} \equiv \{X_j, \ j \neq k\}$

$$X_k(t) \to Y = X_j + Z[X_k(t) - X_j], \tag{4.9}$$

where Z is a random variable drawn from a distribution $g(z)$:

$$g(z) = \begin{cases} \dfrac{1}{\sqrt{z}}, & \text{if } z \in \left[\dfrac{1}{a}, a\right], \\ 0, & \text{otherwise,} \end{cases} \tag{4.10}$$

where a is a constant which is generally set to 2. Then, the new position is accepted with probability:

$$q = \min\left\{Z^{N-1}\frac{p(Y|\mathcal{D})}{p(X_k|\mathcal{D})}, 1\right\}, \tag{4.11}$$

where N is the dimension of the parameter space. The process is then repeated for each walker in the ensemble S. In the CBL the whole procedure is parameterised by means of OpenMP Application Program Interface (OpenMP API), assigning the evolution of each walker to a different CPU and running the tasks in parallel.

4.3 Combination of independent probes

The advent of the wide field surveys marked in the last decade an unprecedented evolution in observational Cosmology and will led to a deeper understanding of our Universe in the next years. The huge amount of information coming from survey such as the SDSS[1] (Sloan Digital Sky Survey) (York et al., 2000; Eisenstein et al., 2011; Blanton et al., 2017a), and DESI (Dark Energy Spectroscopic Instrument) (Hang et al., 2021; Besuner et al., 2021),

---

[1]https://www.sdss.org/

together with data expected from the Vera C. Rubin Observatory LSST[2] (LSST Dark Energy Science Collaboration, 2012) and the ESA *Euclid* mission[3] (Laureijs et al., 2011; Amendola et al., 2018), will shed new light on the underlying cosmological scenario, testing the standard ΛCDM model. It will allow indeed to achieve tighter constraints on the expansion history of the Universe, together with a better understanding of the nature of the DE. Furthermore, we will have the opportunity to study in details the effective mass of neutrinos and the problem of the initial condition of our Universe. Cosmological probes like weak lensing, galaxy clustering, number counts of both galaxy clusters and cosmic voids will be exploited thanks to these new available data-sets to test the cosmological scenario. Their complementary study will lead to even tighter constraints on cosmological parameters, and their combination will possibly break the degeneracies between them.

As an example, Figure 4.1 reports the combination performed in Suzuki et al. (2012), in which the constraints have been derived from different probes: using the luminosity distance measurements of SNe Ia detected at $0.623 < z < 1.415$ with the Hubble Space Telescope, the CMB analysis with WMAP7 data and the BAO measurement from the combined analysis of the 7-th data release of SDSS and the 2dFGRS data. From this plot we can clearly see the great constraining power achieved with the combination of different probes, fundamental to break the strong degeneracies between cosmological parameters.

In the last decade, thanks to the remarkable achievements obtained by the CMB experiments (see e.g. Komatsu et al., 2011; Planck Collaboration et al., 2020), as well as by the Dark Energy Survey (DES) Dark Energy Survey Collaboration et al. (2016); Dark Energy Survey Collaboration (2018), by high-resolution spectroscopic and optical surveys (Blanton et al., 2017b; Ivezić et al., 2019), and deep galaxy surveys (Alam et al., 2017), Cosmology has entered in a new *precision* era, in which we are able to constrain cosmological parameters with sub-percent accuracy. This has led to the rise of the so-called *tensions* between different estimates of the same parameter, such as the Hubble constant tension, which we saw in Section 1.4. Combination of different experiments, when performed properly between *compatible* probes, can represent a very powerful tool in shedding light on these tensions. In particular, given the Bayesian framework described in the previous sections, one can perform a joint analysis on an ensemble of different experiments by setting a unique data vector $\hat{d} = \hat{d}_1 \cup \hat{d}_2 \cup \ldots \hat{d}_N$ as the concatenation of all the different data-sets. Writing the parameters of interest as $\vec{\theta}$ and the

---

[2]Legacy Survey of Space and Time; http://www.lsst.org
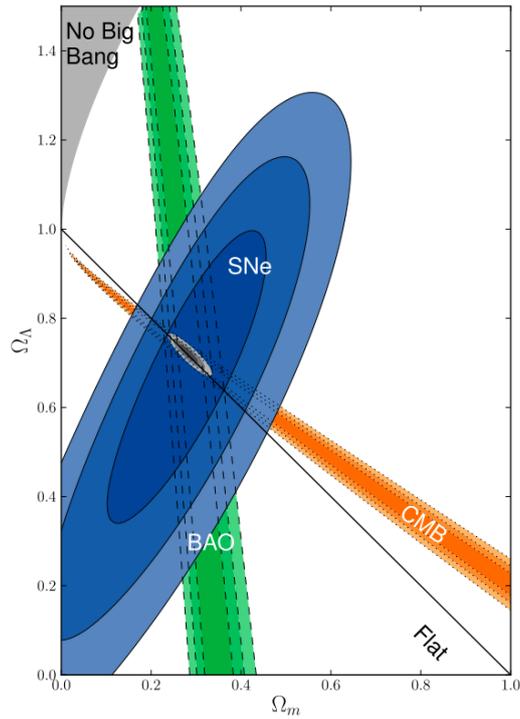[3]http://www.euclid-ec.org

Figure 4.1: Combination of different cosmological probes (black contour) in the $\Omega_m - \Omega_\Lambda$ parameter plane. The blue contours represent the constraints derived from the study of the distance modulus of a sample of SNe Ia in the redshift range $0.623 < z < 1.415$. The orange and green contours represent the constraints derived from the CMB data (WMAP7) and the study of the BAO position peak, respectively. The black solid line represents the equation $\Omega_\Lambda + \Omega_m = 1$, and delineates the transition between an open and a closed Universe. Credits to Suzuki et al. (2012).

nuisance parameters as $\vec{n}$, we can compute the joint posterior probability by comparing the data vector $\hat{d}$ with the theoretical model vector $\mathcal{M}(\vec{\theta}, \vec{n})$, which will include all the different models, and marginalising over nuisance parameters:

$$\mathcal{P}(\vec{\theta}|\hat{d}) \propto \mathcal{P}r(\vec{\theta}) \int \mathcal{L}(\hat{d}|\mathcal{M}(\vec{\theta}, \vec{n}), \mathcal{C}) \mathcal{P}r(\vec{n}) \mathrm{d}\vec{n} . \qquad (4.12)$$

Here $\mathcal{C}$ is the *cross-covariance matrix*, which has to be taken into account in order to model properly the correlation between the different data-sets, $\mathcal{L}$ is the joint likelihood, and $\mathcal{P}r(\vec{\theta})$ and $\mathcal{P}r(\vec{n})$ are the prior distributions on the parameters of interest and the nuisance parameters, respectively. Obtaining the cross-covariance matrix is not always an easy task: an option is to compute it from theoretical models, but it is usually not trivial primarily because of the poorly understanding of the statistical properties of the data, such as the nonlinear regime for galaxy redshift surveys and weak lensing or the galaxy bias in redshift surveys. Another option is to estimate it from simulations or from observations directly, using resampling methods (see Appendix A). We note that valuating the cross-covariance matrix from the data leads generally to high levels of noise in the estimate. To improve the signal to noise ratio it is possible to use different realisations of the same simulation, lowering statistical uncertainty related to this measure. However, this method is generally computationally expensive even for single probe analyses (Taylor et al., 2013; Taylor & Joachimi, 2014; Krause et al., 2017). The main advantage of this method is related to the possibility of including in the estimate of cross-covariance matrix all the expected observational effects, by adding for instance flux limits and survey masks to the analysed simulations.

In this Thesis work, we focus on how to properly perform the combination of independent cosmological probes, where "independent" means that we make the assumption that the different probes are not statistically correlated. More precisely, we consider null off-diagonal terms for both the covariance matrix associated with the data-sets at different redshifts and for the cross-covariance matrix relative to the different cosmological probes. In Appendix A we discuss these assumptions by estimating the covariance matrix between galaxy cluster and void number counts from the analysed mock catalogues, performing a Jackknife resampling.

The combination of different probes will be effective whenever different cosmological constraints overlap in parameter space (like the contours represented in Figure 4.1). A contour is a set of points $\{\vec{\theta}_i\}$ in the parameter space, which gather in the regions where the posterior probability distribution $p(\vec{\theta}|\mathcal{D})$ is peaked. One of the most common methods to combine

different independent probes is through the product of their likelihood functions, or more generally their posterior product (PP). In the following sections we will present other two combination techniques: the *Importance Sampling* (IS) method and the *Posterior as Prior* (PaP) method. Both the PP and the IS methods have been implemented in the new CBL's class called `CombinedPosterior`, belonging to the `statistics` namespace, with which it is possible to manage different objects of the parent class `Posterior`. In particular, an object of the class `Posterior` is built passing to its constructor[4] a `Likelihood` object and a `Prior` object, in agreement with the definition of the Bayes theorem in Eq. (4.1). All these classes follow the hierarchy represented in Figure 4.2. The PaP method can be performed by the user by properly specifying the prior distribution of the free parameters by means of a new implemented constructor of the class `PriorDistribution`.
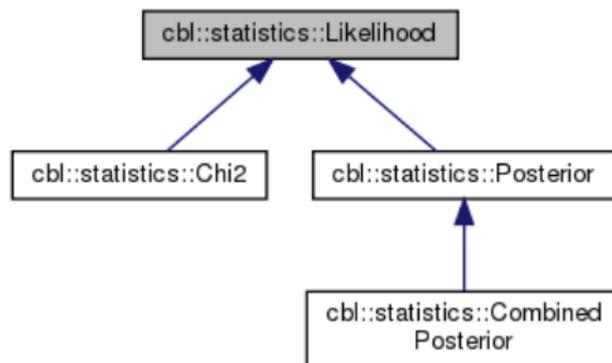
Figure 4.2: Hierarchy diagram of the `Likelihood` class. In this scheme `Chi2` and `Posterior` are directly connected to the parent class `Likelihood`, from which they inherit the public attributes and functions. An analogous inheritance scheme is followed for the `Posterior` and its derived class `CombinedPosterior`.

4.3.1 Posterior product combination

The first method we present is the product of the different posterior probability distributions (reported also with the acronym PP). Let us assume to

---

[4]A constructor is a special member function of a C++ class. Constructors build the objects of the class they belong to, generally setting the values of the fundamental variables of the class.

have a set $\{\mathcal{D}_1, \ldots, \mathcal{D}_N\}$ of N data-sets, with which we want to test N different models. We can compute the single posterior distributions via the Bayes theorem expressed in Eq. (4.1), by specifying a likelihood function $\mathcal{L}(\mathcal{D}_j|\vec{\theta})$ for each of them and a unique prior distribution $\mathcal{P}r(\vec{\theta})$ for the parameters of the models. For each point $\vec{\theta}_i$ obtained during the sampling of the parameter space, the associated value of the un-normalised combined posterior distribution in that point will be the product of the different likelihood functions, times the value of the prior, both computed at $\vec{\theta} = \vec{\theta}_i$:

$$P(\vec{\theta}_i|(\mathcal{D}_1, \ldots, \mathcal{D}_N)) = \mathcal{P}r(\vec{\theta}_i) \prod_{j=1}^{N} \mathcal{L}(\mathcal{D}_j|\vec{\theta}_i). \tag{4.13}$$

Before the implementation of the new codes for probe combination, it was possible to perform the PP method with the CBL by joining manually the data-sets of interests and by building a vector of models $\mathcal{M}$, specifying then the association of each model to each data bin. The main advantage of the new `CombinedPosterior` class for this type of combinatio is that the code automatically manages all the data-sets and the models in input. Indeed, until the hypothesis of independence between the different data-sets holds, the `CombinedPosterior` class allows to create N different `Posterior` objects, given a unique prior distribution for the parameters of interest, and to combine them with a more user-friendly structure.

In the following Listing we show an example of such combination considering $N = 2$ independent cosmological probes:

```python
# import the CosmoBolognaLib and the Python modules
import CosmoBolognaLib as cbl
import numpy as np

# construct the data-sets by reading input files
dataA = cbl.Data1D(path_datasetA)
dataB = cbl.Data1D(path_datasetB)

# set the Likelihood and the Prior type
LikeType = cbl.LikelihoodType__Gaussian_Error_
PriorType = cbl.DistributionType__Uniform_

# construct the likelihood objects
likelihoodA = cbl.Likelihood(dataA, modelA, LikeType)
likelihoodB = cbl.Likelihood(dataB, modelB, LikeType)

# define the priors for the parameters
#(here uniform prior for two parameters)
prior_theta1 = cbl.PriorDistribution(PriorType, min_theta1,
    max_theta1)
```

```
20  prior_theta2 = cbl.PriorDistribution(PriorType, min_theta2,
        max_theta2)
21
22  # put the priors in a single pointer vector
23  prior_distribution = cbl.PriorDistributionPtrVector([
        prior_theta1, prior_theta2])
24
25  # construct the posterior objects
26  PostA = cbl.Posterior(prior_distribution, likelihoodA, seed1)
27  PostB = cbl.Posterior(prior_distribution, likelihoodB, seed2)
28
29  # put them into a pointer vector
30  ptr_posteriors = cbl.PosteriorPtrVector([PostA, PostB])
31
32  # construct the combination object by passing the pointer
        vector
33  CombPost = cbl.CombinedPosterior(ptr_posteriors)
34
35  # initialize the chains
36  CombPost.initialize_chains(chain_size, nwalkers, 1.e-5, start
        )
37
38  # combine by sampling the posterior distribution (Stretch
        Move algorithm)
39  CombPost.sample_stretch_move(2)
```

Listing 4.1: Python example of how to perform the *Posterior Product* combination between two different independent probes.

Here, we start preparing two generic data-sets to be used as inputs for the two distinct probes. Then we build two objects of the `Likelihood` class by using the CBL constructor. It requires to define also a theoretical model for each probe (`modelA` and `modelB`), that can be set by the user also as an external function, and a likelihood type (`Gaussian_Error` in this case). Then we define the priors for the parameters of interest, that here are generically called `theta1` and `theta2`. In this example we set a flat prior distribution for both the parameters, for which we set only the minimum and the maximum values acceptable for our analysis. Finally, we build the `Posterior` objects required for the combined analysis. To do this, we provide to the constructor a vector of pointers to the `Prior` objects, then the `Likelihood` objects prepared to each probe and a seed used to initialise the MCMC. In the end we construct an object of the new class `CombinedPosterior` by giving a pointer vector with the `Posterior` objects. Thanks to this, we can call the functions responsible for the running of the MCMC. `chain_size` and `nwalkers` are the parameters used to set the posterior sampling, while `start` is a vector of values at which the chain will start. The resulting combination,

i.e. the product of the likelihoods times the prior, is performed inside the `sample_stretch_move` function, that has been re-implemented within the `CombinedPosterior` class. The final result of this code will be an output file containing the steps performed by the MCMC for the two parameters of interest, together with other files specifying some characteristics of the sampled combined posterior distribution (e.g. mean, median, standard deviation, quartiles and covariance matrix).

In order to take into account the possible correlations between the parameters of a given model, the user can also exploit this method in combination with a correlated prior distribution. For instance, supposing to know the correlation between the analysed parameters, we can set a multivariate Gaussian distribution as unique prior for these parameter, by specifying their covariance matrix and the vectors of mean values inside the `PriorDistribution` constructor. This method will be applied to perform a different type of probe combination in Section 4.3.3.

4.3.2 Importance sampling combination

The second method we present for the combination of independent probes is the *Importance Sampling* (reported also with the acronym IS). Importance sampling refers to a general technique to extract properties of an underlying distribution $p(x)$, often called *nominal* or *target*, for a random variable X by sampling from another distribution $q(x)$, i.e. $X \sim q$,[5] called *proposal* or *importance* distribution. The expected value of a given function $f(X)$, under the distribution p, indicated as $\mathbb{E}_p[f(X)]$, is:

$$\mu \equiv \int f(x)p(x)\mathrm{d}x = \mathbb{E}_p[f(X)]. \tag{4.14}$$

The *Importance Sampling* is based on the following fundamental identity:

$$\begin{aligned} \mathbb{E}_p[f(X)] = \int f(x)p(x)\mathrm{d}x &= \int \frac{f(x)p(x)}{q(x)}q(x)\mathrm{d}x \\ &= \int w(x)f(x)q(x)\mathrm{d}x = \mathbb{E}_q[w(X)f(X)], \end{aligned} \tag{4.15}$$

where $w(x) \equiv p(x)/q(x)$ is called *importance weight* and $\mathbb{E}_q[\cdot]$ denotes the expectation value with respect to the importance distribution. Eq. (4.15) holds for any probability function q that has its support including the support

---

[5]with this notation we indicate that the random variable X is distributed following the probability distribution function q.

of p. Therefore, one can estimate the expectation value $\hat{\mu}_q$ for the function f by drawing $x_1, \ldots, x_N$ independent samples from q:

$$\hat{\mu}_q = \frac{1}{n} \sum_{j=1}^{N} w(x^j) f(x^j), \quad x^j \sim q. \tag{4.16}$$

One can show that $\mathbb{E}_q[\hat{\mu}_q] = \mu$ (directly from Eq. 4.15), and that the estimate of the variance is $\mathbb{V}\mathrm{ar}_q[\hat{\mu}_q] = \sigma_q^2 / N$. Here σ_q^2 is the variance associated to the importance distribution, which has the following form:

$$\sigma_q^2 = \int \frac{(f(x)p(x))^2}{q(x)} \mathrm{d}x - \mu^2 = \int \frac{\left(f(x)p(x) - \mu q(x)\right)^2}{q(x)} \mathrm{d}x. \tag{4.17}$$

In the case of cosmological calculations, we deal with a set of parameters $\vec{\theta}$ which follow a posterior distribution, constructed from data-sets, models and priors. The IS combination is performed by considering the posterior of one probe as the importance distribution for the other, and viceversa. During the MCMC, one has to compute the importance weight w for each sampled point in the parameter space. For example, having two data-sets, A and B, the normalised importance weights $w_i = w(\vec{\theta}_i)$ computed for each point $\vec{\theta}_i$ during the sampling of the posterior of A (hence by considering p_A as the importance distribution) results:

$$w_i = \frac{\max\{P_A(\vec{\theta}_A)\}}{\max\{P_B(\vec{\theta}_A)\}} \frac{P_B(\vec{\theta}_i)}{P_A(\vec{\theta}_i)}, \tag{4.18}$$

where $\max\{P_\mathrm{x}(\vec{\theta}_A)\}$ indicates the maximum value of the posterior p_x in the points sampled by the probe A. Doing the same procedure considering p_B as the importance distribution, we will finally obtain a value w for each point in the parameter space, thus for each step of the MCMC computed for both A and B. By concatenating the output chains obtained in the two cases, one can compute the final (combined) expectation value of the posterior following the Eq. (4.16).

This procedure is totally performed inside the `importance_sampling()` function, implemented within the `CombinedPosterior` class. This function takes as input the paths at which the user wants to store the final output chains and the parameters required for the MCMC sampling. It provides in output the MCMC produced from the sampling of the posterior of each probe, together with the concatenated chains with the importance weight computed to each step.

A-posteriori importance sampling

To combine two posterior distributions with the IS method described above, one has to provide the data-sets for each probe, as well as their models. However, the user may want to combine the posterior distribution obtained from other experiments (e.g. combining with Planck18 or WMAP constraints, as well as with other probes for which public results have been provided), without the need to re-run the chains. For this reason we implemented another `importance_sampling()` function, which can compute the importance weights for two external, pre-ran chains. In order to calculate the importance weights w_i, we need to evaluate the posterior of the first probe in the points of the parameter space sampled for the second probe.

To accomplish this task, we implemented a function that performs a N-dimensional interpolation, by means of a chain mesh. This function is indeed implemented inside the `cbl::chainmesh::ChainMesh` class, which provides optimised tools to construct a mesh grid starting from an N-dimensional distribution of points. The implemented `interpolate()` function is able to adapt the target distribution to a regular grid (i.e. having cells of the same size for the different sides), assigning each point of the distribution to a given cell. The interpolation at each point of the importance distribution is performed by the `interpolate()` function by computing the average of the closest points belonging to the target distribution. The search for closest points is performed by the `close_objects()` function, which allows to speed up the computation by considering only the objects belonging to the cells that fall inside a given radius, centred on the considered point of the parameter space. The value of this radius is specified by the parameter `rMAX`, which can be set by the user. Other parameters required for the `interpolate()` function are `cell_size` and `distNum`, representing the grid cell side and the number of the closest points used to perform the average operation, respectively.

Figure 4.3 shows an example of this procedure. The two distributions of points are obtained from external chain files, which sample the parameter space given by the generic parameters $\{\theta_1, \theta_2\}$. In this example the points reported in orange represent the importance distribution and the points in blue the target distribution. We choose the point of the importance distribution marked with a "X" and we search for the closest points belonging to the target distribution, by means of a chain mesh. We consider in particular `rMAX` = 3 and `cell_size` = 2, building a 100×100 grid (100 cells for each side). The insert in the corner shows a zoom of the points involved in the computation of the close objects. The searching region is reported with a red circle and the big points are those selected by the algorithm. In particular,
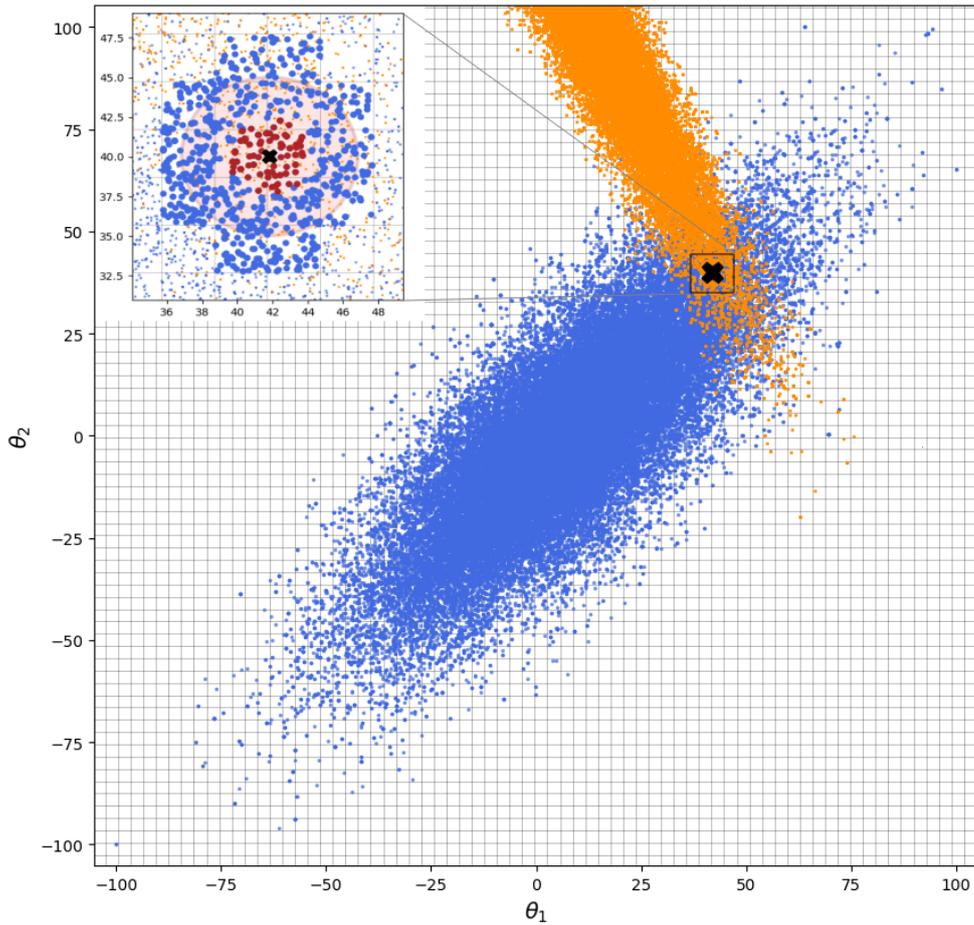
Figure 4.3: Example of analysis of two contours in the parameter space. During the sampling of the orange contour, the interpolated posterior value of the blue contour at a given point (black cross) is the average between the posterior values computed at the closest points (red dots) inside a circle of radius `rMAX`. The points found by the `close_objects()` function are represented as blue dots in the insert in the top-left corner. Among these points, the red ones are those selected to perform the average operation.

these are the points belonging to the grid cells included or intersected by the searching radius. Finally, the big red points are the ones selected, among all the close objects, to perform the average operation required for the interpolation procedure. The number of points selected among the close objects is given by the parameter `distNum`.

Once we have the interpolated values for one posterior distribution at each point of the other, the importance weights are computed by the `importance_sampling()` function as in Eq. (4.18). However, since the importance weight is defined as a ratio, and the posteriors distribution can show some noise in the sampled values, it is possible to encounter numerical fluctuations in the weight distribution. The user can therefore choose to remove the outlayers in the obtained distribution of weights thanks to the parameter `cut_sigma`. With this parameter we set the number of standard deviations that have to be kept in the analysis, starting from the mean value of the distribution (e.g. `cut_sigma`= 5 implies to discard the values over the 5σ of the weight distribution).

The *Importance Sampling* method is the fastest among the methods implemented in this Thesis work, but it is restricted only to $N = 2$ different cosmological probes. Moreover, it has the downside of being dependent on the parameters set to perform the interpolation (`rMAX`, `distNum`, `cut_sigma`), even if in many cases it does not show relevant modifications with the variation of these parameters. Nevertheless, it is also the only method that does not require in input the analysed data-sets and neither theoretical models, but only the resulting MCMC. This allows, in principle, to combine any publicly available chain relative to independent probes.

4.3.3   Posterior as Prior combination

The last method we implemented is the *Posterior as Prior* combination (PaP, hereafter). In this method we consider the posterior distribution of a given cosmological probe as the prior distribution for the other. One can easily show that this procedure, also known as *Bayes update*, is equivalent to the *Posterior Product* combination described in Section 4.3.1. Indeed, let us consider the parameter set $\vec{\theta}$ we want to constrain by testing two models on two different data-sets $\mathcal{D}_1$ and $\mathcal{D}_2$. Considering the result of the sampling of second posterior as the prior for the sampling of the first posterior $p(\vec{\theta}) = p(\vec{\theta}|\mathcal{D}_2)$, it follows from the Bayes theorem Eq. (4.1) that:

$$p(\vec{\theta}|\mathcal{D}_1) = \frac{p(\mathcal{D}_1|\vec{\theta})p(\mathcal{D}_2|\vec{\theta})p(\vec{\theta})}{p(\mathcal{D}_1)p(\mathcal{D}_2)} \equiv p(\vec{\theta}|(\mathcal{D}_1, \mathcal{D}_2)). \qquad (4.19)$$

As a consequence of this, the *Posterior as Prior* combination method is symmetric, meaning that it is equivalent to consider the posterior 1 as a prior for the model 2 or the posterior 2 as the prior for the model 1 (with the proper likelihood function specified for the posterior sampling).

With this method, we want to encapsulate the information achieved by a different experiment in the prior set for the analysed parameters. Let us assume that we want to constrain a set of N different cosmological parameters $\vec{\theta}$, and that we previously sampled a posterior distribution for that parameters using a given cosmological probe. To use the constraints obtained from a different cosmological probe as priors for another Bayesian analysis we implemented two techniques, by means of two new `PriorDistribution` constructors. With the first method, we read the MCMC output given by another probe, exploiting the `interpolate()` function presented in the previous section to have access to the points of the parameter space that are not covered by the sampling procedure. With the second, we approximate the posterior distribution that we want to use as a prior with a multivariate Gaussian.

The former method is in general very convenient since one can consider as prior any type of posterior distribution, not necessarily Gaussian or described by analytical functions. However, this method is affected by the same issues related to the `interpolate()` function described before. In particular, it can be subject to numerical fluctuations given by the scatter of the points which constitute the read posterior distribution. Moreover, it slightly depends on the parameters set for the interpolation procedure, even if the results obtained with this algorithms are fully consistent even for large variations of these parameters.

The second method is based on the assumption that the distribution used as prior can be represented by a multivariate Gaussian. This hypothesis holds for the majority of the cosmological cases, but it is not always appropriate or precise. This method is faster with respect to the former, since it does not require to perform the interpolation procedure for each step of the MCMC, and it is more accurate than considering N independent Gaussian prior distributions for the parameters, since it takes into account also the correlation between the latter. Let us assume of having the data coming from the sampling of the posterior distribution of the first probe. These may be provided by a different experiment or may have been produced by the user. In the latter case, running the MCMC with the CBL, we will also have as output the covariance matrix, Σ, for the different parameters $\vec{\theta}$, which will be used to build the prior we are interested in. With the new implemented version of `PriorDistribution` constructor we can build a prior represented by a multivariate distribution, thus $p(\vec{\theta}) = \mathcal{G}(\vec{\theta}|\vec{\mu}, \Sigma)$. This form is the same

reported in Eq. (4.4): in this case the parameters $\vec{\theta}$ are specified at the first entry, and $\vec{\mu}$ and Σ are the expectation values vector of the parameters $\vec{\mu} = \{\mathbb{E}[\theta_1], \ldots, \mathbb{E}[\theta_N]\}$ and their $N \times N$ covariance matrix, respectively.

In the Listing 4.2 we report an example of the usage of this typology of combination.

```python
from CosmoBolognaLib import DoubleVector as dv
from CosmoBolognaLib import DoubleVectorVector as dvv

# read the distribution of parameters from the external chain
theta1, theta2 = np.genfromtxt(file_chain, usecols=(0,1),
    unpack=True)

# compute the average vector
Mean = np.median(list(zip(theta1, theta2)), axis=0)

# compute the covariance matrix
cov = np.cov(np.stack((theta1, theta2), axis=0))

# store the cov. matrix in a vector of vectors
covM = dvv([list(cov[0]),list(cov[1])])

# set the prior as multivariate Gaussian distributed
prior = cbl.PriorDistribution(cbl.DistributionType__Gaussian_
    , dv(Mean), covM, dv([min_theta1,min_theta2]), dv([
    max_theta1,max_theta2]))

# put the prior into a single pointer vector
prior_distribution = cbl.PriorDistributionPtrVector([prior])

# construct the Posterior object by passing the prior pointer
    vector and the likelihood
Post = cbl.Posterior(prior_distribution, likelihood, 696)

# initialize chains by choosing a starting point
Post.initialize_chains(chain_size, nwalkers, 1.e-5, start)

# sample the posterior with the Stretch Move algorithm
Post.sample_stretch_move(2)
```

Listing 4.2: Python example on how to consider an external posterior distribution as the prior for a given cosmological probe.

We first read the MCMC relative to the sampling of the posterior distribution of a probe, in which `theta1` and `theta2` are the free parameters of the model. We then compute the mean and the covariance matrix of these parameters from the chains. With these values we can now build a `PriorDistribution` object, which will constitute the multivariate Gaussian distribution used as

prior. Consequently, the statistical Bayesian can proceed as described in the Listing 4.3.1, by defining the `Likelihood` object for the second probe by means of a data-set, a model and a likelihood type, and finally running the posterior sampling method.

All the algorithms implemented in this Thesis work are described in details and applied to a general example in the following jupyter notebook, which can also be found inside the CosmoBolognaLib C++/Python libraries. Moreover we present a further application of the a-posteriori importance sampling technique in this jupyter notebook, by combining publicly available data obtained from the WMAP experiment (Komatsu et al., 2011) and from the analysis of the BAO peak.

Chapter 5

Cosmological constraints from the combination of cluster and void number counts

In this Chapter we shall apply the algorithms implemented in this Thesis work and presented in Chapter 4, to combine galaxy cluster and void number counts. The data-sets are extracted from simulated catalogues, which are obtained from the Magneticum Pathfinder simulations, described in Section 5.1. In order to extract constraints from these cosmological probes, we compare the simulated galaxy cluster number density with the Despali model described in Section 3.5.2, and the void number density with the Vdn model described in Section 3.6.1. We perform a Bayesian statistical MCMC analysis by sampling the posterior distribution of the parameters Ω_m and σ_8. The considered models will have to be calibrated to match the halo finder assumptions. Moreover, it will be necessary to take into account the bias of galaxies inside and around cosmic voids (see Section 3.6.4). The calibrations are described in Section 5.2 for the HMF model and in Section (**?**) for the VSF, while in Section 5.4 we present the constraints obtained from each probe alone. Finally, in Section 5.5 we apply and compare the combination methods described in the previous Chapter on the constraints obtained independently from the two considered cosmological probes.

5.1 The Magneticum simulations

In this Thesis work we make use of simulated galaxy and galaxy cluster catalogues extracted from the Magneticum Pathfinder Simulations[1] (Dolag

---

[1] http://www.magneticum.org/
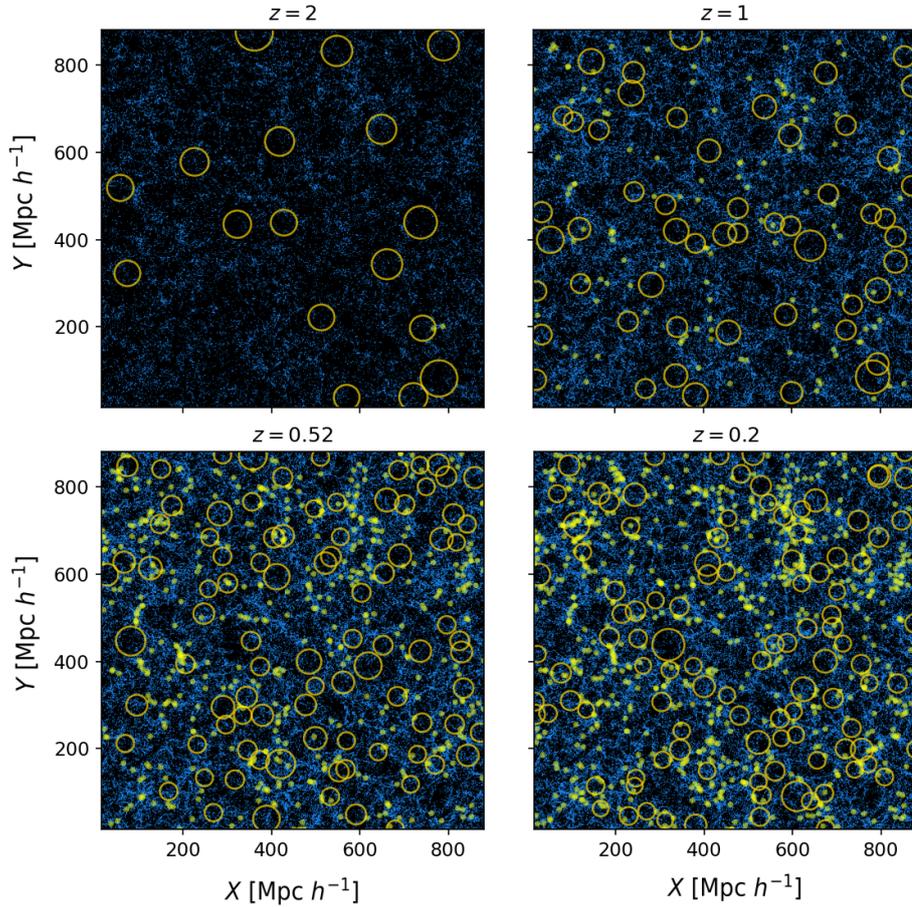
Figure 5.1: Evolution of large-scale structures in the Magneticum Pathfinder simulations, from redshift $z = 2$ to $z = 0.2$. Blue dots represent the distribution of (simulated) galaxies having stellar mass $M_* \geq 10^{10} \ M_\odot \ h^{-1}$, inside a comoving box of volume $(896 \ \text{Mpc} \ h^{-1})^3$ (Box1). The yellow dots represent galaxy clusters with mass $M_{500c} \geq 5 \times 10^{13} \ M_\odot \ h^{-1}$, while yellow empty circles represent the cosmic voids identified in the distribution of galaxies. The axes are expressed in comoving coordinates, and the Z-coordinate is represented as a slice that ranges from $520 \ \text{Mpc} \ h^{-1}$ to $580 \ \text{Mpc} \ h^{-1}$.

et al., in preparation), a large set of cosmological, hydrodynamical simulations, having box volumes going from $(12 \text{ Mpc}/h)^3$ (Box6) to $(2688 \text{ Mpc}/h)^3$ (Box0). These boxes have different resolutions, and thus allow us to study the formation and evolution of both large-scale structures as well as phenomena that occur on smaller scales, from the motion of galaxies to the physics of gas inside them. The Magneticum simulations have been run with the parallel code P-GADGET3, which is an updated version of the TreeSPH GADGET-2 code presented in Springel (2005). In the latter, gravitational forces are computed through a TreePM algorithm, while the hydrodynamics is modelled through a an SPH algorithm (see Section 2.6).

For our analysis we consider Box1, which is a large simulation that follows the evolution of 2×1526^3 particles in a comoving volume of $(896 \text{ Mpc}/h)^3$, with which it is possible to make a detailed statistical analysis of galaxy clusters and cosmic voids. Furthermore the main baryonic physics phenomena are implemented in these hydrodynamic simulations, following the methods presented in Springel & Hernquist (2003), such as the cooling of gas, the star formation and supernovae feedback. In addition, black holes and AGN feedback are included (Di Matteo et al., 2008), as well as thermal conduction, stellar population and chemical enrichment models (Tornatore et al., 2003; Tornatore & Borgani, 2007).

The simulation outputs, from which cluster and void catalogues are constructed, are selected at four redshift bins, $z = (0.2, 0.52, 1, 2)$, based on the spatially-flat ΛCDM model, with parameters fixed to the seven-year Wilkinson Microwave Anisotropy Probe (WMAP7) data (Komatsu et al., 2011), with matter density $\Omega_m = 0.272$, power spectrum normalisation $\sigma_8 = 0.809$, Hubble constant $H_0 = 70.4$ km s^{-1}Mpc^{-1}, DE density $\Omega_\Lambda = 0.728$ and primordial spectral index $n_s = 0.963$. Figure 5.1 shows the evolution of the large-scale distribution of galaxies inside a slice of thickness 60 Mpc h^{-1}, extracted from the comoving cosmological boxes we considered for the analysis. We also report the galaxy clusters identified in these simulation boxes by means of the SUBFIND algorithm, i.e. the SO halo finder (see Section 3.5) employed inside the Magneticum and the cosmic voids identified with the void finder VIDE (see Section 3.6.2) in combination with the cleaning procedure (see Section 3.7).

5.2 Calibration of the halo mass function

Before comparing the number density of clusters and the HMF model described in Section 3.5.2, a calibration of the latter is necessary, because of the differences in the halo definitions between the Magneticum simulations

and that of the theoretical Despali model. Indeed, in the former, DM haloes are detected with the SUBFIND algorithm (Springel et al., 2001; Dolag et al., 2009) which finds haloes through a SO halo finder, considering an overdensity threshold of $\Delta = 500\rho_c$, after having assembled DM particles together with a FoF algorithm with a linking length $b = 0.16$. As described in Section 3.5.2, the Despali model has been calibrated, instead, on haloes detected with a SO method considering overdensities equal to Δ_{vir}. Despali et al. (2016) provided fitting formulae, Eq. (3.45), for the parameters of their model in order to predicts the HMF at different thresholds with respect to Δ_{vir}. Nevertheless, the calibration performed by the authors considered DM-only cosmological simulations, while the Magneticum simulations also follow the baryonic component. For these reasons, we re-calibrate the parameters of the HMF model, by assuming the validity of the functional relation given by Eq. (3.45). This translates into finding other coefficients for the relations $a = a(x), p = p(x)$ and $A = A(x)$, where $x \equiv \log(\Delta(z)/\Delta_{vir}(z))$.

As shown in Despali et al. (2016), while the parameters A and a exhibit a very regular trend, described by Eq. (3.45), the behaviour of p is very uncertain. Moreover, the HMF results more sensitive to the variation of a and A, with respect to the variation of p. For these reasons, we choose to consider p as fixed to the value obtained by the authors in the most general fitting case, that is $p = 0.2536$. Lastly, we test the calibration with two different minimum masses for the galaxy cluster catalogues: $M_{500c} \geq 5 \times 10^{13} \; M_\odot h^{-1}$ and $M_{500c} \geq 10^{14} \; M_\odot h^{-1}$. This acts as a simplified selection function, which we adopted in order to mimic the range of masses for galaxy clusters in real surveys, and because for smaller masses we verified the presence of incompleteness in the cluster counts. We do not focus on any observable-mass relation in this work, and we consider cluster masses as a direct observable. The first minimum mass cut is applied to all the 4 redshift snapshots, while for the other selection no clusters were found at $z = 2$. Hence we decided to remove the latter from our analysis.

Writing the general form of the fitting formula (3.45) as:

$$\begin{aligned}
a &= a_1 x^2 + a_2 x + a_3 \\
A &= -A_1 x + A_2 \\
p &= 0.2536 \; ,
\end{aligned} \qquad (5.1)$$

we calibrate the new relations for the model parameters by performing a Bayesian statistical MCMC analysis on the simulated HMF, by sampling the posterior distribution of the a_i and A_i coefficients, fixing the cosmological parameters to the WMAP7 results. In particular, we consider flat priors for these coefficients and a Gaussian likelihood function (Section 4.2.1) for

the fitting. The result of the calibration for the two different mass cuts is presented in Figure 5.2, where it is shown the posterior distribution for the coefficients of Eq. (5.1). We recover the Despali a_i parameters values within the 95% confidence regions, while the A_i parameters are consistent with those found by the authors within the 68% confindence regions. However, for all the parameters, the errors, as well as the degeneracies between them, are very high. Moreover, the posterior distribution for the parameters are Gaussian for the lowest mass cut, while this no longer holds for the highest mass cut. Indeed the a_i parameters show a slight level of asymmetry.

Considering clusters having mass $M_{500c} \geq 5 \times 10^{13}\ M_\odot h^{-1}$, we found the following fitting relations for the Despali parameters:

$$a = (5 \pm 2)\ x^2 - (5 \pm 2)\ x + (2.1 \pm 0.6)$$
$$A = (-0.1 \pm 0.1)\ x + (0.27 \pm 0.09) \tag{5.2}$$
$$p = 0.2536\ ,$$

while for $M_{500c} \geq 10^{14}\ M_\odot h^{-1}$ we found:

$$a = (7 \pm 4)\ x^2 - (7 \pm 4)\ x + (3 \pm 1)$$
$$A = (-0.2 \pm 0.4)\ x + (0.3 \pm 0.2) \tag{5.3}$$
$$p = 0.2536\ .$$

Considering the latter as our new fitting formulae for the Despali parameters, we can compare the simulated numerical densities, extracted from the Magneticum simulation, with the theoretical Despali HMF, for the 4 different redshifts considered. The number counts are divided into equidistant logarithmic mass bins, with a Poissonian error associated for each of them. The comparison between model and simulated data, before and after the re-calibration, can be seen for the two mass cuts in Figures 5.3 and 5.4. From the residuals, which are computed as the ratio between the difference *data-model* and the *data errors*, one can clearly see how the re-parameterised model fits better the simulated data, particularly at low redshift.

5.3 Calibration of the void size function

The other observable we considered is the abundance of cosmic voids as a function of their radius. We considered voids identified in the simulated galaxy catalogues. We consider only galaxies having a stellar mass $M_* \geq 10^{10}\ M_\odot\ h^{-1}$, following the choice made in Marulli et al. (2017) who analysed the same cosmological simulations. We apply the VIDE void finder described
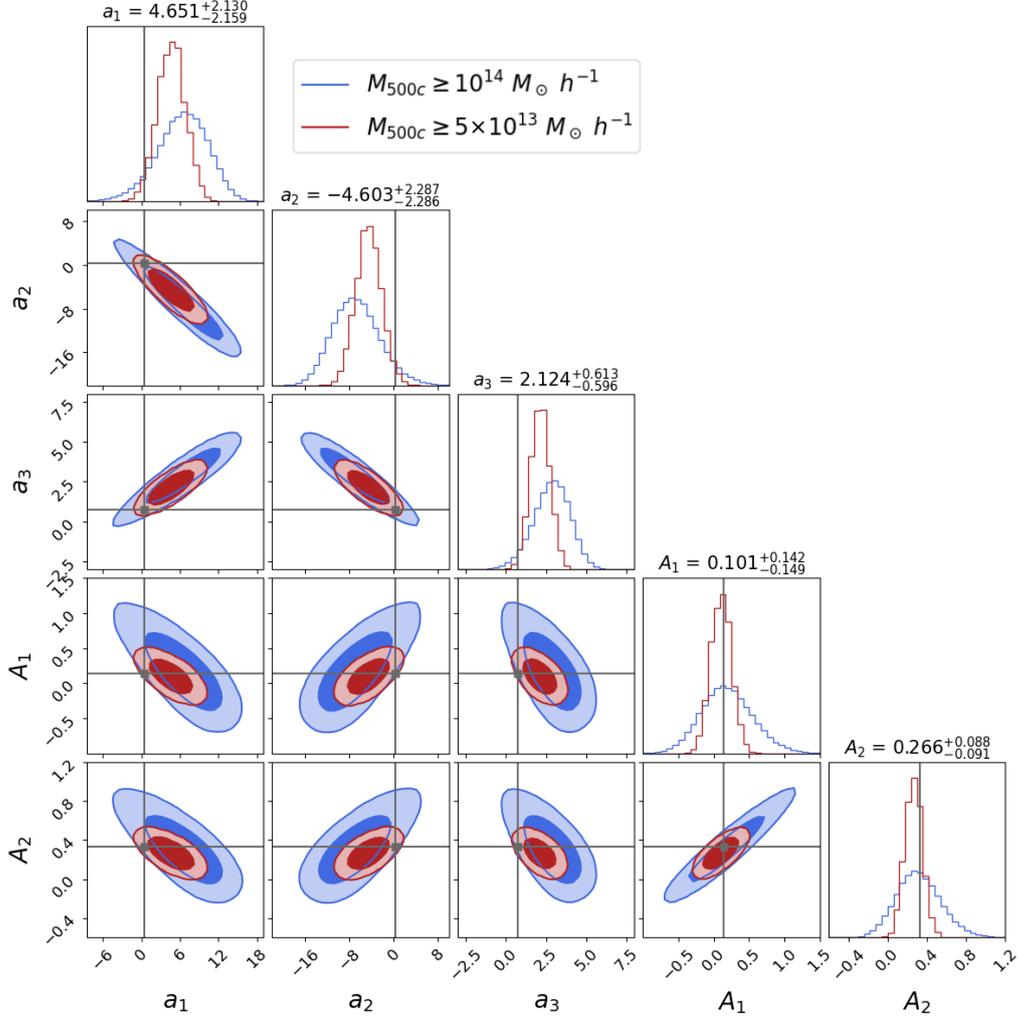
Figure 5.2: Posterior probability distribution of the coefficients of the fitting relation between the Despali parameters (a, A) and $x = \log_{10}(\Delta(z)/\Delta_{vir}(z))$, for two different mass cuts (red contours for $M_{500c} \geq 5 \times 10^{13}\ M_\odot h^{-1}$, blue contours for $M_{500c} \geq 10^{14}\ M_\odot h^{-1}$). Dark and light areas show the 68% and 95% confidence regions, while the solid gray lines represent the values for the coefficients found in Despali et al. (2016). The top of each column states the mean and standard deviation of the 1D marginal distributions of the parameters for the lowest mass cut.
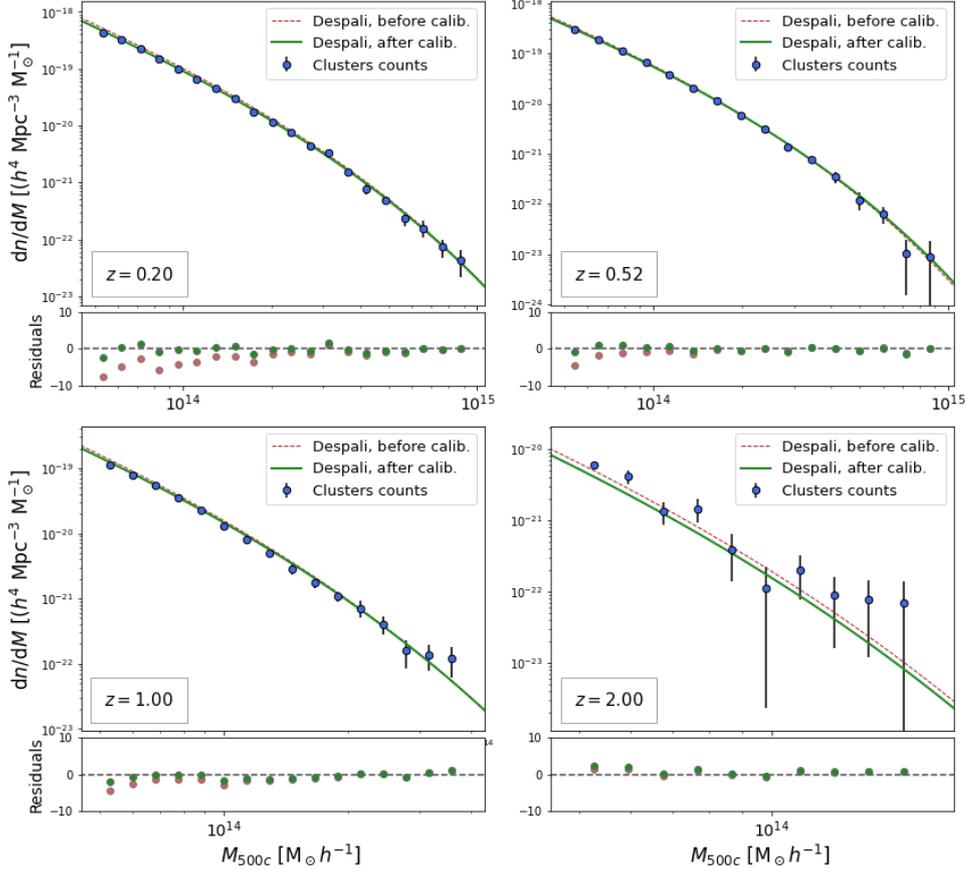
Figure 5.3: The measured mass function of galaxy clusters (blue dots) identified in the Magneticum Simulation (Box1), having $M_{500c} \geq 5 \times 10^{13} \ M_\odot h^{-1}$, at redshifts $z = 0.2, 0.52, 1, 2$. *Upper sub-panels*: the red dashed line represents the theoretical HMF with model parameters expressed with the fitting formula provided by Despali et al. (2016), while the green solid line represents the HMF after the model parameters re-calibration. *Lower sub-panels*: the residuals of the cluster counts, computed as the ratio between the difference *data - model* and the *data errors*. The different colours represent the residuals with respect to the two models considered in the upper sub-panels.
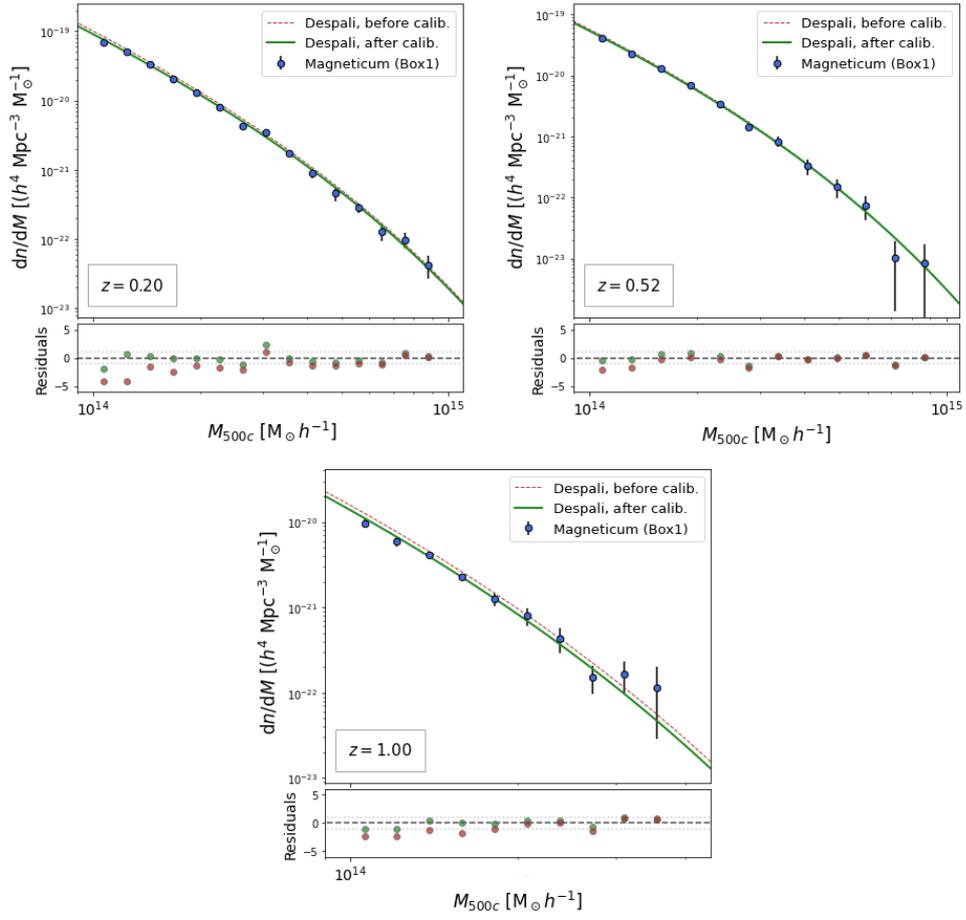
Figure 5.4: As in Fig. 5.2 but for galaxy clusters having $M_{500c} \geq 10^{14} \, M_{\odot} h^{-1}$.

in Section 3.6.2 in order to extract the void catalogues, and we cleaned them with the algorithm described in Section 3.7 and exploited in other recent works (Ronconi & Marulli, 2017; Ronconi et al., 2019; Contarini et al., 2019, 2020).

As did for the HMF, also the VSF model has to be calibrated in order to allow a comparison to the simulated data. We make use of the Vdn model described in Section 3.6, which predicts only the number densities of voids, as a function of their radius, in an un-biased DM distribution. However, we deal with voids detected from galaxy catalogues, thus it is necessary to take into account the bias of galaxies inside them, as described in Section 3.6.4, by modifying properly the shell-crossing threshold that enters in the model. The procedure described in Section 3.6.4, and exploited in Contarini et al. (2019, 2020), considers DM haloes as biased tracers of the matter distribution. Therefore we need to re-calibrate the $\mathcal{F}(b_{\mathrm{eff}})$ relation between the effective bias computed at large scales (b_{eff}) and the bias computed inside cosmic voids (b_{punct}, Eq. 3.58).

Let us write this relation in a general linear form:

$$b_{\mathrm{punct}} \equiv \mathcal{F}(b_{\mathrm{eff}}) = b_{\mathrm{slope}} \cdot b_{\mathrm{eff}} + b_{\mathrm{off}} , \qquad (5.4)$$

where b_{slope} and b_{off} are the slope and the offset of this relation, and $b_{\mathrm{eff}} = b_{\mathrm{eff}}(z)$, where z is the redshift. Similarly to what we have done with the Despali parameters, here the re-calibration consists into finding new values for the parameters b_{slope} and b_{off} in the above relation.

In order to do this, we have to measure b_{eff} from the galaxy distribution of our catalogues. We follow the same prescriptions as in Marulli et al. (2018) and Contarini et al. (2019), measuring the 2PCF of galaxies in our simulated catalogues and performing a Bayesian statistical analysis in order to infer the value of b_{eff} at different redshifts.

In order to estimate the 2PCF of galaxies we use the Landy-Szalay estimator given by Eq. (2.33), which provides an un-biased estimator of $\xi(r)$ in the limit of $N_R \to \infty$, where N_R is the total number of objects in the random catalogues. We constructed the latter to be three times larger than the Magneticum galaxy catalogues, preserving their geometry and their three-dimensional coverage, in order to limit the poissonian error in the data-random counts with respect to the errors from the data-data counts. We divided the data of the 2PCF in 20 comoving separation bins, considering them to be inside the range of interest for voids, that is 10 Mpc $h^{-1} < r < 60$ Mpc h^{-1}, at all the redshifts considered.

We finally performed a Bayesian MCMC analysis on the 2PCF, consid-

ering the following Gaussian likelihood function $\mathcal{L}$:

$$-2\ln\mathcal{L} = \sum_{i=1}^{N}\sum_{j=1}^{N}(\xi_i^d - \xi_j^m)C_{i,j}^{-1}(\xi_j^d - \xi_j^m) \ , \tag{5.5}$$

where $C_{i,j}$ is the covariance matrix, which measures the correlation between the different bins of the 2PCF, N is the number of bins and the superscripts d and m stand for *data* and *model*, respectively. The covariance matrix is computed with the Bootstrap method (Efron, 1979), by dividing each data catalogues in 125 sub-catalogues, constructing $n_R = 100$ realisations by resampling from the sub-catalogues. Specifically, it is defined as

$$C_{i,j} = \mathcal{F}\sum_{k=1}^{n_R}(\xi_i^k - \hat{\xi}_i)(\xi_j^k - \hat{\xi}_j) \ , \tag{5.6}$$

where k runs over the different realisations while i and j run over the 2PCF bins. Here $\hat{\xi}$ represents the average of the 2PCF of the n_R samples and $\mathcal{F}$ is the normalisation factor, necessary in order to take into account the possible correlation between the different n_R realisations (Norberg et al., 2009), which is $\mathcal{F} = 1/(n_R - 1)$ for the Bootstrap method.

The 2PCF model of galaxies $\xi^m(r)$ is computed as

$$\xi^m(r) = b_{\text{eff}}^2\xi_{DM}(r) \ , \tag{5.7}$$

where $\xi_{DM}(r)$ is the DM 2PCF, which we computed by Fourier transforming the DM power spectrum $P_{DM}(k)$, measured using the Code for Anisotropies in the Microwave Background (CAMB)[2]. The results are shown in Figure 5.5, where the data points represent the square root of the ratio between the 2PCF of galaxies and that of DM. The values of b_{eff} used for the VSF calibration are estimated as the best-fit values obtained in the range 20 Mpc h^{-1} < r < 40 Mpc h^{-1}, where the data are more stable. These are $b_{\text{eff}} = \{1.32 \pm 0.02, 1.54 \pm 0.02, 1.98 \pm 0.02, 3.16 \pm 0.04\}$ for the four redshifts considered.

We can now move to the re-calibration of the $\mathcal{F}(b_{\text{eff}})$ relation, which can be performed by sampling the posterior distribution of the parameters b_{slope} and b_{off} via a Bayesian statistical analysis. We tested three different effective radius cuts: $R_{\text{eff}} \geq 2.5\lambda_{mps}(z)$, $R_{\text{eff}} \geq 2.75\lambda_{mps}(z)$ and $R_{\text{eff}} \geq 3\lambda_{mps}(z)$, where λ_{mps} is the mean particle separation of galaxies in the catalogues, which depends on the given redshift. In particular, we found $\lambda_{mps} = 6.55, 6.92, 7.81, 10.52$ Mpc h^{-1}. We let b_{slope} and b_{off} to vary freely between flat priors, considering a Gaussian likelihood function. All the data-sets
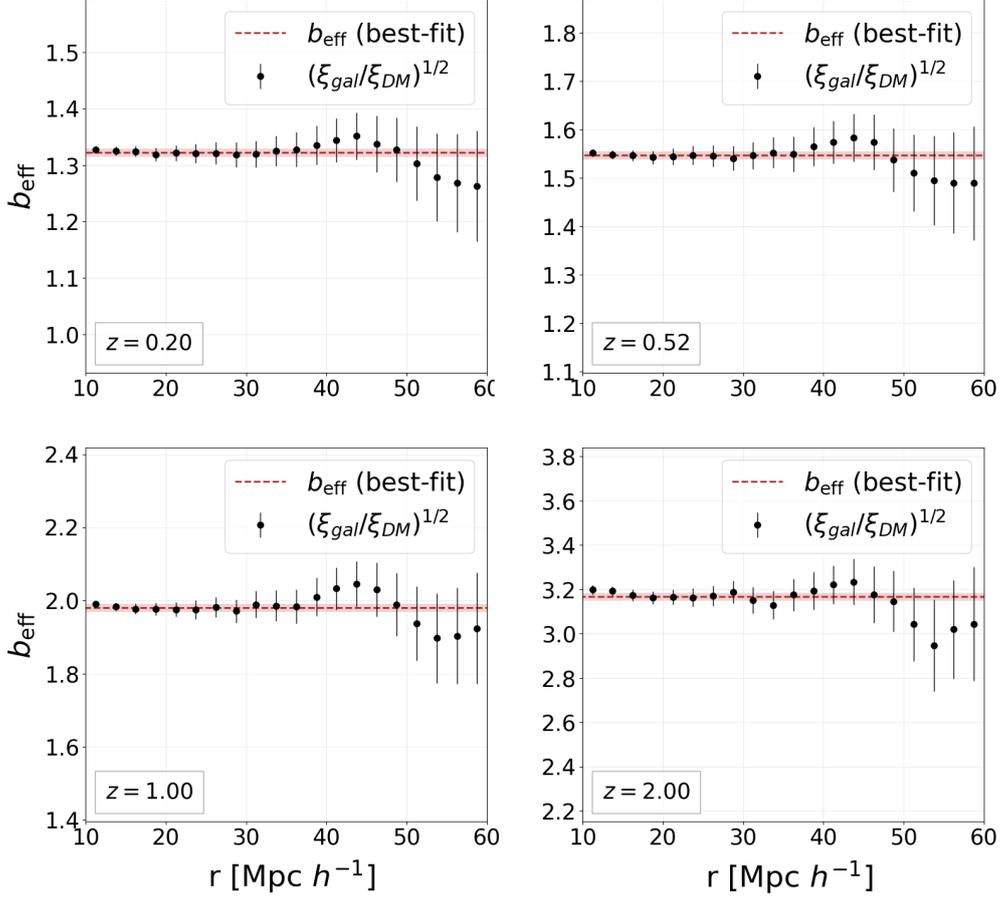
---

[2]For information about CAMB: https://camb.info/

Figure 5.5: The large-scale bias measured in the galaxy catalogues, having minimum mass $M_{min} = 10^{10}\ M_\odot h^{-1}$, extracted from the Magneticum simulation Box1 at redshift $z = 0.2, 0.52, 1, 2$. The black dots represent the square root of the ratio between the 2PCF of galaxies, obtained with the Landy-Szalay estimator, and the 2PCF of the underlying DM matter distribution, computed with CAMB. The errorbars correspond to the diagonal elements of the covariance matrix (Eq. 5.6). The dashed red lines represent the best-fit obtained by fitting the data between $r = 20$ Mpc h^{-1} and $r = 40$ Mpc h^{-1}, while the dashed red areas represent the 1σ uncertainties of each best-fit.
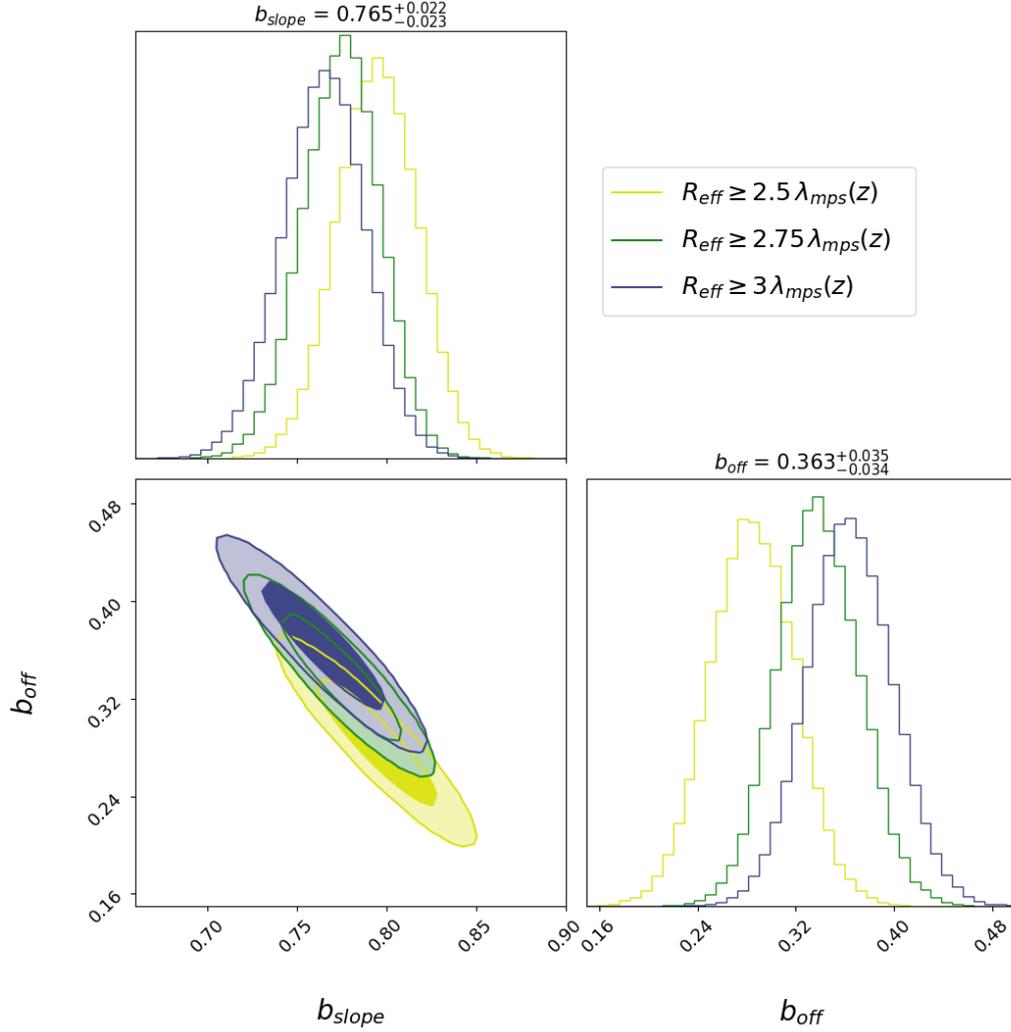
Figure 5.6: Posterior probability distribution of b_{slope} and b_{off}, coefficients of the relation $\mathcal{F}(b_{\text{eff}})$ (Eq. 5.4), for the three different minimum radii considered. Here $\lambda_{mps}(z)$, which depends on redshift, represents the mean particle separation of galaxies in the simulated catalogues. Dark and light areas show the 68% and 95% confidence regions, while at the top of each column we report the mean and the standard deviation of the 1D marginal distribution for the strongest radius cut.

at different redshifts have been treated as independent, and we marginalised over the uncertainty of b_{eff}. As done for the HMF re-calibration, in this procedure we fix the underlying cosmological model at the WMAP7 cosmological parameters. The results of the re-calibration are shown in Figure 5.6. The new bias relations found with this calibration, depending on the low radii cut, are:

$$\mathcal{F}(b_{\text{eff}}) = (0.80 \pm 0.02)b_{\text{eff}} + (0.28 \pm 0.03), \quad R_{\text{eff}} \geq 2.5\lambda_{mps}(z)$$
$$\mathcal{F}(b_{\text{eff}}) = (0.78 \pm 0.02)b_{\text{eff}} + (0.34 \pm 0.04), \quad R_{\text{eff}} \geq 2.75\lambda_{mps}(z)$$
$$\mathcal{F}(b_{\text{eff}}) = (0.77 \pm 0.02)b_{\text{eff}} + (0.36 \pm 0.04), \quad R_{\text{eff}} \geq 3\lambda_{mps}(z) \ .$$

As one can see, the slope of the relation grows for higher cuts, while the offset decreases. Nevertheless, the uncertainties on the two parameters remain the same for every cut. Even if the parameters are all consistent with each other in all the three cases, we found a strong dependence on the radius cut for the $\mathcal{F}(b_{\text{eff}})$ relation.

5.4 Constraints from cluster counts and void counts

Comparing the HMF and VSF models with the simulated data-sets, we can finally extract constraints on the cosmological parameters Ω_m and σ_8. As already explained, numerical densities as a function of the mass, for the HMF, and the radius, for the VSF, are considered as statistically independent, meaning that we do not take into account the cross-covariance matrix between the two probes (see Appendix A for the estimation of the cross-covariance matrix). The constraints are obtained by concatenating the data-sets at different redshifts into a unique vector. Consequently, we perform a Bayesian MCMC analysis on it, considering flat priors on Ω_m and σ_8, and marginalising over the non-cosmological parameters, considering them as nuisances. Regarding the latter, i.e. (a_i, A_i) for the HMF model and $(b_{\text{slope}}, b_{\text{off}})$ for the VSF model, we consider a multivariate normal distribution as their prior, centred on their average values obtained from the calibrations, specifying their covariance matrix. In this way, we are able to take into account all the correlations between the nuisance parameters, which arise in the calibration procedure.

We consider the errors on our data as Gaussian distributed, and we run the MCMC in order to sample the joint posterior distribution of Ω_m and σ_8. We show the resulting contours for the HMF model for the two different minimum mass considered, $M_{500c} \geq 5\times10^{13}\ M_\odot h^{-1}$ and $M_{500c} \geq 10^{14}\ M_\odot h^{-1}$,
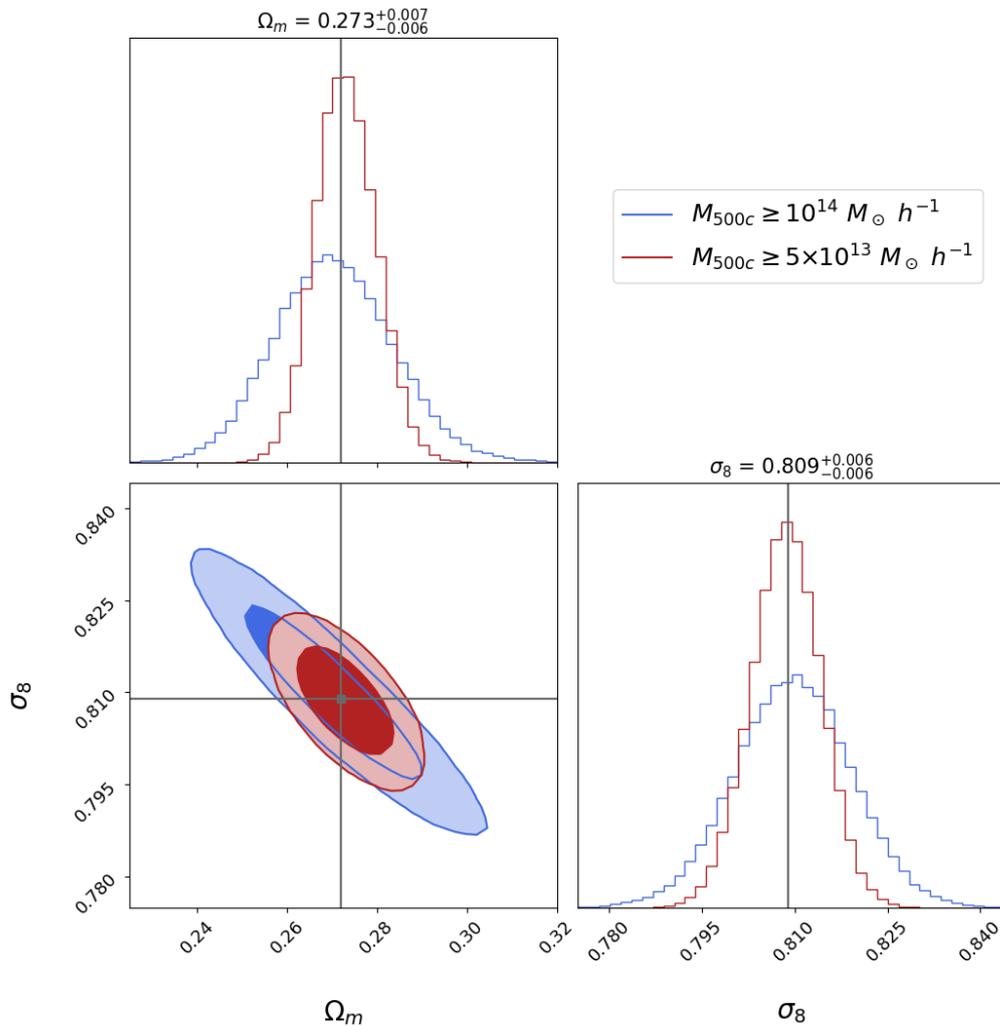
Figure 5.7: Constraints on Ω_m and σ_8 from the HMF, considering two different minimum masses for the clusters in our simulated catalogues. The red contour shows the posterior probability, marginalised over the Despali parameters (a, A), for $M_{500c} \geq 5 \times 10^{13}\ M_\odot h^{-1}$, while the posterior probability in the case of $M_{500c} \geq 10^{14}\ M_\odot h^{-1}$ is represented in blue. The solid gray lines correspond to $\Omega_m = 0.272$ and $\sigma_8 = 0.809$, representing the truth underlying cosmological model of the simulations (WMAP-7).
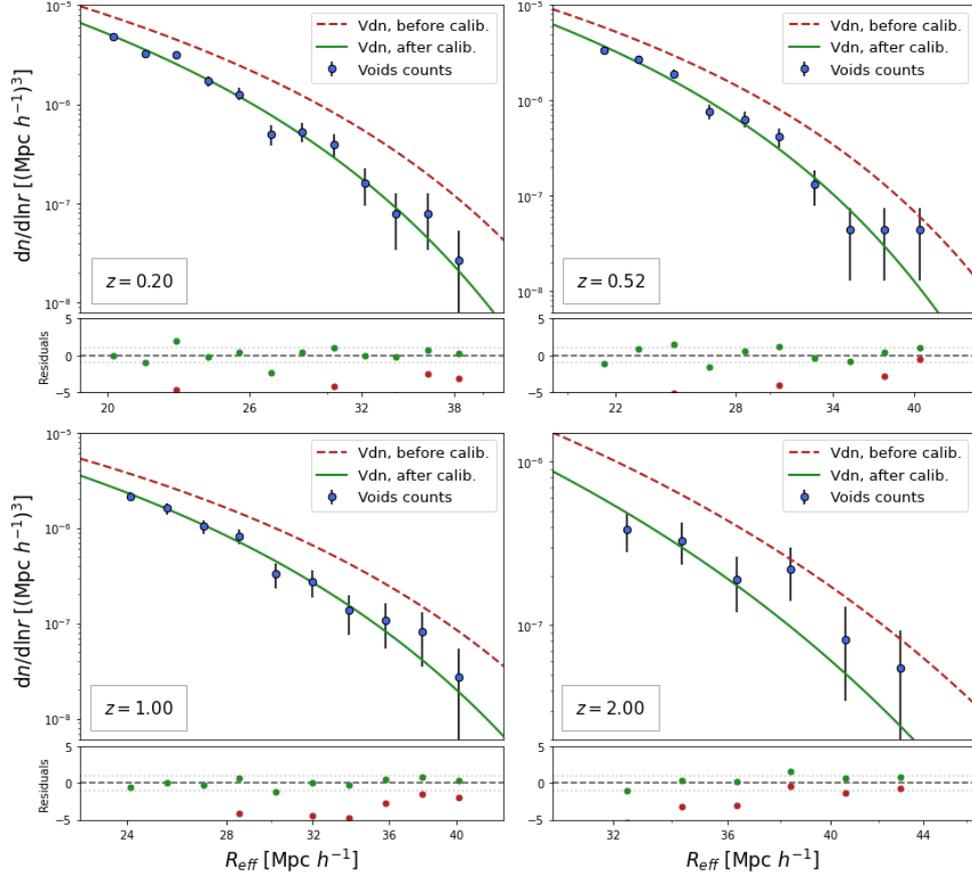
Figure 5.8: The measured size function of voids (blue dots) identified in the distribution of galaxies, from the Magneticum Simulation (Box1) at redshifts $z = 0.2, 0.52, 1, 2$, for $R_{\mathrm{eff}} \geq 3.0 \cdot \lambda_{mps}(z)$, where $\lambda_{mps}(z)$ is the mean particle separation of the simulated galaxies. The green lines represent the Vdn model, properly re-scaled in order to compare it with the data. The dashed red lines represent the Vdn model rescaled using the calibration performed in Contarini et al. (2019) for DM haloes. The lower sub-panels report the residuals of the void counts, computed as the ratio between the difference *data - model* and the *data errors*, which are Poissonian.

in Figure 5.7, while in Figure 5.9 are shown the different contours obtained for the three cuts described above. Unlike the HMF contours, which remain centred on the truth values for Ω_m and σ_8, the VSF contours approach the latter when considering the highest radius cut. Indeed, only the $R_{eff} \geq 3.0 \cdot \lambda_{mps}(z)$ cut presents a good centring, while for the others the truth

values lie outside the 95% confidence regions, though they shrink up to a factor of 2.

In our case we know a-priori the truth values for the cosmological parameters, therefore we choose to keep only the strongest cut, in order to properly compare it with the HMF contour. The comparison between the re-scaled VSF model and the simulated numerical densities of voids for $R_{eff} \geq 3.0 \cdot \lambda_{mps}(z)$ can be seen in Figure 5.8. The two final contours, which can be seen together in Figure 5.10, show different type of degeneracies, demostrated by the fact that they have different orientations in the $\Omega_m - \sigma_8$ parameter plane. In particular, they are almost perpendicular, which is a powerful feature when we consider their combination. The parameter values obtained by marginalising over the Despali parameters, constraining the HMF alone, are:

$$\Omega_m = 0.273 \pm 0.007,$$
$$\sigma_8 = 0.808 \pm 0.006, \tag{5.8}$$

for the weakest mass selection ($M_{500c} \geq 5 \times 10^{13} \ M_\odot h^{-1}$) and

$$\Omega_m = 0.270 \pm 0.013,$$
$$\sigma_8 = 0.810 \pm 0.009 \tag{5.9}$$

for the strongest one ($M_{500c} \geq 10^{14} \ M_\odot h^{-1}$).

The constraints obtained by exploiting the VSF alone for voids having radius $R_{eff} \geq 3.0 \cdot \lambda_{mps}(z)$, marginalising over b_{slope} and b_{off}, are:

$$\Omega_m = 0.266 \pm 0.016,$$
$$\sigma_8 = 0.803 \pm 0.015 \ . \tag{5.10}$$

For both models the truth values $\Omega_m = 0.272$ and $\sigma_8 = 0.809$ are recovered within 1σ. The VSF contour is marginally decentred in Ω_m, even though it is broader than the HMF contours. The uncertainties on the two cosmological parameters are comparable for the strongest mass cut applied, while for clusters having mass $M_{500c} \geq 5 \times 10^{13} \ M_\odot h^{-1}$, the VSF contour is $2-3$ times broader than the HMF one.

5.5   Constraints from the probe combinations

Finally, we are ready to combine the two cosmological probes applying the three methods implemented in this Thesis work and described in the previous Chapter: *Posteriors Product*, *Importance Sampling* and *Posterior as Prior* methods. All the results of the combination methods can be seen together

in Figure 5.11, for clusters with $M_{500c} \geq 5 \times 10^{13} \ M_\odot h^{-1}$, and in Figure 5.12 for the other mass selection. Starting with the classic *Posterior Product* method, we sample the posterior distribution of Ω_m and σ_8 by considering flat priors on them, and marginalising over all the nuisance parameters for the two models, that are Despali post-calibration and Vdn post-calibration. For each point of the parameter space in the sampling procedure, the combined posterior value in that point is the product of the likelihood functions of the two models, times the prior, following Eq. (4.13). For the *Importance Sampling* method, we compute the importance weight for each point of the sampling procedure by considering firstly the VSF posterior distribution as the importance distribution and then we repeat for the HMF posterior. By concatenating these distributions along with the importance weights specified for each point, we are able to compute the average of the weighted points in the parameter space as well as the weighted standard deviation. Even though the combination methods are all consistent with each other, the standard deviation for the IS results slightly wider than that of the posteriors product, although the difference is in the fourth digit. Lastly, the *Posterior as Prior* method is performed by considering multivariate Gaussian priors for the joint distribution of Ω_m and σ_8, specifying their covariance matrix which derives from the calibration procedures, first in the case of the VSF and then for the HMF calibration. The parameters of the model does not enter in the multivariate prior are considered as nuisance, and we marginalise on them in order to obtain the final constraints. This method is applicable in two directions, by considering one probe's posterior as the prior of the other and vice-versa. However, this method is symmetric, and the final constraints are practically the same, as one can see in Figure 5.13 where we compare the methods performed in the two directions for the highest mass cut. The results are well consistent with the *Posterior Product* method, and with the latter they represent the methods with the most strong constraining power. Our best guess for the constraints, thus, are the following:

$$\Omega_m = 0.271 \pm 0.004$$
$$\sigma_8 = 0.809 \pm 0.003 \ . \tag{5.11}$$

Considering the different mass cuts for the galaxy clusters, the overall combination contours are well placed in the concordance region of the two cosmological probes, and they are very well centred on the truth values of the simulation. Moreover, the average posterior values as well as their standard deviations for the combination contours in the two cases are consistent with each other. Comparing the last results with those obtained by exploiting individually the HMF and VSF models, the combined constraints show a relative improvement of about $4 - 5$ times in terms of standard deviation.
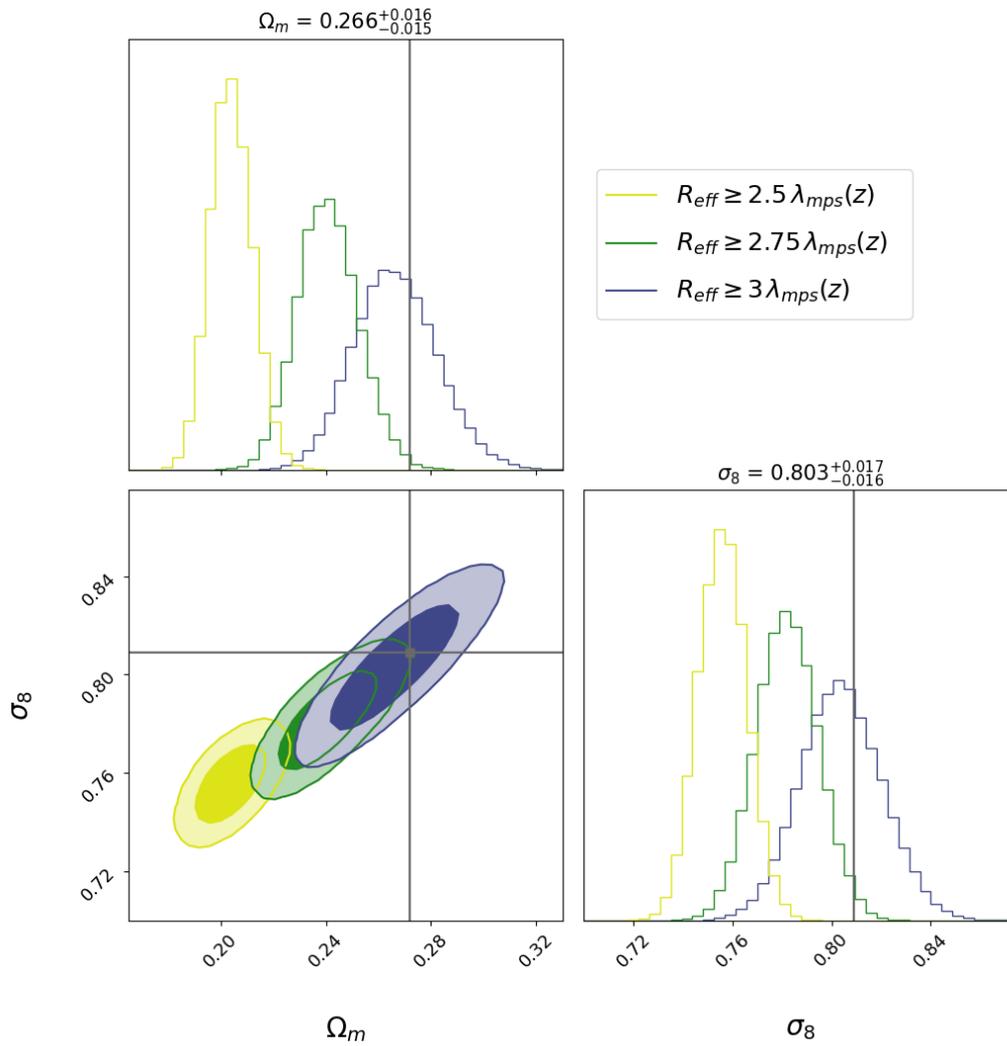
Figure 5.9: Constraints on Ω_m and σ_8 from the VSF, considering three different minimum effective radii.
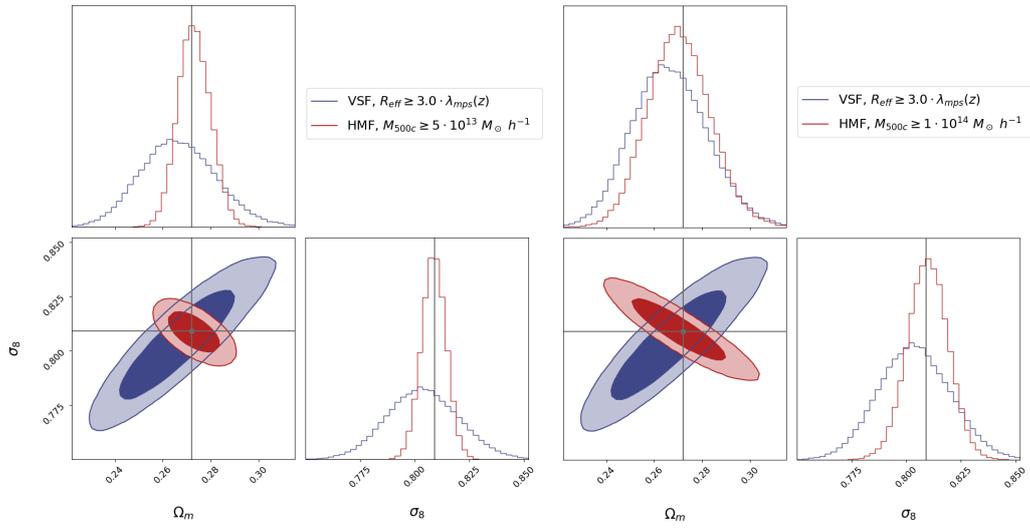
Figure 5.10: Comparison of the constraints on Ω_m and σ_8 obtained by marginalising over the nuisance parameters of the HMF and VSF models. The left contour plot shows in red the constraints for the HMF considering a minimum mass for the clusters sample of $M_{500c} \geq 5 \times 10^{13}\ M_\odot h^{-1}$, while for the right panels the minimum mass considered is $M_{500c} \geq 10^{14}\ M_\odot h^{-1}$. For the VSF contour, which is shown in blue, the effective radius considered is $R_{\mathrm{eff}} \geq 3\lambda_{\mathrm{mps}}(z)$ Mpc h^{-1} for both plots. Dark and light areas show the 68% and 95% confidence regions.
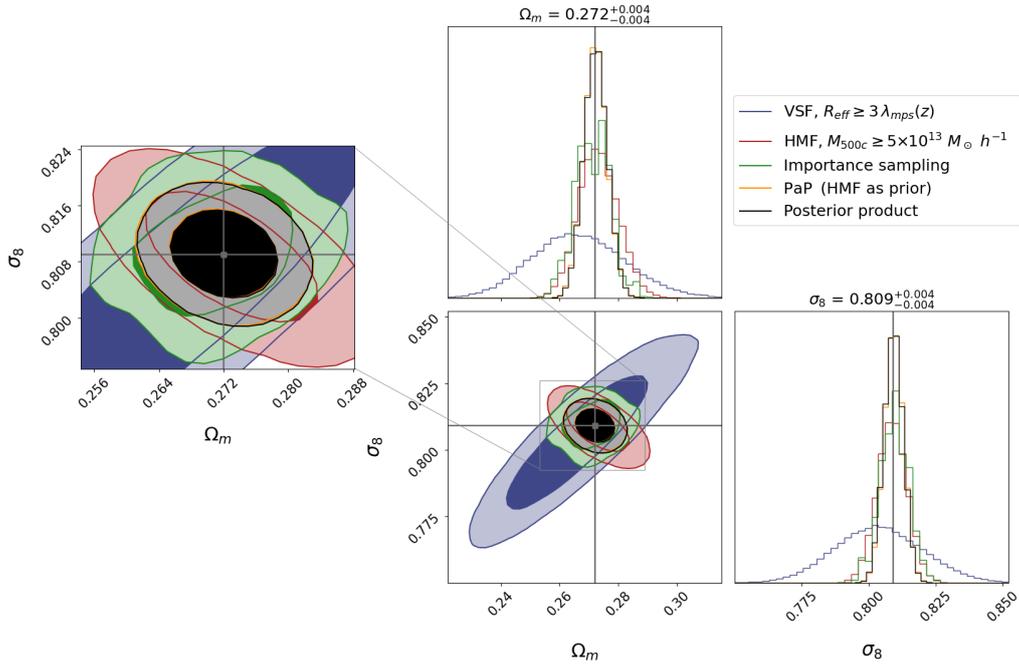
Figure 5.11: Combination contours for the methods implemented in this Thesis work, for the lowest mass cut considered ($M_{500c} \geq 5 \times 10^{13} \; M_\odot h^{-1}$). The green contour represents the posterior distribution obtained with the *Importance Sampling* method (read from external chains), while the *Posterior as Prior* and the posteriors product methods are represented by orange and black contours, respectively. The blue contour represents the VSF posterior distribution while the HMF one is represented in red.
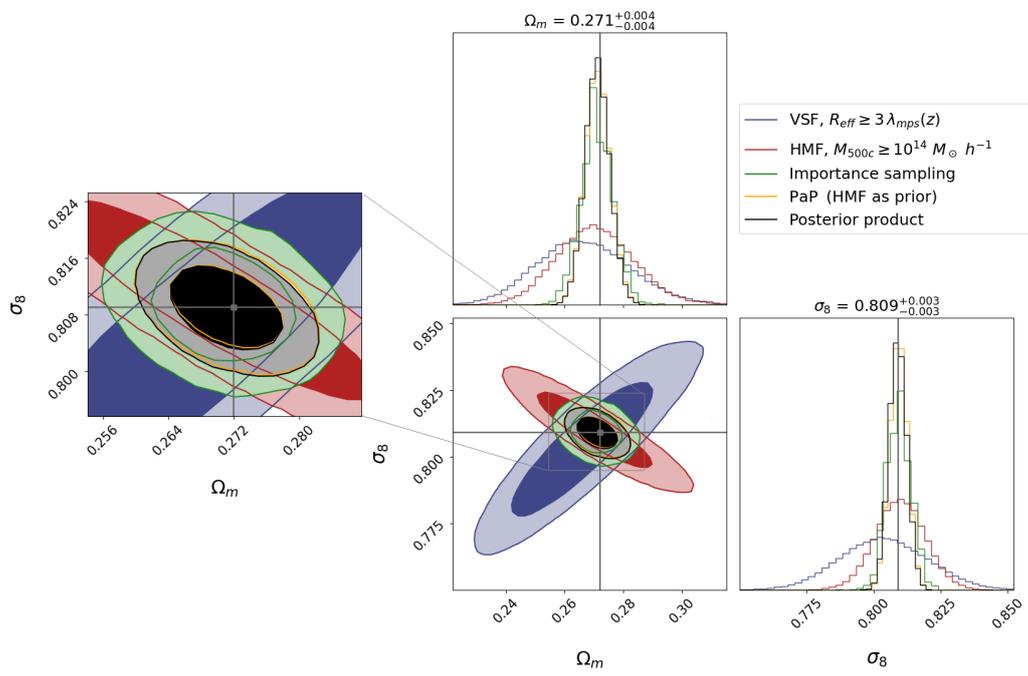
Figure 5.12: As Fig. 5.11 but for the case $M_{500c} \geq 10^{14}\ M_\odot h^{-1}$.
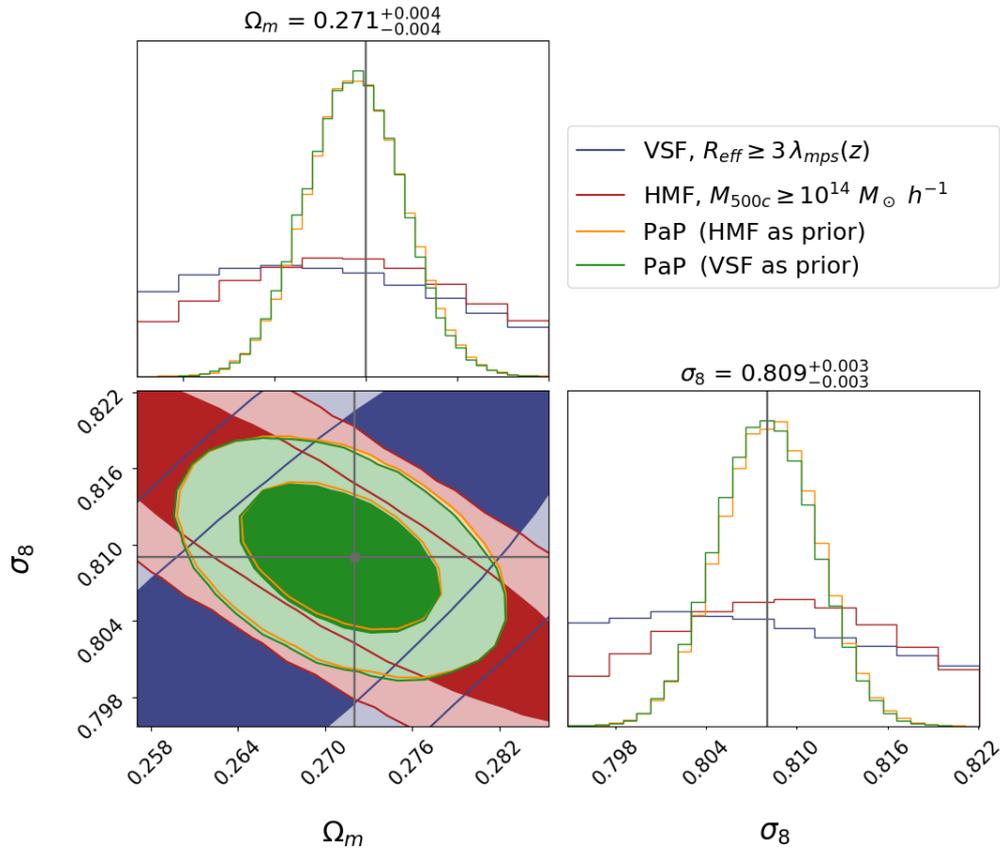
Figure 5.13: Comparison between the results obtained with the *Posterior as Prior* (PaP) methods. The green contour represents the PaP method when considering the HMF posterior probability distribution on $\Omega_m - \sigma_8$ as the prior for the VSF probe. The orange contour shows the same method but in the opposite direction, i.e when considering the VSF posterior as the HMF prior.

Chapter 6

Discussion and conclusions

6.1 The scientific problem

The last decades have seen enormous advances in the study of the large-scale structures, which have deepened our knowledge of the spatial properties of the Universe and their evolution with cosmic time. Deep cosmological surveys have mapped, with an astonishing precision, the distribution of visible matter in the local Universe, allowing for high-accuracy measurements of the cosmological parameters. Among the most famous achievements, there are the results from the Planck mission (Planck Collaboration et al., 2020). This satellite investigated the Universe when it was only 400'000 years old, analysing in great details the cosmic microwave background anisotropies, shedding light on the mechanism that initially generated the cosmic-web pattern that we observe today. Moreover, large cosmological simulations, realised thanks to the increasing power of modern computers, are capable to produce more and more realistic representations of our Universe. Being able to build these simulations with different sizes allows us to study in details the astrophysical processes that are at the base of the formation and evolution of galaxies, as well as to perform accurate statistical analyses of the distribution of large-scale structures.

The model that better predicts the properties of our Universe on large scales is nowadays the ΛCDM, which has been continuously tested with increasingly precise measurements, becoming the most popular and widespread model in Cosmology. In this framework, the Universe is modelled as a Friedmann-Lemaître-Robertson-Walker expanding spacetime, in which early density perturbations grow as a consequence of the attractive nature of gravity, leading to the formation of haloes, i.e. clumps of matter, inside which stars, as well as galaxies, form and evolve. The latter, over cosmic time,

143

create the most massive, gravitationally bounded objects that we observe today, i.e. galaxy clusters. These structures usually contain thousands of galaxies in a compact volume of few Mpc. Being the most massive virialised objects in the Universe, galaxy clusters can be exploited in order to put tight constraints on the main cosmological parameters, to better understand both the expansion history of our Universe and its matter content. Indeed, this can be achieved by studying their number counts as well as their clustering properties and density profiles over different redshifts.

While galaxy clusters are associated with the growth of posiive perturbations of the spacetime, the growth of depressions in it leads to the formation of cosmic voids, i.e. the most underdense regions of our Universe. These objects, having radii up to tens of Mpc, fill the remaining volume of the Universe, and are becoming a competitive cosmological probe when studying their number counts, their density profiles, or the redshift-space distortions of galaxies within them. Moreover, as being limited to the physical minimum for the density contrast $\delta = -1$, these objects remain mildly nonlinear during their evolution, in contrast with their overdense counterpart (dark matter haloes) which can reach extremely nonlinear amplitudes. Being objects almost empty, voids are very sensitive to the effect of dark energy and the most elusive kind of matter such as neutrinos. Indeed, voids are also considered to be promising laboratories for testing alternative models of dark energy and modified gravity theories. Nonetheless, a complete theoretical understanding about their life-cycle still lacks. In fact, we do not know how to model their growth rates as well as their formation and merger history. Moreover, though void size function (VSF) models have been proposed during the years (see Section 3.6), they are only applicable to un-biased distributions of matter, and iare inadequate when predicting the number counts of voids identified in biased tracers distributions, like e.g. galaxies. In order to exploit the VSF as a cosmological probe, we have to re-scale the VSF consistently with the bias of the tracers, as show by Contarini et al. (2019).

In this Thesis work we presented the implementation of new codes inside the publicly available C++/Python libraries CosmoBolognaLib, which allow the user to perform a combined analysis of multiple independent cosmological probes. The combination analysis is a very powerful tool, useful to achieve strong constraints on cosmological parameters, and has been widely performed in the last decades in several works. The great precision in measuring cosmological parameters has led to some observational tensions in the ΛCDM between the estimated values of the same parameters from different probes (see Section 1.10.1). In particular, these tensions arise by comparing measurements of the early Universe, such as the data coming from the CMB anisotropies, and low-z cosmological probes, paving the way for the

requirement to include new physics in the Standard Cosmological Model. In this context, the combination technique can shed new light on these tensions, disentangling the degeneracies between parameters by providing tighter constraints on them. In addition, since the ΛCDM model is dominated by dark components, most of the fundamental questions of Cosmology remain unanswered. Probably, the biggest problem on which we still have to investigate is the evidence for the presence of a dark field, which contributes to about 70 % of the Universe's content, responsible for its present accelerated expansion.

6.2 Procedure and results

In this work, we exploited the combination of two cosmological probes: the mass function of galaxy clusters and the size function of cosmic voids. The data have been obtained by analysing a simulated volume of $(896\ \mathrm{Mpc}\ h^{-1})^3$ from the Magneticum Pathfinder simulations at 4 different redshifts in the range $0.2 < z < 2.0$. The simulated galaxy cluster number densities have been compared to the theoretical halo mass function (HMF) model described in Section 3.5.2 and presented in Despali et al. (2016). The theoretical model has been previously calibrated using the Magneticum simulations, given the differences between the halo finder algorithms employed to calibrate the model presented in the literature and the one used to detect clusters in these simulations. We presented our analysis by considering two mass cuts for the clusters in our catalogues, specifically $M_{500c} \geq 5 \times 10^{13}\ M_\odot\ h^{-1}$ and $M_{500c} \geq 10^{14}\ M_\odot\ h^{-1}$, which act as simplified selection functions, since no observable-mass relations have been considered in this work. Regarding voids, we identified them in the simulated distributions of galaxies, by considering the latter having stellar mass $M_* \geq 10^{10}\ M_\odot\ h^{-1}$. We built the void catalogues by running the void finder algorithm VIDE, in combination with the cleaning procedure described in Ronconi & Marulli (2017) and introduced in Jennings et al. (2013), aimed at preparing these catalogues to match the theoretical definition used to develop the void size function model. We considered only voids having an effective radius 3 times larger than the mean particle separation of galaxies in our catalogues, and we show how a different choice for this minimum radius could give biased constraints for the cosmological parameters. We compared the cleaned void number densities with the theoretical Vdn model developed in Jennings et al. (2013) and described in Section 5.8. As did for the HMF, we calibrated the VSF model in order to take into account the effect of the tracer bias on the void abundance. For this purpose, we followed the prescriptions presented in Contarini et al. (2019, 2020): we measured the effective large-scale bias of galaxies in

our catalogues, b_{eff}, and computed a re-parameterisation of the characteristic underdensity threshold of the Vdn model, by means of a linear function of the bias b_{eff}. This new calibration of the model allows us to predict the abundance of voids traced by galaxies unlike the one proposed in Contarini et al. (2019, 2020), who performed the same calibration for void identified in the DM halo distribution.

Then, by performing a Bayesian MCMC analysis, we obtained the posterior probability distributions for the matter density parameter Ω_m and the today perturbation amplitude, σ_8, marginalising over the nuisance parameters of the considered models. We showed that the HMF and the void size function can be considered as complementary probes, providing a nearly perpendicular intersection in the $\Omega_m - \sigma_8$ parameter space, which is a desired property when we want to extract as much information as possible from the probe combination. During this analysis, we considered the and the VSF as independent probes, since they model the peaks and the depths of the matter density field, respectively. This hypothesis has been tested in Appendix A by estimating the cross-covariance matrix between the different data-sets analysed.

In order to obtain the final combined contours in parameter space, we applied three new numerical methods, implemented inside the CosmoBolognaLib C++/Python libraries. In particular, the most tight constraints are those deriving from the application of the *Posterior Product* and *Posterior as Prior* methods. Following the former, the final posterior distribution for the parameter of interest is computed as the product of the single likelihood functions times a unique prior vector (see Eq. 4.13). In the latter, instead, the posterior distribution obtained by exploiting one cosmological probe is used as the prior distribution for the other. We implemented also a third method, the *Importance Sampling* technique, in which one extracts information about a given probability distribution, for a certain set of parameters, by sampling from another distribution. Following this procedure, we obtain the combined posterior distributions of two probes by performing a concatenation of their resulting MCMC chains, with specific weights. In particular, with this method we associate to each step of the chains an *importance weight*, defined as the ratio of the two un-normalised posterior distributions in the corresponding point of the parameter space. Though the latter method presents a weaker constraining power, it is totally in agreement with the others in terms of posterior mean and standard deviation. Moreover, the *Importance Sampling* is the fastest among the methods presented, and it can be applied to whatever external posterior distributions.

All the implemented methods gave consistent results and, by combining number counts of galaxy clusters and cosmic voids, we achieved significantly

tighter constraints on the cosmological parameters, with respect to those obtained with the single cosmological probes.

6.3   Future perspectives

Beyond the independence assumptions we have described so far, there are several caveats in this work. Firstly, we assumed a perfect knowledge about the distribution of tracers in the simulated catalogues. Dealing with real data catalogues we have to take into account the errors associated to the mass and the position of the tracers, as well as the effects of peculiar velocity of the these objects and the uncertainties related to the cosmological parameters used to compute cosmological distances, thus redshift-space and geometrical distortions. In addition, we do not take into account any observable-mass scaling relation for the mass of the considered simulated tracers (both galaxies and galaxy clusters), as if it were perfectly measured. We also neglected the super-sample covariance. The latter concerns the fact that our simulations or observations map a limited portion of the Universe, from which the measure of the background density of matter can be intrinsically biased. This is caused by the fact that one usually does not consider perturbations having Fourier modes larger than the size of the simulation/survey, potentially causing a change in the relative improvements of the combination analysis. Finally, despite our study was limited to the constraints on Ω_m and σ_8, it will be fundamental to consider all the remaining cosmological parameters on which the analysed models strongly depend. We will address all these issues in future works.

In order to obtain well grounded statistical results from the analysis of large-scale structures, it would be interesting to search for galaxy clusters and cosmic voids in much larger cosmological volumes. Upcoming galaxy surveys like Euclid and DESI will map a very large fraction of our Universe, with volumes of the order of 10^2 Gpc h^{-1}, improving our knowledge about the underlying cosmological model. Moreover, in order to improve the combination analysis of multiple cosmological probes, it will be necessary to add the possibility of setting the cross-covariance matrix to the implemented codes presented in this work, in order to extend the combination analysis also to dependent cosmological probes.

Appendix A

Estimation of cross-correlations between galaxy cluster and void number counts

In this Thesis work, we based our combined analysis on two assumptions, the independence of the data-sets between the redshifts, and the independence between the number densities of galaxy clusters and cosmic voids. These assumptions can be tested by estimating their cross-covariance matrix and check whether its off-diagonal elements are consistent with zero.

Instead of computing the covariance matrix theoretically, one can opt for two alternative ways. The first concerns the estimation from a number of independent realisations of the same cosmological simulation, while the second concerns the Jackknife or Bootstrap re-sampling of the observed (or simulated) data itself. These latter methods can be used to generate new hypothetical samples that are representative of an underlying population, i.e. the analysed catalogues in this work, whose statistics we want to estimate. In this Appendix we make use of the Jackknife technique, which we are going to describe.

Let us assume to divide the catalogue volume into $N_{\rm sub}$ equally sized sub-regions, where $\alpha \sim (i_M, i_z)$ is defined as the generic bin in mass and redshift, obtained from cluster catalogues, and $\beta \sim (i_R, i_z)$ is the generic bin in radius and redshift, obtained from the void catalogues. A value for the number densities is associated for each of these bins, as well as the pure number counts, of clusters (N_α) and voids (N_β). The Jackknife technique works by sequentially deleting one sub-sample from the entire data-set, re-computing the desired statistic on the remaining samples. To do so, we can define a

Jackknife sample to be the total simulated volume minus one subsample:

$$N_\alpha^{\rm jk}(i) = N_\alpha^{\rm tot} - N_\alpha(i). \tag{A.1}$$

Consequently, the Jackknife covariance matrix between two generic bins can be estimated as follows:

$$\mathcal{C}^{\rm jk}(N_\alpha, N_\beta) = \frac{N_{\rm sub}}{N_{\rm sub} - 1} \sum_{i=1}^{N_{\rm sub}} \left(N_\alpha^{\rm jk}(i) - \bar{N}_\alpha^{\rm jk}\right)\left(N_\beta^{\rm jk}(i) - \bar{N}_\beta^{\rm jk}\right), \tag{A.2}$$

where:

$$\bar{N}_\alpha^{\rm jk} = \frac{1}{N_{\rm sub}} \sum_{i=1}^{N_{\rm sub}} N_\alpha^{\rm jk}(i) \tag{A.3}$$

is the average of the Jackknife counts in a specific bin.

The cross-covariance matrix, as expressed in Eq. (A.2), is a block-matrix, i.e. it is subdivided in submatrices. It takes into account all the correlations between cluster counts and void counts at different redshifts. Consequently, the total cross-covariance matrix can be written as:

$$\mathcal{C}^{\rm jk}(N_\alpha, N_\beta) = \left(\begin{array}{c|c} \mathcal{C}_{H,H} & \mathcal{C}_{V,H} \\ \hline \mathcal{C}_{H,V} & \mathcal{C}_{V,V} \end{array} \right), \tag{A.4}$$

where $\mathcal{C}_{H,H}, \mathcal{C}_{V,H}, \mathcal{C}_{H,V}$ and $\mathcal{C}_{V,V}$ are the covariance matrices between mass bins, mass bins and radius bins (and viceversa) and radius bins at all redshifts ($z = 0.2, 0.52, 1, 2$), respectively. All these matrices are themselves block matrices, with the different elements representing the covariances between different redshift data-sets. For example, considering $\mathcal{C}_{H,H}$, this is:

$$\mathcal{C}_{H,H} = \begin{pmatrix} \mathcal{C}_{\rm HMF}^{0.2,0.2} & \mathcal{C}_{\rm HMF}^{0.52,0.2} & \mathcal{C}_{\rm HMF}^{1,0.2} & \mathcal{C}_{\rm HMF}^{2,0.2} \\ \mathcal{C}_{\rm HMF}^{0.2,0.52} & \mathcal{C}_{\rm HMF}^{0.52,0.52} & \mathcal{C}_{\rm HMF}^{1,0.52} & \mathcal{C}_{\rm HMF}^{2,0.52} \\ \mathcal{C}_{\rm HMF}^{0.2,1} & \mathcal{C}_{\rm HMF}^{0.52,1} & \mathcal{C}_{\rm HMF}^{1,1} & \mathcal{C}_{\rm HMF}^{2,1} \\ \mathcal{C}_{\rm HMF}^{0.2,2} & \mathcal{C}_{\rm HMF}^{0.52,2} & \mathcal{C}_{\rm HMF}^{1,2} & \mathcal{C}_{\rm HMF}^{2,2} \end{pmatrix}, \tag{A.5}$$

where all these sub-matrices are computed with the Jackknife resampling of Eq. (A.2).

The assumption about the independence of the different data-sets across the redshifts can be tested by checking if the off-diagonal matrices $\mathcal{C}_{V,H}$ and $\mathcal{C}_{H,V}$ are zero matrices (i.e. with null entries). For a graphic reason, we choose to represent the Pearson correlation coefficients, denoted as ρ_{ij}, computed from the covariance matrix elements $\mathcal{C}_{ij}$ as

$$\rho_{ij} = \frac{\mathcal{C}_{ij}}{\sqrt{\mathcal{C}_{ii} \times \mathcal{C}_{jj}}}, \tag{A.6}$$

and we refer to the matrix composed by the ρ_{ij} elements as the *correlation* matrix. These coefficients can only assume values in the $(-1, +1)$ range, and measure the linear correlation between two different variables, in our case mass bins and radius bins.

We present the internal correlation matrices for the cluster counts ($\mathcal{C}_{H,H}$), computed with the two mass cuts considered in this Thesis work ($M_{500c} \geq 5 \times 10^{13} \ M_\odot h^{-1}$ and $M_{500c} \geq 10^{14} \ M_\odot h^{-1}$) in Figure A.1, while the internal correlation matrix for the void counts is shown in Figure A.2. For both cases we consider $N_{\mathrm{sub}} = 125$ sub-samples.

As one can clearly see, the internal correlation matrix of galaxy cluster counts, for both the mass cuts, present non-vanishing off-diagonal elements. This tells us that the assumption of total independence of redshift for these number counts could be incorrect and groundless. These features, however, are probably due to the fact that we are considering the same sub-samples for the different redshifts so that we are considering always the same objects that evolve with time. In order to check this hypothesis, we computed again the covariance matrix for cluster number counts by considering, for each redshift, a different cosmological sub-box. In particular, we divided the native cosmological box in 4 blocks. The results are shown in Figure A.3, and would seem to validate our hypothesis. Though this procedure permits us to check for non-spurious covariances, it highly reduces the statistic of our galaxy cluster sample. The fact that we do not see evident correlation features in the void count correlation matrix could be due to the fact that the data are noisier and, maybe, that the procedure of cleaning of the void catalogues reduces the possibility of following the evolution of the same void with cosmic time.

Finally, we present the total cross-correlation matrix, for both the mass selections considered, in Figure A.4 and A.5. The left upper blocks represent the internal correlation matrix for cluster counts, while the right lower blocks represent the one associated with void counts. The other two off-diagonal blocks are the ones that can be used to check for eventual correlation between the two cosmological probes at different redshifts. As one can see the level of correlation of these two matrices is lower than the diagonal matrices, therefore the assumption of independence between the two cosmological probes seems to hold. About the absence of correlation features in the void count correlation matrix, one can see that it presents, nevertheless, a slightly higher correlation factor with respect to the off-diagonal matrices.
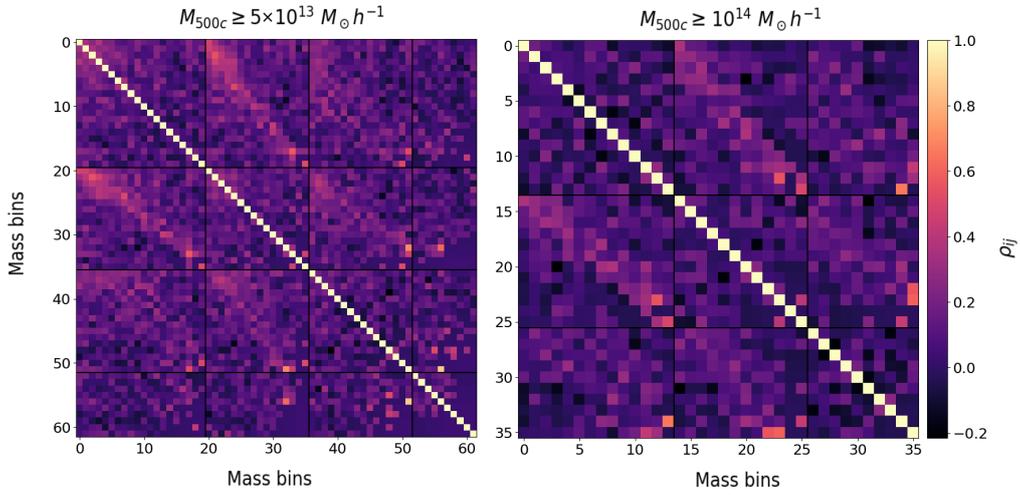
Figure A.1: Jackknife correlation matrices for the cluster counts at different redshifts, considering the two mass cuts exploited in this Thesis work. The solid black lines separate data-sets at different redshifts.
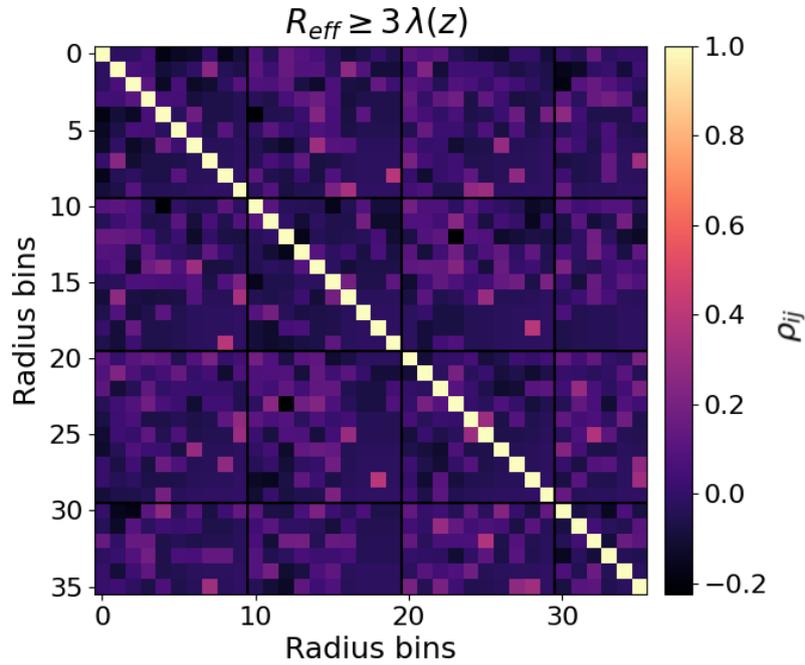

Figure A.2: Jackknife correlation matrix for void counts at different redshifts. The solid black lines separate data-sets at different redshifts.
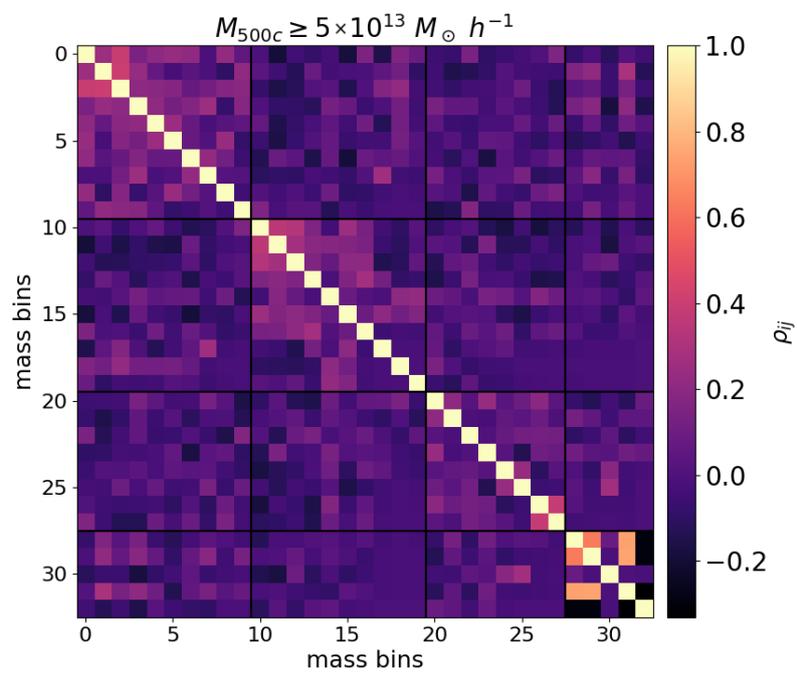
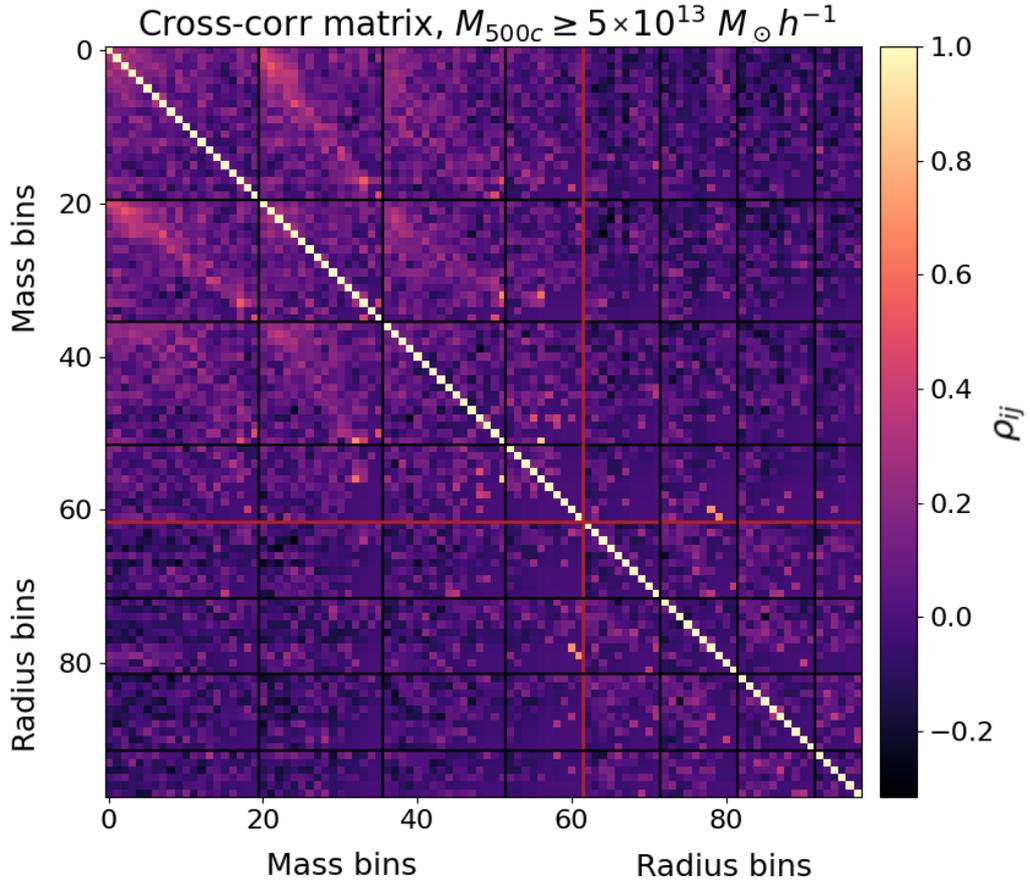Figure A.3: Jackknife Correlation matrix for cluster counts considering different sub-boxes for each redshift.

Figure A.4: Jackknife cross-correlation matrix for cluster counts and void counts, considering clusters having mass $M_{500c} \geq 5 \times 10^{13} \, M_\odot h^{-1}$. data-sets at different redshifts are separated by black solid lines, while mass bins and radius bins are divided by solid red lines.
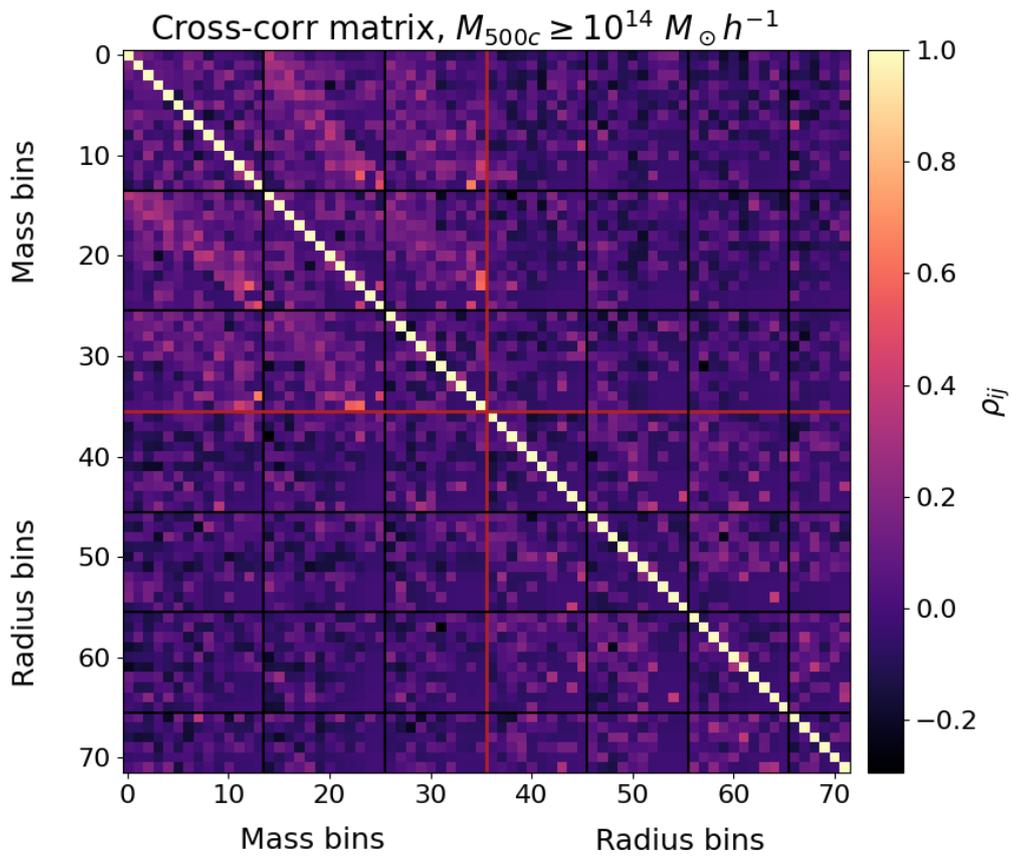
Figure A.5: As Fig. A.4 but for $M_{500c} \geq 10^{14} \ M_\odot h^{-1}$.

Bibliography

Aarseth S. J., 1963, Mon. Not. R. Astron. Soc., 126, 223

Achitouv I. E., Corasaniti P. S., 2012, J. Cosm. Astro-Particle Phys., 2012, 002

Alam S., et al., 2017, Mon. Not. R. Astron. Soc., 470, 2617

Alam S., et al., 2021, Phys. Rev. D, 103, 083533

Alcock C., et al., 1997, Astrophys. J., 486, 697

Amendola L., et al., 2018, Living Reviews in Relativity, 21, 2

Anderson L., et al., 2014, Mon. Not. R. Astron. Soc., 441, 24

Aubert M., et al., 2020, arXiv e-prints, p. arXiv:2007.09013

Baldi M., 2012, Mon. Not. R. Astron. Soc., 422, 1028

Baldi M., Villaescusa-Navarro F., 2016, arXiv e-prints, p. arXiv:1608.08057

Baldi M., Pettorino V., Robbers G., Springel V., 2010, Mon. Not. R. Astron. Soc., 403, 1684

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, Astrophys. J., 304, 15

Barnes J., Hut P., 1986, Nature, 324, 446

Barreira A., Cautun M., Li B., Baugh C. M., Pascoli S., 2015, J. Cosm. Astro-Particle Phys., 2015, 028

Benevento G., Hu W., Raveri M., 2020, Phys. Rev. D, 101, 103517

Bernardeau F., 1994, Astrophys. J., 427, 51

Besuner R., et al., 2021, arXiv e-prints, p. arXiv:2101.11794

Blake C., Glazebrook K., 2003, Astrophys. J., 594, 665

Blanton M. R., et al., 2017a, Astron. J., 154, 28

Blanton M. R., et al., 2017b, Astron. J., 154, 28

Blumenthal G. R., da Costa L. N., Goldwirth D. S., Lecar M., Piran T., 1992, Astrophys. J., 388, 234

Bocquet S., et al., 2019, Astrophys. J., 878, 55

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, Astrophys. J., 379, 440

Bos E. G. P., van de Weygaert R., Dolag K., Pettorino V., 2012, Mon. Not. R. Astron. Soc., 426, 440

Bosma A., 2000, in Combes F., Mamon G. A., Charmandaris V., eds, Astronomical Society of the Pacific Conference Series Vol. 197, Dynamics of Galaxies: from the Early Universe to the Present. p. 91

Brandbyge J., Hannestad S., Haugbølle T., Wong Y. Y. Y., 2010, J. Cosm. Astro-Particle Phys., 2010, 014

Brandt T., 2017, in APS April Meeting Abstracts. p. U5.003

Bridle S. L., Eke V. R., Lahav O., Lasenby A. N., Hobson M. P., Cole S., Frenk C. S., Henry J. P., 1999, Mon. Not. R. Astron. Soc., 310, 565

Chan K. C., Hamaus N., Biagetti M., 2019, Phys. Rev. D, 99, 121304

Chevallier M., Polarski D., 2001, International Journal of Modern Physics D, 10, 213

Colberg J. M., et al., 2008, Mon. Not. R. Astron. Soc., 387, 933

Coles P., 2003, Statistical Properties of Cosmological Fluctuations. p. 237

Contarini S., Ronconi T., Marulli F., Moscardini L., Veropalumbo A., Baldi M., 2019, Mon. Not. R. Astron. Soc., 488, 3526

Contarini S., Marulli F., Moscardini L., Veropalumbo A., Giocoli C., Baldi M., 2020, arXiv e-prints, p. arXiv:2009.03309

Costanzi M., et al., 2019, Mon. Not. R. Astron. Soc., 488, 4779

Costanzi M., et al., 2021, Phys. Rev. D, 103, 043522

Courtin J., Rasera Y., Alimi J. M., Corasaniti P. S., Boucher V., Füzfa A., 2011, Mon. Not. R. Astron. Soc., 410, 1911

Crocce M., Fosalba P., Castander F. J., Gaztañaga E., 2010, Mon. Not. R. Astron. Soc., 403, 1353

Croft R. A. C., Weinberg D. H., Bolte M., Burles S., Hernquist L., Katz N., Kirkman D., Tytler D., 2002, Astrophys. J., 581, 20

Croton D. J., et al., 2005, Mon. Not. R. Astron. Soc., 356, 1155

DES Collaboration 2019, Physical Review Letters, 122, 171301

Dark Energy Survey Collaboration 2018, Phys. Rev. D, 98, 043526

Dark Energy Survey Collaboration et al., 2016, Mon. Not. R. Astron. Soc., 460, 1270

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, Astrophys. J., 292, 371

Despali G., Giocoli C., Angulo R. E., Tormen G., Sheth R. K., Baso G., Moscardini L., 2016, Mon. Not. R. Astron. Soc., 456, 2486

Dhawan S., Jha S. W., Leibundgut B., 2018, Astron. Astrophys., 609, A72

Di Matteo T., Colberg J., Springel V., Hernquist L., Sijacki D., 2008, Astrophys. J., 676, 33

Di Valentino E., Melchiorri A., Silk J., 2016, Phys. Rev. D, 93, 023513

Di Valentino E., et al., 2020a, arXiv e-prints, p. arXiv:2008.11283

Di Valentino E., et al., 2020b, arXiv e-prints, p. arXiv:2008.11284

Di Valentino E., et al., 2020c, arXiv e-prints, p. arXiv:2008.11285

Di Valentino E., et al., 2020d, arXiv e-prints, p. arXiv:2008.11286

Dolag K., Borgani S., Murante G., Springel V., 2009, Mon. Not. R. Astron. Soc., 399, 497

Dolag K., Gaensler B. M., Beck A. M., Beck M. C., 2015, Mon. Not. R. Astron. Soc., 451, 4277

Doroshkevich A. G., 1970, Astrofizika, 6, 581

Efron B., 1979, The Annals of Statistics, 7, 1

Efstathiou G., Davis M., White S. D. M., Frenk C. S., 1985, Astrophys. J. Suppl., 57, 241

Einasto J., Joeveer M., Saar E., 1980, Mon. Not. R. Astron. Soc., 193, 353

Einasto J., et al., 2011, Astron. Astrophys., 534, A128

Einstein A., 1917

Eisenstein D. J., et al., 2011, Astron. J., 142, 72

Elyiv A. A., Karachentsev I. D., Karachentseva V. E., Melnyk O. V., Makarov D. I., 2013, Astrophysical Bulletin, 68, 1

Elyiv A., Marulli F., Pollina G., Baldi M., Branchini E., Cimatti A., Moscardini L., 2015, Mon. Not. R. Astron. Soc., 448, 642

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, Publ. Astron. Soc. Pac., 125, 306

Forero-Romero J. E., Hoffman Y., Gottlöber S., Klypin A., Yepes G., 2009, Mon. Not. R. Astron. Soc., 396, 1815

Freedman W. L., et al., 2001, Astrophys. J., 553, 47

Freedman W. L., et al., 2019, Astrophys. J., 882, 34

Frenk C. S., White S. D. M., Davis M., Efstathiou G., 1988, Astrophys. J., 327, 507

Friedman A., 1922, Zeitschrift fur Physik, 10, 377

Fuchs B., 2001, in Klapdor-Kleingrothaus H. V., ed., Dark Matter in Astro- and Particle Physics. p. 25 (arXiv:astro-ph/0010358)

Gawiser E., Silk J., 1998, Science, 280, 1405

Goodman J., Weare J., 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Granett B. R., Neyrinck M. C., Szapudi I., 2008, arXiv e-prints, p. arXiv:0805.2974

Gunn J. E., Gott J. Richard I., 1972, Astrophys. J., 176, 1

158

Guth A. H., 1981, Phys. Rev. D, 23, 347

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

Hamaus N., Pisani A., Choi J.-A., Lavaux G., Wandelt B. D., Weller J., 2020, J. Cosm. Astro-Particle Phys., 2020, 023

Hang Q., Alam S., Peacock J. A., Cai Y.-C., 2021, Mon. Not. R. Astron. Soc., 501, 1481

Hastings W., 1970, Biometrika, 57, 97

Hockney R. W., Eastwood J. W., 1981, Computer Simulation Using Particles

Holanda R. F. L., Gonçalves R. S., Alcaniz J. S., 2012, J. Cosm. Astro-Particle Phys., 2012, 022

Hoyle F., Vogeley M. S., Pan D., 2012, Mon. Not. R. Astron. Soc., 426, 3041

Hu W., Kravtsov A. V., 2003, Astrophys. J., 584, 702

Hubble E., 1925, Contributions from the Mount Wilson Observatory / Carnegie Institution of Washington, 304, 1

Hubble E., 1926, Contributions from the Mount Wilson Observatory / Carnegie Institution of Washington, 324, 1

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Ivezić Ž., et al., 2019, Astrophys. J., 873, 111

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, Mon. Not. R. Astron. Soc., 321, 372

Jennings E., Li Y., Hu W., 2013, Mon. Not. R. Astron. Soc., 434, 2167

Joudaki S., et al., 2017, Mon. Not. R. Astron. Soc., 471, 1259

Kaiser N., 1984, Astrophys. J. Lett., 284, L9

Kaiser N., 1992, Astrophys. J., 388, 272

Karwal T., Kamionkowski M., 2016, Phys. Rev. D, 94, 103523

Kilbinger M., 2015, Reports on Progress in Physics, 78, 086901

159

Kitayama T., Suto Y., 1996, Astrophys. J., 469, 480

Knebe A., et al., 2013, Mon. Not. R. Astron. Soc., 435, 1618

Komatsu E., et al., 2009, Astrophys. J. Suppl., 180, 330

Komatsu E., et al., 2011, Astrophys. J. Suppl., 192, 18

Krause E., et al., 2017, arXiv e-prints, p. arXiv:1706.09359

Kreckel K., Platen E., Aragón-Calvo M. A., van Gorkom J. H., van de Weygaert R., van der Hulst J. M., Beygu B., 2012, Astron. J., 144, 16

Kreisch C. D., Pisani A., Carbone C., Liu J., Hawken A. J., Massara E., Spergel D. N., Wandelt B. D., 2019, Mon. Not. R. Astron. Soc., 488, 4413

Kuijken K., et al., 2015, Mon. Not. R. Astron. Soc., 454, 3500

LSST Dark Energy Science Collaboration 2012, arXiv e-prints, p. arXiv:1211.0310

Lacey C., Cole S., 1993, Mon. Not. R. Astron. Soc., 262, 627

Lahav O., 2001, in von Hippel T., Simpson C., Manset N., eds, Astronomical Society of the Pacific Conference Series Vol. 245, Astrophysical Ages and Times Scales. p. 617 (arXiv:astro-ph/0105352)

Landy S. D., Szalay A. S., 1993, Astrophys. J., 412, 64

Laureijs R., et al., 2011, arXiv e-prints, p. arXiv:1110.3193

Lavaux G., Wandelt B. D., 2010, Mon. Not. R. Astron. Soc., 403, 1392

Leavitt H. S., Pickering E. C., 1912, Harvard College Observatory Circular, 173, 1

Lesci G. F., et al., 2020, arXiv e-prints, p. arXiv:2012.12273

Lewis A., Challinor A., 2006, Phys. Rept., 429, 1

Liddle A. R., 1999, in Masiero A., Senjanovic G., Smirnov A., eds, High Energy Physics and Cosmology, 1998 Summer School. p. 260 (arXiv:astro-ph/9901124)

Lilje P. B., Lahav O., 1991, Astrophys. J., 374, 29

Linde A. D., 1983, Physics Letters B, 129, 177

Linder E. V., 2003, Physical Review Letters, 90, 091301

Marulli F., Carbone C., Viel M., Moscardini L., Cimatti A., 2011, Mon. Not. R. Astron. Soc., 418, 346

Marulli F., Veropalumbo A., Moresco M., 2016, Astronomy and Computing, 14, 35

Marulli F., Veropalumbo A., Moscardini L., Cimatti A., Dolag K., 2017, Astron. Astrophys., 599, A106

Marulli F., et al., 2018, Astron. Astrophys., 620, A1

Marulli F., Veropalumbo A., García-Farieta J. E., Moresco M., Moscardini L., Cimatti A., 2020, arXiv e-prints, p. arXiv:2010.11206

Massara E., Villaescusa-Navarro F., Viel M., Sutter P. M., 2015, J. Cosm. Astro-Particle Phys., 2015, 018

Mather J. C., et al., 1994, Astrophys. J., 420, 439

Micheletti D., et al., 2014, Astron. Astrophys., 570, A106

Mo H. J., White S. D. M., 1996, Mon. Not. R. Astron. Soc., 282, 347

Mukhanov V. F., Feldman H. A., Brandenberger R. H., 1992, Phys. Rept., 215, 203

Mészáros P., 1974, Astron. Astrophys., 37, 225

Neyrinck M. C., 2008, Mon. Not. R. Astron. Soc., 386, 2101

Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, Mon. Not. R. Astron. Soc., 396, 19

Padmanabhan N., et al., 2007, Mon. Not. R. Astron. Soc., 378, 852

Peacock J. A., et al., 2001, Nature, 410, 169

Peebles P. J. E., 1970, Astron. J., 75, 13

Peebles P. J. E., 1980, The large-scale structure of the universe

Peebles P. J. E., 2001, Astrophys. J., 557, 495

Perlmutter S., et al., 1999, Astrophys. J., 517, 565

Pierpaoli E., Borgani S., Scott D., White M., 2003, Mon. Not. R. Astron. Soc., 342, 163

Pisani A., Sutter P. M., Hamaus N., Alizadeh E., Biswas R., Wandelt B. D., Hirata C. M., 2015, Phys. Rev. D, 92, 083531

Pisani A., et al., 2019, British Association for American Studies, 51, 40

Planck Collaboration et al., 2014, Astron. Astrophys., 571, A16

Planck Collaboration et al., 2016a, Astron. Astrophys., 594, A1

Planck Collaboration et al., 2016b, Astron. Astrophys., 594, A14

Planck Collaboration et al., 2016c, Astron. Astrophys., 594, A24

Planck Collaboration et al., 2020, Astron. Astrophys., 641, A6

Platen E., van de Weygaert R., Jones B. J. T., 2007, Mon. Not. R. Astron. Soc., 380, 551

Poulin V., Smith T. L., Karwal T., Kamionkowski M., 2019, Physical Review Letters, 122, 221301

Press W. H., Schechter P., 1974, Astrophys. J., 187, 425

Refregier A., Rhodes J., Groth E. J., 2002, Astrophys. J. Lett., 572, L131

Riess A. G., et al., 1998, Astron. J., 116, 1009

Riess A. G., Casertano S., Yuan W., Macri L. M., Scolnic D., 2019, Astrophys. J., 876, 85

Ronconi T., Marulli F., 2017, Astron. Astrophys., 607, A24

Ronconi T., Contarini S., Marulli F., Baldi M., Moscardini L., 2019, Mon. Not. R. Astron. Soc., 488, 5075

Ross A. J., Samushia L., Howlett C., Percival W. J., Burden A., Manera M., 2015, Monthly Notices of the Royal Astronomical Society, 449, 835

Ryden B., 2003, Introduction to cosmology

Schneider P., 2005, arXiv e-prints, pp astro–ph/0509252

Schuster N., Hamaus N., Pisani A., Carbone C., Kreisch C. D., Pollina G., Weller J., 2019, J. Cosm. Astro-Particle Phys., 2019, 055

Seo H.-J., Eisenstein D. J., 2003, Astrophys. J., 598, 720

Shandarin S. F., 2011, J. Cosm. Astro-Particle Phys., 2011, 015

Shandarin S. F., Zeldovich Y. B., 1989, Rev. Mod. Phys., 61, 185

Shandarin S., Feldman H. A., Heitmann K., Habib S., 2006, Mon. Not. R. Astron. Soc., 367, 1629

Sheth R. K., Tormen G., 1999, Mon. Not. R. Astron. Soc., 308, 119

Sheth R. K., Tormen G., 2002, Mon. Not. R. Astron. Soc., 329, 61

Sheth R. K., van de Weygaert R., 2004, Mon. Not. R. Astron. Soc., 350, 517

Sheth R. K., Mo H. J., Tormen G., 2001, Mon. Not. R. Astron. Soc., 323, 1

Sievers J. L., et al., 2003, Astrophys. J., 591, 599

Spolyar D., Sahlén M., Silk J., 2013, Phys. Rev. Lett., 111, 241103

Springel V., 2005, Mon. Not. R. Astron. Soc., 364, 1105

Springel V., Hernquist L., 2003, Mon. Not. R. Astron. Soc., 339, 289

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, Mon. Not. R. Astron. Soc., 328, 726

Springel V., et al., 2005, Nature, 435, 629

Starobinskiǐ A. A., 1979, Soviet Journal of Experimental and Theoretical Physics Letters, 30, 682

Sunyaev R. A., Zeldovich Y. B., 1972, Comments on Astrophysics and Space Physics, 4, 173

Suto Y., Sato K., Sato H., 1984, Progress of Theoretical Physics, 71, 938

Sutter P. M., et al., 2015, Astronomy and Computing, 9, 1

Suzuki N., et al., 2012, Astrophys. J., 746, 85

Taylor A., Joachimi B., 2014, Mon. Not. R. Astron. Soc., 442, 2728

Taylor A., Joachimi B., Kitching T., 2013, Mon. Not. R. Astron. Soc., 432, 1928

Tegmark M., et al., 2004, Astrophys. J., 606, 702

Tikhonov A. V., Karachentsev I. D., 2006, Astrophys. J., 653, 969

Tikhonov A. V., Klypin A., 2009, Mon. Not. R. Astron. Soc., 395, 1915

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, Astrophys. J., 688, 709

Tornatore L., Borgani S., 2007, in Emsellem E., Wozniak H., Massacrier G., Gonzalez J. F., Devriendt J., Champavert N., eds, EAS Publications Series Vol. 24, EAS Publications Series. pp 139–144, doi:10.1051/eas:2007021

Tornatore L., Borgani S., Springel V., Matteucci F., Menci N., Murante G., 2003, Mon. Not. R. Astron. Soc., 342, 1025

Uzan J.-P., Aghanim N., Mellier Y., 2004, Phys. Rev. D, 70, 083533

Vázquez J. A., Padilla L. E., Matos T., 2018, arXiv e-prints, p. arXiv:1810.09934

Verde L., Treu T., Riess A. G., 2019, Nature Astronomy, 3, 891

Verza G., Pisani A., Carbone C., Hamaus N., Guzzo L., 2019, J. Cosm. Astro-Particle Phys., 2019, 040

Viel M., Haehnelt M. G., Springel V., 2004, Mon. Not. R. Astron. Soc., 354, 684

Vikhlinin A., et al., 2009, Astrophys. J., 692, 1060

Villaescusa-Navarro F., Vogelsberger M., Viel M., Loeb A., 2013, Mon. Not. R. Astron. Soc., 431, 3670

Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, Mon. Not. R. Astron. Soc., 433, 1230

Webster M., Bridle S. L., Hobson M. P., Lasenby A. N., Lahav O., Rocha G., 1998, arXiv e-prints, pp astro–ph/9802109

Wechsler R. H., Tinker J. L., 2018, Annu. Rev. Astron. Astrophys., 56, 435

White S. D. M., Zaritsky D., 1992, Astrophys. J., 394, 1

Yang L. F., Neyrinck M. C., Aragón-Calvo M. A., Falck B., Silk J., 2015, Mon. Not. R. Astron. Soc., 451, 3606

Yoo J., Watanabe Y., 2012, International Journal of Modern Physics D, 21, 1230002

York D. G., et al., 2000, Astron. J., 120, 1579

Zel'Dovich Y. B., 1970, Astron. Astrophys., 500, 13

Zeldovich Y. B., 1972, Mon. Not. R. Astron. Soc., 160, 1P

Zeldovich I. B., Einasto J., Shandarin S. F., 1982, Nature, 300, 407

Zwicky F., 1937, Astrophys. J., 86, 217

de Sitter W., 1917