Alma Mater Studiorum · Università di Bologna

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

Back-in-time void finder: dynamical detection of cosmic voids for precision Cosmology

Tesi di Laurea Magistrale

Presentata da:
Simone Sartori

Relatore:
Prof. Federico Marulli

Co-Relatori:
Prof. Lauro Moscardini
Dr.ssa Sofia Contarini

Anno Accademico 2021/2022

Abstract

Cosmic voids are vast and underdense regions emerging between the elements of the cosmic web (i.e. galaxy clusters, filaments and walls) and dominating the large-scale structure of the Universe in terms of volume. They evolve from the perturbations of the primordial matter density field towards a mildly nonlinear regime, gradually becoming less dense and more spherical (Icke, 1984). Analogously to their overdense counterpart (i.e. galaxy clusters), void number counts and density profiles have been demonstrated to provide powerful cosmological probes (see Pisani et al., 2019, and references therein). Indeed, thanks to their low-density nature and they very large sizes (up to hundreds of megaparsec), voids represent natural laboratories to test alternative dark energy scenarios, modifications of gravity and the presence of massive neutrinos (see Contarini et al., 2021, and references therein).

Despite the increasing use of cosmic voids in Cosmology, a commonly accepted definition for these objects has not yet been reached. For this reason, different void finding algorithms have been proposed during the years. The latter can be divided in three main classes, according to the criterion on which they are based: *density*, *geometry* and *dynamics* (see Colberg et al., 2008, for a cross-comparison). Voids finder algorithms based on density or geometrical criteria are affected by the shot noise, i.e. the intrinsic uncertainties arising from the sparsity of the objects used as tracers of the density field. In recent years, new solutions have been explored to face these issues. Among these the most interesting is based on the innovative idea to identify void positions through the dynamics of the mass tracers, without performing any direct reconstruction of the density field.

The goal of this Thesis is to provide a performing void finder algorithm based on dynamical criteria. In particular, we have improved and extend the work introduced by Elyiv et al. (2015), adding the new codes to the `CosmoBolognaLib` (Marulli et al., 2016), a large set of *free software* C++/Python numerical libraries providing an efficient environment for a large variety of cosmological calculations.

The `Back-in-time void finder` (`BitVF`) we present exploits a Lagrangian representation of the mass tracers, interpreting their motion as a cosmological fluid. Tracers are used as test particles and their orbits are reconstructed from their actual clustered configuration to an homogeneous and isotropic distribution, expected for the Universe early epoch. The displacement field reconstruction is based on the Zel'dovich Approximation (Zel'dovich, 1970), which admits the extension of the linear theory also to perturbations having density contrast bigger than one. With this method, we exploit the principle of least action to trace the back-in-time evolution of tracers, approximating their trajectories as rectilinear. Once the displacement field, $\mathbf{\Psi}$, is reconstructed, the divergence field is computed as $\mathbf{\Theta} = \nabla \cdot \mathbf{\Psi}$, which represents the density at each point. Consequently, void centres are identified as

local minima with negative divergence and will correspond to the points from which mass tracers locally depart with maximum velocity. Then a radius is assigned to each void as the radius of a sphere having a certain fraction of the Universe mean density.

In order to extract cosmological information, it is necessary to be able to compare the sample of detected voids with theoretical predictions. This result is achieved through a procedure of void sample cleaning, based on three basic steps (Jennings et al., 2013), in which voids overlapping or having internal densities over a certain threshold are rejected, moreover voids radii are rescaled to a fixed internal spherical density contrast. An improved and more stable version of the cleaning algorithm, based on the work of Ronconi & Marulli (2017), is presented in this Thesis. The cleaned catalogue emerging from this procedure is then ready to be used for cosmological analyses, and is especially suited to exploit the void abundance as cosmological probe.

In this Thesis work we applied the developed void finding algorithm to the DUSTGRAIN-pathfinder simulations, using different types of mass tracers (dark matter particles and dark matter haloes), at various redshifts ($z = 0,\ 0.5,\ 1,\ 2$) and mass tracer subsample values. From the resulting void samples we computed different void statistics, comparing the results to those obtained with `VIDE` (Sutter et al., 2014a), currently the most widely used void finder in literature. In particular, void density profiles and counts have been extensively investigated. `BitVF` proved to be able to produce a more reliable void sample, since the statistics that we derived using this code are more accurate than those obtained with `VIDE` voids. This is true in particular when we apply these algorithms to biased tracer distributions. We demonstrated indeed that only `BitVF` provides void catalogues whose statistics can be used to accurately predict the values of σ_8 and $\Omega_{\rm m}$, with confidence contours remarkably tight. The `BitVF` algorithm is thus a fundamental tool for precision Cosmology, particularly in the perspective of future wide-field galaxy surveys like the ESA *Euclid* mission (Laureijs et al., 2011), the NASA *Nancy Grace Roman Space Telescope* (NGRST, formerly called WFIRST, Green et al. 2012) and the *Vera C. Rubin Observatory* (LSST, Collaboration 2012).

Sommario

I vuoti cosmici sono regioni vaste e sottodense che emergono tra gli elementi della cosiddetta ragnatela cosmica (ovvero gli ammassi di galassie, i filamenti e le distribuzioni planari galassie) dominando la struttura a grande scala dell'Universo in termini di volume. I vuoti si evolvono dalle perturbazioni del campo di densità primordiale della materia verso un regime leggermente non lineare, diventando gradualmente meno densi e più sferici (Icke, 1984). Analogamente alla loro controparte sovradensa (cioè gli ammassi di galassie), il numero di vuoti e i profili di densità forniscono potenti *probes* cosmologiche (vedi Pisani et al. 2019 e riferimenti). Infatti, grazie alla loro natura sottodensa e alle loro dimensioni (fino a centinaia di megaparsec), i vuoti rappresentano laboratori naturali per testare scenari alternativi di energia oscura, gravità modificata e neutrini massivi (vedi Contarini et al. 2021, e riferimenti).

Nonostante il crescente utilizzo dei vuoti cosmici in Cosmologia, non si è ancora giunti a una definizione comunemente accettata di questi oggetti. Per questo motivo, nel corso degli anni sono stati proposti diversi algoritmi per la ricerca dei vuoti. Questi ultimi possono essere suddivisi in tre classi principali, a seconda del criterio su cui si basano: densità, geometria e dinamica (si veda Colberg et al. 2008 per un confronto diretto tra questi metodi). Gli algoritmi di identificazione dei vuoti basati su criteri di densità o geometrici risentono del cosidetto *shot noise*, ovvero delle incertezze intrinseche derivanti dalla dispersione degli oggetti utilizzati come traccianti del campo di densità. Per questo negli ultimi anni sono state studiate nuove soluzioni per far fronte a questi problemi intrinsici. Tra questi il più interessante è basato sull'idea innovativa di identificare le posizioni dei vuoti attraverso la dinamica dei traccianti, senza effettuare alcuna ricostruzione diretta del campo di densità.

L'obiettivo di questa Tesi è quello di presentare un algoritmo di ricerca dei vuoti basato su criteri dinamici, funzionante ed efficiente. In particolare, abbiamo migliorato ed esteso il lavoro introdotto da Elyiv et al. (2015), aggiungendo i nuovi codici alle `CosmoBolognaLib` (Marulli et al., 2016), un ampio insieme di librerie C++/Python *free software* che forniscono un contesto all'avanguardia per una grande varietà di calcoli cosmologici.

Il void finder che andiamo a presentare, chiamato `Back-in-time void finder` (`BitVF`), sfrutta una rappresentazione lagrangiana dei traccianti di massa, interpretando il loro moto come quello di un fluido cosmologico. I traccianti sono utilizzati quindi come particelle di prova e le loro orbite sono ricostruite dalla loro attuale configurazione clusterata a una distribuzione omogenea e isotropa, cioè quella prevista per l'epoca iniziale dell'Universo. La ricostruzione del campo di spostamento si basa sull'approssimazione di Zel'dovich (Zel'dovich, 1970), che ammette l'estensione della

teoria lineare anche a perturbazioni con contrasto di densità maggiore di uno. Con questo metodo, sfruttiamo il principio di minima azione per tracciare l'evoluzione a ritroso nel tempo dei traccianti, approssimando le loro traiettorie come rettilinee. Una volta ricostruito il campo di spostamento, $\mathbf{\Psi}$, si calcola il campo di divergenza come $\mathbf{\Theta} = \nabla \cdot \mathbf{\Psi}$, rappresentante la densità puntuale. Di conseguenza, i centri dei vuoti sono identificati come minimi locali con divergenza negativa e corrispondono ai punti da cui i traccianti fluiscono con la massima velocità, quando considerato il loro spostamento indietro nel tempo. A questo punto si assegna una dimensione a ogni vuoto interpretandoli come sfere contenenti una certa frazione della densità media dell'Universo.

Per estrarre informazioni cosmologiche è necessario poter confrontare il campione di vuoti individuati con le predizioni teoriche. Questo risultato si ottiene attraverso una procedura chiamata *cleaning*, basata su tre passi fondamentali (Jennings et al., 2013), nei quali vengono eliminati i vuoti sovrapposti tra loro e quelli con densità superiori a una certa soglia, e vengono inoltre riscalati i vuoti modellandoli come sfere contenenti uno specifico contrasto di densità. In questa Tesi viene presentata una versione migliorata e più stabile dell'algoritmo di *cleaning*, basata sul lavoro di (Ronconi & Marulli, 2017). Questa procedure consente di costruire un catalogo di vuoti pronto per essere utilizzato per analisi cosmologiche, particolarmente adatto per utilizzare l'abbondanza di vuoti come *probe* cosmologica.

In questi lavoro di Tesi, applichiamo l'algoritmo di identificazione di vuoti che abbiamo sviluppato alla simulazione DUSTGRAIN-pathfinder, considerando diversi tipi di traccianti (particelle di materia oscura e aloni di materia oscura), a vari redshift ($z = 0, 0.5, 1, 2$) e diversi valori di sottocampionamento dei traccianti. Le statistiche dei vuoti risultanti sono state confrontate con quelle ottenute tramite l'applicazione di `VIDE` (Sutter et al., 2014a), l'algoritmo di ricerca di vuoti più diffuso in letteratura. In particolare, sono stati confrontati i profili di densità, i conteggi e la capacità di vincolare i parametri cosmologici dei vuoti identificati con questi due algoritmi. `BitVF` si è dimostrato in grado di produrre un campione di vuoti più affidabile, in quanto da esso abbiamo derivato statistiche più accurate rispetto a quelle ottenute con `VIDE`, specie quando applicato a distribuzioni traccianti con bias, riuscendo a prevedere correttamente i valori di σ_8 e Ω_m, con contorni di confidenza notevolmente stringenti. L'algoritmo `BitVF` è destinato a diventare uno strumento fondamentale per la Cosmologia di precisione, in particolare in vista delle future *surveys* di galassie ad ampio campo, come la *missione Euclid* di ESA Laureijs et al. 2011, il *telescopio spaziale Nancy Grace Roman* della NASA (NGRST, precedentemente chiamato WFIRST, Green et al. 2012) e *l'osservatorio Vera C. Rubin* (LSST, LSST Collaboration 2012).

Contents

Introduction

The last hundred years have revolutionised the human view of the Universe. If, Hubble and Lemaître already in 1929 and 1931, respectively, contributed with their work to definitively rejecting the vision of a static and immutable Universe (Edwin 1929, Lemaître 1931), in the last twenty-five years the observation of distant type Ia supernovae has definitively established the accelerated expanding Universe as the prevailing model (Riess et al. 1998, Schmidt et al. 1998, Perlmutter et al. 1999). In the following years, observations of the anisotropies of the cosmic microwave background and of the large-scale structures of the Universe have confirmed this scenario, now widely accepted among the scientific community (Eisenstein et al. 2005, Planck Collaboration 2016, Planck Collaboration 2020).

This model, named ΛCDM after its main components, is therefore currently considered the standard cosmological model. This paradigm assumes, first of all, the validity of the cosmological principle, i.e. that the Universe is homogeneous and isotropic at large scales, as well as the validity of the Einstein's General Relativity. Moreover, it predicts the presence of a significant amount of non-collisional and non-relativistic matter, about 27% of the total mass, called *cold dark matter* (CDM). Furthermore, this model interprets the accelerated expansion of the Universe as being due to the action of a component, of constant density in space and time, called *cosmological constant* Λ. This component currently represents about 70% of the total energy density of the Universe. The cosmological constant can be also interpreted as the so-called *dark energy* (DE), which could be thought of as the vacuum energy. However, the mechanisms regulating this physical process are unknown to date, as is the exact nature of dark matter (DM). Due to these still open problems, this scenario has often been questioned. In addition, further theoretical and observational problems have emerged over the years. The former, for instance, consist of the problems of coincidence and fine-tuning (Weinberg 1989, but see Bianchi & Rovelli 2010 for an alternative point of view), the latter consist in particular on the problem of the tension in the measurements of the value of Hubble's constant (which parameterises the expansion velocity of the cosmos) when measured in the early or late Universe (Bernal et al. 2016, Moresco & Marulli 2017).

The scientific community has been busy in recent years in investigating alternative models to the ΛCDM, such as, for instance, models involving a time dependency on the density of DE or involving modifications to the theory of General Relativity. In this context, a large number of cosmological probes are being exploited in an attempt of constraining cosmological parameters, such as the cosmic microwave background (Planck Collaboration 2016, Planck Collaboration 2020), the clustering of galaxies, in particular baryonic acoustic oscillations (BAO, Crocce & Scoccimarro 2008), and weak lensing effects due to large-scale structure (Hildebrandt et al., 2016).

However, despite the great variety of cosmological probes, many of the main problems of the cosmological model still remain unsolved. It is therefore necessary to investigate possible new probes that explore the less known regions of the Universe, providing information that is orthogonal to what is currently available, thus breaking the degenerations that arise in the constraints of cosmological parameters. A possible class of objects that match this definition are *cosmic voids*. Cosmic voids are large structures that dominate the volume of the Universe, occupying the underdense regions that partition the cosmic web. These objects constitute a unique cosmological probe: analogously to their overdense counterparts (i.e. galaxy clusters), void number counts and density profiles have been demonstrated to represent powerful cosmological probes (see Pisani et al., 2019, and references therein). Moreover, their interiors, spanning a large range of scales and being characterised by with low matter density, make them particularly suited to study DE and modified gravity, as well as massive neutrinos, primordial non-Gaussianity, and possible Physics beyond the Standard Model. However, the study of voids, to be statistically significant, requires galaxy surveys that sample large volumes of the Universe in sufficient detail to maximise the number of objects observed. For this reason, only in recent years cosmic voids have started to gain popularity as cosmological probes. In the near future, the upcoming sky surveys such as the ESA *Euclid* mission (Laureijs et al., 2011), the NASA *Nancy Grace Roman Space Telescope* (NGRST, formerly called WFIRST, Green et al. 2012) and the *Vera C. Rubin Observatory* (LSST, Collaboration 2012) will provide a large amount of high quality data, making the cosmological study performed with these objects significantly more effective.

The identification of cosmic voids, however, is non-trivial. In fact, numerous difficulties are present: the lack of a formal definition of cosmic voids, the scarcity of tracers present in these underdense zones, the discretisation of the mass (the density field is sampled by the observed luminous tracers), and the shot noise of the tracer distribution. Over time, several identification methods have been proposed by the scientific community. Following the scheme discussed by Lavaux & Wandelt (2009), we can classify these methods into three different classes, according to the exploited detection criteria: *density*, *geometrical* and *dynamical*. The work presented in this Thesis aims to overcome the difficulties associated with void identification through the development of a new dynamic void finder, called `back-in-time void finder` (`BitVF`), resuming and expanding the work of Elyiv et al. (2015). Dynamic methods, unlike the other approaches, use tracers to sample the velocity field, instead of the density field, overcoming the difficulties associated with the small number of available tracers inside voids.

This Thesis work accompanies the reader from the acquisition of the cosmological tools necessary to understand the subject matter, to an overview of existing identification methods, and finally to a description and validation of the new algorithm developed, providing a complete overview of the topic. In particular, the Thesis is organised as follows:

- in **Chapter 1** we provide the fundamentals for the mathematical description of the modern cosmological models. We supply the main elements of the theory of General Relativity, going through the derivation of the Friedmann Equations. Finally, we illustrate the main features of the currently adopted standard cosmological model, the so-called ΛCDM model.

- in **Chapter 2** we review the Jeans theory, which provides the modelling of structure formation, and we present the theoretical description of the linear and nonlinear evolution of the Universe. Lastly, we introduce the reader to the cosmological simulations, fundamental to test the predictions of the proposed cosmological models and overcome the statistical limits imposed by the uniqueness of the Universe.

- in **Chapter 3** we provide a comprehensive overview of cosmic voids, describing the main methods for their identification, and then illustrate the main statistics associated with these objects, their observable features, and their potential as cosmological probes.

- in **Chapter 4** we describe our void finder in detail, analysing step by step the algorithm which, starting from the position of cosmic tracers, leads to the identification of voids. Furthermore, we analyse the application of `BitVF` to simulations to validate the method and the results.

- in **Chapter 5** we compare the properties of voids obtained through the application of our finder on different distributions of cosmological tracers with those obtained through the application of the most commonly used void finder in literature, with the aim of exploring the validity of our work and the possibility of a future use of `BitVF` for finding voids through which to constrain cosmological parameters.

- in **Chapter 6** we provide an overview of our Thesis work and illustrate future prospects and outstanding issues.

Chapter 1

Cosmological framework

In this first chapter, we explore the physical and mathematical background that will enable us to understand the topics discussed in this thesis. The aim is to familiarise with modern cosmological models based on the theory of General Relativity (GR, Sec. 1.1). In order to do that, we introduce in Sec. 1.2 the *Friedmann-Lemaître-Robertson-Walker* (FLRW) metric, which is needed to describe the curvature of space-time in homogeneous and isotropic universes. At this point, we are able to define both the *Hubble-Lemaître* law and the cosmological redshift (Sec. 1.3), and to derive the *Friedmann* equations as solutions of the *Einstein*'s field equation (Sec. 1.4). With these equations we are able to illustrate, in Sec. 1.6 and Sec. 1.7, models for an open, flat and closed universes and to provide an overview of the so-called Standard Cosmological Model.

1.1 Fundamentals of General Relativity

From planetary interactions to large-scale structure formation, the Universe is dominated by Gravity. For this reason, it is natural that the main cosmological models are based on the action of this force. At present, the most consistent description of Gravity is the one provided by Albert Einstein's theory of General Relativity (GR, Einstein 1915). For GR, gravity is actually a deformation effect of space-time given by the presence of mass and energy. The geometry of space-time is described by the *metric tensor*, $g_{\mu\nu}$. The infinitesimal distance ds^2 between two events $x^\mu = (ct, x, y, z)$ and $x^\nu = x^\mu + dx^\mu = (c(t + dt), x + dx, y + dy, z + dz)$ can be expressed as

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \qquad (\mu, \nu = 0, 1, 2, 3) \ , \tag{1.1}$$

where x, y, and z, represent the spatial coordinates, c is the speed of light and t the coordinate time. Equation (1.1) can be expanded as

$$ds^2 = g_{00} dt^2 + 2 g_{0i} dx^i dt + g_{ij} dx^i dx^j \ , \tag{1.2}$$

where $g_{00} dt^2$ is the time component, $g_{ij} dx^i dx^j$ are the spatial components and $2 g_{0i} dx^i dt$ are the mixed components.

We now define the *Riemann–Christoffel Tensor*, $R^\mu_{\alpha\beta\gamma}$, a tensor of type (1,3) which is the most common way to encode the curvature of a Riemannian manifold

$$R^\mu_{\alpha\beta\gamma} = \frac{d\Gamma^\mu_{\alpha\gamma}}{dx^\beta} - \frac{d\Gamma^\mu_{\alpha\beta}}{dx^\gamma} + \Gamma^\mu_{\sigma\beta}\Gamma^\sigma_{\gamma\alpha} - \Gamma^\mu_{\sigma\gamma}\Gamma^\sigma_{\beta\alpha} \ . \tag{1.3}$$

The Riemann-Christoffel Tensor can be contracted in $R_{\mu\nu}$ and R, defined as:

$$R_{\alpha\beta} \equiv R^{\mu}_{\alpha\beta\gamma} \, , \tag{1.4}$$

$$R \equiv R^{\mu}_{\mu} = g^{\mu\nu} R_{\mu\nu} \, , \tag{1.5}$$

respectively the *Ricci Tensor* and *Ricci Scalar*.

The geometry of space-time is related to the presence of mass and energy, described by the *energy-momentum tensor*, $T_{\mu\nu}$, through the *Einstein's field equation*

$$R_{\nu\mu} - \frac{1}{2} g_{\nu\mu} R = \frac{8\pi G}{c^4} T_{\mu\nu} \tag{1.6}$$

where G is the gravitational constant. Looking at the left hand of the equation (1.6), we can define the *Einstein tensor*

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \, , \tag{1.7}$$

which gives us the description of the geometry of space-time. Combining equations (1.6) and (1.7), we obtain a compact equation for the *Einstein's gravitational field equation*

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \, , \tag{1.8}$$

which is one of the pillars of Cosmology. The quantity $8\pi G/c^4$ ensures to obtain the Poisson's equation in the weak gravitational field limit.

1.2 The Friedmann–Lemaître–Robertson–Walker Metric

The most common cosmological models are based on the assumption of *isotropy* (the universe is identical in every direction) and *homogeneity* (the universe is identical everywhere) on large scales[1]. This assumption is called *Cosmological Principle* (CP). Once the validity of this hypothesis is assumed, it is necessary to construct a model for a universe that respects the CP. It is possible to assume that there exists a certain time instant at which the metric is the same at each spatial position. Thanks to the isotropy hypothesis, the terms g_{0i} of Eq. (1.2) become null. The general form of the metric is thus obtained under the CP conditions:

$$ds^2 = (cdt)^2 - g_{ij} dx^i dx^j = (cdt)^2 - dl^2 \, . \tag{1.9}$$

We can imagine the Universe as a continuous fluid in which each element is labelled by three spatial and one temporal coordinates, x^i, with $i = (1, 2, 3)$, and t. The former are called *comoving coordinates* while the latter is called *proper time*. These coordinates are intended to be at rest with respect to the reference system of the fluid of the Universe.

[1]Nowadays, large scales are defined as dimensions of hundreds of Mpc, where 1 Mpc $= 3.086 \times 10^{18}$ cm.

To determine g_{ij} we have to find a spatial 3D metric which follows the requirements of homogeneity and isotropy. Thanks to the assumption of the CP, the tensor R^i_{jkl} does not depend on the derivatives of the metric. Through the symmetry properties of the deriving form of the Riemann tensor and the transformation from Cartesian coordinates (x_1, x_2, x_3) to spherical polar coordinates (ρ, ϕ, θ),

$$\begin{cases} x_1 = \rho \sin(\theta) \cos(\phi) \\ x_2 = \rho \sin(\theta) \sin(\phi) \\ x_3 = \rho \cos(\theta) \end{cases} , \tag{1.10}$$

where the ranges of these values are $0 \leq \rho < \infty$, $0 \leq \phi < 2\pi$, $0 \leq \theta < \pi$ and ρ is considered dimensionless, we can introduce the most general form of equation (1.9), the so-called *Friedmann-Lemaître-Robertson-Walker* (FLRW) metric:

$$ds^2 = (cdt)^2 - a^2(t) \left[\frac{dr^2}{1 - \kappa r^2} + r^2(\sin^2 \theta d\phi^2 + d\theta^2) \right] , \tag{1.11}$$

where $a(t)$ is the *cosmic scale factor* (or the *expansion parameter*), having the dimensions of a length, and κ is the *adimensional curvature parameter*.

The values of κ and $a(t)$ can be derived from Einstein's field equation, once the energy-momentum tensor is given. In particular, calculating the value of κ is essential for determining the geometry of the considered universe. There can be three types of curvature: *positive*, *negative* and *null*. Correspondingly, κ can assume only three values, each associated with a particular type of geometry:

- $\kappa = 1 \rightarrow$ **elliptic geometry**. Space is closed but with no boundaries. It can be thought as a three-dimensional hypersphere or some other spherical 3-manifold (such as the Poincaré dodecahedral space);
- $\kappa = 0 \quad \rightarrow$ **flat geometry**. Space is infinite and Euclidean;
- $\kappa = -1 \rightarrow$ **hyperbolic geometry**. Space is open and infinite, analogously to the 2D surface of a saddle.

1.3 The Hubble–Lemaître law

The *proper distance*, D_{pr}, is defined as the time distance between two points P and P_0 in a generic coordinate system. This quantity depends on time through the already defined term $a(t)$ and can be expressed as

$$D_{pr} = a(t) \int_0^r \frac{dr'}{\sqrt{1 - \kappa r'^2}} = a(t) F_\kappa(r) . \tag{1.12}$$

Henceforth, using the convention of indicating with the subscript 0 the quantities calculated at the present time, $t = t_0$, we can link the distance at a generic time t with that at the present time through the relation

$$D_C = D_{pr}(t_0) = a_0 F_\kappa(r) . \tag{1.13}$$

The quantity D_C, called *comoving distance*, remains constant with the expansion of the universe. The connection between D_{pr} and D_C is provided by the relationship

$$D_{pr} = \frac{a(t)}{a_0} D_C \ . \tag{1.14}$$

We can visualise the expanding Universe as a 3D version of the surface of an inflating balloon: the distance between each pair of points on the surface increases with the flow of time. It is possible to define a *radial velocity* v_r for each pair of points as the derivative of D_{pr} with respect to time t:

$$v_r = \frac{d}{dt} D_{pr} = \frac{d}{dt}[a(t)F(r)] = \dot{a}(t)F(r) + a(t)\dot{F}(r) \ . \tag{1.15}$$

Given the time-independence of the term $F(r)$, the previous relation becomes the well-known *Hubble-Leimaître's law*,

$$v_r = \dot{a}(t)F(r) = \frac{\dot{a}(t)}{a(t)} D_{pr} = H(t) D_{pr} \ . \tag{1.16}$$

$H(t)$ is a function of time and is supposed to have the same value across the Universe at a given cosmic time. Its value at present time, $H_0 = H(t = t_0)$ is called the *Hubble constant* and provides a measure of the isotropic expansion rate of the present Universe. The value of H_0 is generally provided, through the inclusion of a dimensionless parameter h, in the following way:

$$H_0 = 100h \ \mathrm{km\,s^{-1}\,Mpc^{-1}}. \tag{1.17}$$

In the current state of the art, the measured value of H_0 is $\sim 70 \ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, but it is affected by a certain tension ($\sim 4\sigma$) between the value inferred from the cosmic microwave background (CMB) analysis and the one measured in the local Universe (see Verde et al. 2019, Di Valentino et al. 2021 and Abdalla et al. 2022). For example, $H_0 = 67.4 \pm 0.5 \ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, from the CMB angular spectrum (Planck Collaboration, 2020), $H_0 = 67.7^{+4.3}_{-0.42} \ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ from the analysis of gravitational waves (Mukherjee et al., 2020) and $H_0 = 74.03 \pm 1.42 \ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, by using distance ladders as Cepheids or SNIa (Riess et al., 2019). Based on the approximation of $H(t)$ being constant in time and having H_0 expressed in units of $\mathrm{s^{-1}}$, a rough estimate of the age of the Universe can be obtained by inverting the value of the Hubble parameter.

1.3.1 Definition of redshift

The main consequence of having an isotropically expanding Universe is obviously the distancing of faraway objects from the observer, net of their own peculiar motions. We call the global motion of objects in the Universe with respect to each other the *Hubble Flow*. The observational consequence of this phenomenon is the reddening of the electromagnetic radiation due to the shift to longer wavelengths. This phenomenon is called *cosmological redshift* of the electromagnetic spectrum and it is not to be confused with the Doppler effect.

In general, consider a moving source relatively to the observer that emits monochromatic radiation at a wavelength λ_{em} relative to its own reference system. Let λ_{oss}

denote the wavelength of the electromagnetic radiation measured by the observer and z the relative difference between the two wavelengths:

$$z \equiv \frac{\lambda_{oss} - \lambda_{em}}{\lambda_{em}} = \frac{\Delta\lambda}{\lambda_{em}} \ . \tag{1.18}$$

z can assume negative values (*blueshift*) when the source is approaching the observer and positive values (*redshift*) when the source is receding.

By definition, since photons are non-massive particles, they move along null geodesics $ds^2 = 0$. Now consider an observer at distance d from a light source and consider again Eq. (1.11), i.e. the FLRW metric in polar coordinates. Assume $d\theta = d\phi = 0$ for simplicity. We obtain:

$$ds^2 = (cdt)^2 - a^2(t)\frac{dr^2}{1-\kappa r^2} = 0 \ . \tag{1.19}$$

If a photon is emitted from a point with coordinates $(r, 0, 0)$, at some time t_{em}, and it is observed at time t_{obs}, integrating the metric along the path, we obtain:

$$\int_{t_{em}}^{t_{obs}} \frac{cdt}{a(t)} = \int_0^r \frac{dr'}{\sqrt{1-\kappa r^2}} = F(r) \ . \tag{1.20}$$

Let us consider the emission of a second photon from the light source at $t'_{em} = t_{em} + \delta t_{em}$ that reaches the observer at $t'_{obs} = t_{obs} + \delta t_{obs}$. Since we know that, in a system of comoving coordinates, $F(r)$ is independent of the expansion of the Universe, the difference between the two photon paths is given only in terms of time:

$$\int_{t_{em}}^{t_{obs}} \frac{cdt}{a(t)} = F(r) = \int_{t'_{em}}^{t'_{obs}} \frac{cdt}{a(t)} \ . \tag{1.21}$$

Thus, if the time intervals δt_{em} and δt_{obs} are small enough, $a(t)$ can be considered almost constant. It follows that

$$\frac{\delta t_{em}}{a(t_{em})} = \frac{\delta t_{obs}}{a(t_{obs})} \ . \tag{1.22}$$

Now, recalling the definition of z (Eq. (1.18)) and how wavelength and frequency are related by the relationship $\lambda = c/\nu$, with $\nu = 1/\delta t$, we can define:

$$1 + z \equiv \frac{a_0}{a(t)} \ . \tag{1.23}$$

The latter equation can be regarded as one of the pillars of Cosmology, as it manages to link an easily observable quantity (the redshift z) to the expansion parameter $a(t)$. However, it is important to remember that there are further contributions to the wavelength distortion of photons emitted by distant objects, such as proper motions, effects due to the presence of mass (e.g. gravitational lensing), interactions with scattered matter (e.g. the Sunyaev-Zel'dovich effect) and other relativistic effects.

1.4 The Friedmann equations

From the assumption of the validity of the Cosmological Principle and General Relativity, considering the energy-momentum tensor applied to a perfect fluid with pressure p, density ρ and four-velocity u,

$$T_{\mu\nu} = -pg_{\mu\nu} + (p + \rho c^2)u_\mu u_\nu \ , \tag{1.24}$$

it is possible to apply the FLRW metric to solve Einstein's field equations. These, set of 16 equations, of which only two are independent (*time-time*, *space-space*), provide the time evolution for $a(t)$ and thus the dynamic evolution of the Universe. The resulting equations are called the *first and the second Friedmann Equations* and can be expressed as follows:

$$\ddot{a} = -\frac{4\pi}{3}G\left(\rho + \frac{3p}{c^2}\right)a \quad (time) \ , \tag{1.25}$$

$$\dot{a}^2 + \kappa c^2 = \frac{8\pi}{3}G\rho a^2 \quad (space) \ . \tag{1.26}$$

These equations are linked together by the *adiabatic condition*

$$d\mathcal{U} = -pdV \ , \tag{1.27}$$

where $\mathcal{U}$ and V represent the internal energy and the volume of the Universe, respectively. This condition can be expressed, considering the evolution of the Universe, as

$$d(\rho c^2 a^3) = -pda^3, \tag{1.28}$$

from which we derive

$$\dot{\rho} + 3\left(\rho + \frac{p}{c^2}\right)\frac{\dot{a}}{a} = 0 \ . \tag{1.29}$$

The density, ρ, and the pressure, p, in these equations have to be considered as the sum of all the densities and all the pressures of the different components present in the Universe.

1.5 The cosmological constant and the Einstein model

When the russian mathematician Alexander Friedmann derived his equations in 1922, the most popular idea was that of a static Universe without preferred times and directions, as predicted assuming the *Perfect Cosmological Principle*. In order to satisfy this hypothesis, we must have that $\ddot{a} = \dot{a} = 0$. The only way this can be verified is by having, from Eqs. (1.25) and (1.26),

$$\rho = -3\frac{P}{c^2} \ . \tag{1.30}$$

In other words, either density or pressure must assume a negative value, but this leads to an unphysical condition. Einstein, who at the time was deeply convinced

of the static nature of the Universe, decided to solve the problem by introducing in his equations a constant term, the *cosmological constant* Λ. The field equations are modified by this term in the following way:

$$R_{\nu\mu} - \frac{1}{2} g_{\nu\mu} R - \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \ . \tag{1.31}$$

We can re-write Eq. (1.31) by modifying the energy-momentum tensor, so as to obtain a similar form to the original one:

$$R_{\nu\mu} - \frac{1}{2} g_{\nu\mu} R = \frac{8\pi G}{c^4} \tilde{T}_{\mu\nu} \ , \tag{1.32}$$

where the new tensor is given by

$$\tilde{T}_{\mu\nu} \equiv T_{\mu\nu} + \frac{\Lambda c^4}{8\pi G} g_{\mu\nu} = -\tilde{p} g_{\mu\nu} + (\tilde{p} + \tilde{\rho} c^2) u_\mu u_\nu \ , \tag{1.33}$$

with $\tilde{p}$ and $\tilde{\rho}$ *effective pressure* and *effective density* respectively, defined as:

$$\tilde{p} \equiv p - \frac{k c^4}{8\pi G} \ , \qquad \tilde{\rho} \equiv \rho + \frac{\Lambda c^2}{8\pi G} \ . \tag{1.34}$$

It is straightforward to show that

$$\Lambda = \frac{k}{a^2} \quad [L^{-2}] \ , \tag{1.35}$$

which demonstrates how, in Einstein's model, Λ is linked to the curvature parameter. This value, in order to make physical sense, must be positive. In other words, it must be verified that $k = +1$. Finally, we can calculate what is the value of $\Lambda = \Lambda_e$ that makes the Universe static:

$$\Lambda_e = \frac{4\pi G \rho}{c^2} \ . \tag{1.36}$$

The universe imagined by Einstein therefore has spherical geometry and is static as long as the condition $\Lambda = \Lambda_e$ is verified.

The concept of the cosmological constant has long been regarded by Einstein (but more generally by the scientific community) as his greatest mistake. Not only the observations made by Hubble, just a few years later in 1929, disproved the hypothesis of a static Universe, but this type of model is afflicted by two major weaknesses:

- the value of Λ required to his purpose should have been extremely fine-tuned;
- the solutions of Eq. (1.31) would have led to an unstable equilibrium (Bianchi & Rovelli, 2010).

Nowadays, the concept of the cosmological constant has been rehabilitated. Not only do we know that the Universe is not static, but even since the 1990s, through observations of distant type Ia supernovae (SNIa), we know that the Universe is expanding at an accelerated rate (Riess et al., 1998). This type of expansion cannot be justified in any way by the normal field equations. It was necessary to reintroduce the idea of the cosmological constant, as a term with a repulsive effect. More specifically, Λ takes different meanings depending on its position in the Einstein's field equation:

- **left-hand side**: Λ can be interpreted as a modification of gravity;
- **right-hand side**: Λ can be interpreted as an additional negative energy contribution called *dark energy* (DE).

1.6 Friedmann models

Equations (1.25) and (1.26), introduced by Friedmann, describe the various possible evolutions of the universe depending on the assumed curvature. However, in order to achieve an analytical solution, it is necessary to introduce an *equation of state* which links together the various types of fluid that compose the Universe. In Cosmology, it is possible to assume the condition of a perfect fluid. This results in the following formulation for the equation of state:

$$p = w\rho c^2 \tag{1.37}$$

where w is defined so that the sound speed is:

$$c_s = \left(\frac{\partial p}{\partial \rho}\right)_s^{1/2} = c\sqrt{w} \ . \tag{1.38}$$

where w, to have physical meaning, must take a value between

$$0 \leq w \leq 1 \ . \tag{1.39}$$

This is known as the *Zel'dovich interval*. The speed of sound must therefore be positive but can never be greater than c. The value of w depends on the type of component considered. In particular, for non-relativistic matter, such as dust, $w \sim 0$ and therefore the pressure can be considered negligible. For non-degenerate relativistic matter, $w = 1/3$. This is the case for radiative fluid or more generally for non-massive or relativistic particles such as photons and neutrinos. A more exotic case occurs when we assume an equation of state for rigid matter (for example, when we are dealing with a Bose-Einstein condensate). In this case $w = 1$, and therefore $p = \rho c^2$. Finally, despite not being predicted by classical physics, negative values of w ($-1 \leq w < 0$) can be obtained. In particular, the cosmological constant Λ is defined by having $w = -1$.

With these definitions it is now possible to express the energy-momentum tensor as the sum of all the i components:

$$T_{\mu\nu} \equiv \sum_i T^i_{\mu\nu} \ . \tag{1.40}$$

By combining Eqs. (1.28) and (1.37), we can derive how the density of each cosmological component varies over time with the expansion of the Universe:

$$\rho_w \propto a^{-3(1+w)} \propto (1+z)^{3(1+w)} \tag{1.41}$$

This relation shows that the universe was dominated by different components at different times.

It is possible to derive the value of the density, called the *critical density* ρ_{crit}, for a Universe with a flat geometry:

$$\rho_{crit}(t) \equiv \frac{3}{8\pi G}\left(\frac{\dot{a}}{a}\right)^2 = \frac{3\dot{H}^2(t)}{8\pi G} \ . \tag{1.42}$$

This value must be extremely fine-tuned to obtain a flat geometry. In the case where $\rho < \rho_{crit}$ an open Universe is obtained and consequently an evolution leading to an eternal expansion takes place. In the case where $\rho > \rho_{crit}$, a closed Universe is obtained, implying a deceleration leading to a contraction. The present value of ρ_{crit}, depending on Hubble constant H_0, is (Planck Collaboration, 2020):

$$\rho_{crit}(t = t_0) \equiv \rho_{crit,0} \simeq 1.9h^2 \times 10^{-29}\ \mathrm{g\,cm^{-3}} \ . \tag{1.43}$$

Thanks to this definition, we can rewrite equation (1.26) to obtain a formulation for the curvature κ:

$$\frac{\kappa}{a^2} = \frac{1}{c^2}\left(\frac{\dot{a}}{a}\right)^2\left(\frac{\rho_{crit}}{\rho} - 1\right) \ . \tag{1.44}$$

From the definition of critical density, a dimensionless parameter Ω, the *density parameter*, can be defined:

$$\Omega(t) \equiv \frac{\rho(t)}{\rho_{crit}(t)} \ . \tag{1.45}$$

which can be expressed for each component, Ω_{w_i}. By definition, the *total density parameter* is the sum of all the fluid components permeating the universe:

$$\Omega_{tot} = \sum_i \Omega_{w_i} \ . \tag{1.46}$$

It is immediately understandable that for a flat Universe $\Omega_{tot} = 1$, while for an open Universe $\Omega_{tot} < 1$ and for a closed Universe $\Omega_{tot} > 1$.

By using the definitions (1.16) and (1.45), the Eq. (1.44) can be reformulated as:

$$1 - \Omega(t) = \frac{\kappa c^2}{a^2(t)H^2(t)} \ . \tag{1.47}$$

The sign of the right-hand side of the equation is governed by the value of κ, which cannot change over time. It follows that the left-hand side cannot change sign either, and therefore a Universe governed by the Friedmann equations cannot modify its geometry during its evolution.

From Friedmann's second equation, rewritten in terms of H, Ω and z, which are more representative parameters of the observable Universe, we obtain a formulation that parameterises the H evolution as a function of redshift z:

$$H^2(z) = H_0^2(1+z)^2\left[1 - \sum_i \Omega_{0,w_i} + \sum_i \Omega_{0,w_i}(1+z)^{1+3w_i}\right] = H_0^2 E^2(z) \ . \tag{1.48}$$

The term $1 - \Omega_{tot}$ is the so-called *curvature density parameter*.

1.6.1 The Big Bang singularity

Assume that the Universe consists of only one fluid component and that this respects ordinary physics ($0 \leq w \leq 1$). Friedmann's second equation (1.26) becomes

$$\ddot{a} = -\frac{4\pi}{3} G\rho(1+3w)a \ , \tag{1.49}$$

$\ddot{a}$ must be strictly negative: the function $a(t)$ has downward concavity and cannot have an inflection point (only in the case of a fluid with $w < -1/3$ an inflection can occur). Moreover, we know from observations that the Universe is expanding, so $\dot{a} > 0$. The function $a(t)$ is therefore monotonously increasing and, going back in time, we have that inevitably this function intersects the time axis, i.e. there must exist a some finite time in the past at which $a(t) = 0$ and in which temperature, density and expansion rate diverge:

$$\lim_{t\to 0} \rho(t) = \lim_{t\to 0} \left(\frac{a_0}{a}\right)^{-3(1+w)} \to +\infty \ . \tag{1.50}$$

This singularity is called the *Big Bang* (BB). The true physical conditions at the time of the BB are unknown. For times $t < t_p$, where $t_p \sim 10^{-43}\,\mathrm{s}$ is the *Planck time*, it is no longer possible to neglect the quantum effects of gravity. To date, a universally accepted treatment of quantum gravity has not yet been developed. Note that the expansion of the universe emerging from the BB model is a result of the initial conditions describing a homogeneus and isotropic Friedmann Universe, and it is not due to pressure.

1.6.2 The Einstein-de Sitter model vs open and closed universes

A generic model that includes the hypothesis of mono-component fluid and that assumes a flat geometry ($\kappa = 0$, $\Omega = 1$) is called *Einstein-de Sitter Model* (EdS). With these assumptions, equation (1.48) reduces to

$$H(z) = H_0(1+z)^{\frac{3(1+w)}{2}} \ . \tag{1.51}$$

Our Universe is made up of three main components: radiation, matter and Λ. Each of these three components, being associated with different values of w, becomes dominant at a certain stage in the evolution of the Universe itself (as shown in Fig. 1.1, which plots showing the evolution of the density of the various components of the Universe as a function of redshift), which, within each of these stages, can therefore be approximated as composed of a single type of fluid. So we can divide the history of the Universe into epochs based on which component was the dominant in that time interval. In particular, at early times the dominant component results to be the radiation (*radiation-dominated era*) while at late times the matter component becomes the most relevant (*matter-dominated era*). Moreover, interpreting the dark energy component as a fluid with $w = -1$, we can demonstrate that its density is independent of the time and starts to be dominant only at very recent epochs (*DE-dominated era*).

The Table 1.1 provides some useful relationships for an EdS universe in the general case and when it is dominated by radiation or matter. From these relations we can derive the density evolution for the matter component ($w = 0$):

$$\rho_m = \rho_{0,m}(1+z)^3 \ , \tag{1.52}$$

and for the radiation component (w = 1/3):

$$\rho_r = \rho_{0,r}(1+z)^4 \ . \tag{1.53}$$

By equalising these relationships we find the moment of *matter-radiation equivalence*, the moment when the densities of the two components had identical values:

$$z_{eq} = \frac{\rho_{0,m}}{\rho_{0,r}} - 1 \simeq 3 \cdot 10^3 \ . \tag{1.54}$$

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

Table 1.1: Dependencies obtained for an EdS universe in the general case, in the case of a matter-dominated universe and in the case of a radiation-dominated universe.

Let us now consider the cases of curved universes, i.e. models with a single component and $\Omega \neq 1$. Consider Eq. (1.48) in the case of a single component. The curvature term $(1 - \Omega_0)$ is constant. The second term depends on time. We look for the value $a = a^*$ that equals the curvature term:

$$|1 - \Omega_0| = \Omega_0 \left(\frac{a_0}{a^*}\right)^{1+3w} \ . \tag{1.55}$$

When $a \ll a^*$ or $z \gg z^*$, i.e. in the phases immediately following the Big Bang, the equation becomes

$$H(z) = H_0 \Omega_0^{1/2} (1+z)^{\frac{3(1+w)}{2}} \ , \tag{1.56}$$

which differs from Eq. (1.51) only by the constant factor $\Omega_0^{1/2}$. All universes, in their initial phases, behave in a similar way to the EdS one and their geometry is assumably flat.

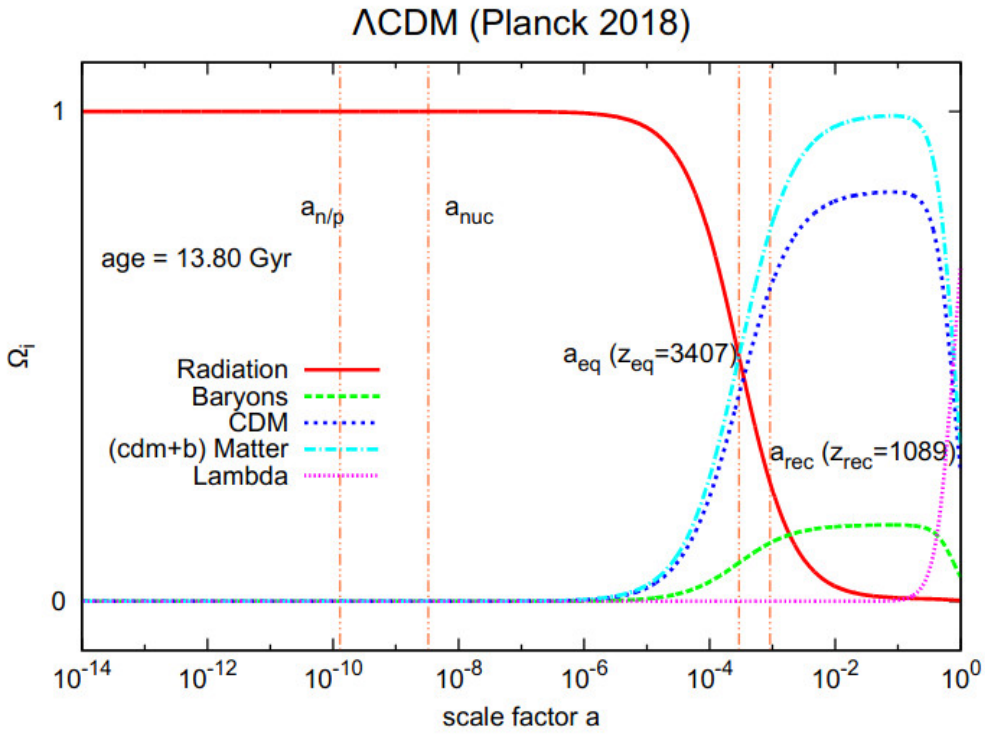

Figure 1.1: Evolution of the density parameters of the various components modelled assuming a ΛCDM Cosmology and Planck data. Credits: Rindler-Daller (2020).

1.7 The Standard cosmological model

A one-component model is unable to fully describe the properties of the present Universe. Since the beginning of the 21st century, the ΛCDM model has gradually become established. As the simplest model in agreement with observations, it is commonly referred to as the *standard cosmological model*. The flat-ΛCDM model, or more simply ΛCDM, takes its name from the assumption of an almost flat geometry and from the main components of the cosmological fluid at $z = 0$, which are assumed to be the cosmological constant (Λ), $\Omega_{0,\Lambda} \simeq 0.7$, and the cold dark matter (CDM), $\Omega_{0,CDM} \simeq 0.25$, a non-collisional and non-relativistic matter component. In addition, a small component of baryonic matter, $\Omega_{0,b} \simeq 0.05$, and a negligible component of radiative fluid, $\Omega_{0,r} \simeq 10^{-5}$, consisting of all relativistic particles (mainly photons and neutrinos) are present. Since DM does not interact with baryonic matter except gravitationally, it does not emit photons and is therefore extremely difficult to detect. However, despite the fact that its nature is still unknown, there are strong indications of its presence, such as, for example, the measures of the rotation curves of disk galaxies (Bosma, 1999), the distortion of light due to the phenomenon of gravitational lensing (Ellis 2010, Gilman et al. 2018) and the analysis of the CMB's black body spectrum (Galli et al., 2011). However, it should be pointed out that it is possible to relate the model to observations without including DM by using modified gravity models (Famaey & McGaugh, 2012).

According to this model, the evolution of the Universe is effectively described

by the evolution of its temperature. At present, the various components are not in thermal equilibrium and therefore their relative temperatures take on different values. At high redshifts, however, the situation was different. In fact, the temperature was high enough to keep matter ionised, ensuring a sufficient number of interactions to maintain equilibrium between matter and radiative components. Due to its continued expansion, around $z = 1500$ the temperature of the Universe dropped to a point where protons and electrons recombined, causing the two cosmological fluids to decouple and allowing photons to freely propagate. The probability for a photon to be scattered for the last time by the primordial plasma has a peak at $z = 1100$. This moment is called the *redshift of the last scattering*. The radiation emitted at this moment is now observable, enormously reddened by the expansion of the Universe, and takes the name of *Cosmic Microwave Background radiation* (CMB). The currently measured temperature of the CMB is $T = 2.7255 \pm 0.0006\, K$ (Planck Collaboration, 2020).

In order to justify the observed flat geometry and uniformity of the CMB, it is convenient to assume a period of inflationary expansion, i.e. a phase of exponential accelerated expansion, in the early epoch of the Universe. Moreover, the inflationary era gives rise to the density perturbations which originate the cosmic structures observed today.

As introduced previously, the fundamental component of the Universe predicted by the standard cosmological model is the cosmological constant Λ, in form of dark energy. Not only the inclusion of this component justifies the flatness of the Universe, but, due to its characteristics, it also justifies the accelerated expansion. However, the nature of DE is still unknown and is one of the most important cosmological problems still open.

Chapter 2

Structure formation

Looking at the present Universe on the Megaparsec scales, it is evident how the assumptions of homogeneity and isotropy are violated. Strong fluctuations in the density field are present, up to several orders of magnitude from the average. However, looking at the CMB, it can be seen that the situation in the early stages of the Universe evolution was very different. Assuming adiabatic perturbations, for which the relation is therefore valid:

$$\frac{\delta T}{\overline{T}} \sim \frac{\delta \rho}{\bar{\rho}} \,, \tag{2.1}$$

we observe that in the CMB:

$$\frac{\delta T}{\overline{T}} \approx 10^{-5} \,, \tag{2.2}$$

where $\overline{T}$ represents the average black body temperature of the cosmic microwave background. Therefore, we can conclude that the Universe was almost homogeneous at this time.

The nonlinear structures observed nowadays are a consequence of gravitational instability that allows perturbations to grow at a sufficient rate. The evolution of perturbations subjected to gravitational force was described analytically by Jeans as early as 1902. Jeans' theory, illustrated in Sec. 2.1, predicts that the small oscillations in the density field in the early Universe grow with time, resulting in the collapsed structures observed today. However, the description given by this model is valid only as long as the perturbations remain in a linear regime. The analytical description of evolution in the nonlinear regime is very complex and is only available for two simple cases: spherical evolution and the Zel'dovich approximation (Sec. 2.2). To date, this type of solution is normally treated by N-body simulations (Sec. 2.4).

2.1 Linear theory

Jeans theory is applicable for non-relativistic fluids with perturbations on a maximum scale corresponding to the cosmological horizon, defined as the sphere causally connected to the observer and parameterised as

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

At scales $r > R_H$, perturbations evolve exclusively under the action of gravity, but they must be treated with relativistic theory. In this case, the perturbations can always grow and give rise to collapse. For scales $r < R_H$, on the other hand, where the microphysical processes become important, the Jeans model provides a reliable description of these phenomena in linear theory.

2.1.1 Jeans instability in a static universe

Consider a simplified model in which the Universe is assumed to be static. Assume a homogeneous and isotropic background, that is, we have density ρ_0, pressure p_0, gravitational potential ϕ_0 and velocity $v_0 = 0$ as the background quantity of the fluid. The condition whereby both potential and density are spatially constant is non-physical as it only satisfies Poisson's equation for $\rho_0 = 0$. However, it is equally assumable for the purposes of the model and immediately resolved by the introduction of the expansion factor in the case of an expanding universe. The equations of motion of such a fluid, in the Newtonian approximation, are:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0 \qquad \text{continuity equation} \tag{2.4}$$

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

$$\nabla^2\Phi = 4\pi G\rho \qquad \text{Poisson equation} \tag{2.6}$$

$$\frac{\mathrm{d}S}{\mathrm{d}t} = 0 \qquad \text{adiabatic condition} \tag{2.7}$$

$$p = p(\rho, S) = p(\rho) \qquad \text{equation of state} \tag{2.8}$$

where S is the entropy of a fluid element.

Small adiabatic perturbations δq ($\delta q/q \ll 1$, where q is a generic quantity) are introduced to the equilibrium state. The system formed by the Eqs. (2.4), (2.5), (2.6), where the substitution $q \to q_0 + \delta q$ took place, can be linearised (small oscillations). We also define the *density contrast* δ as

$$\delta(\vec{x}, t) = \frac{\delta\rho(\vec{x}, t)}{\rho_0} \ . \tag{2.9}$$

Solutions/perturbations, separated in space-time dependence, are sought in Fourier space:

$$S(\vec{x}, t) = S_k e^{i(\vec{k}\cdot\vec{x} - \omega t)} \ . \tag{2.10}$$

The solving relation for the density contrast is a differential equation, which in Fourier space has the form:

$$\ddot{\delta}_k + (k^2 c_s^2 - 4\pi G\rho_0)\delta_k = 0 \ , \tag{2.11}$$

where $k = |\vec{k}|$ is the absolute value of the wavenumber, δ_k the amplitude of the Fourier transform of δ and $c_s = \sqrt{\partial p/\partial\rho}$ the speed of sound. Equation (2.11), known as the *dispersion relation*, has two indipendent solutions

$$\delta_k \propto \exp(\pm i\omega(k)t) \ , \tag{2.12}$$

where

$$\omega(k) = \sqrt{k^2 c_s^2 - 4\pi G\rho_0}\ . \tag{2.13}$$

The behavior of the solutions depends on the sign of $\omega^2(k)$. If $\omega^2 < 0 \rightarrow \omega \in \mathbb{C}$, two exponential solutions exist, one increasing and one decreasing. If $\omega^2 > 0 \rightarrow \omega \in \mathbb{R}$, we obtain two logitudinal small amplitude waves with opposite direction. The limit case $\omega^2 = 0$ allows to define two important quantities:

$$k_J := \sqrt{\frac{4\pi G\rho_0}{c_s^2}} \qquad \text{Jeans wavenumber} \tag{2.14}$$

$$\lambda_J := \frac{2\pi}{k_J} = c_s\sqrt{\frac{\pi}{G\rho_0}} \qquad \text{Jeans length} \tag{2.15}$$

Thus, the propagation and evolution of perturbations depends on how their wavelength relates to the characteristic one, λ_J. If:

- $\lambda < \lambda_J$ (or $k > k_J$), the perturbation propagates as a sound wave with constant amplitude and with a phase velocity $c_{ph} = \omega/k$. This velocity tends to the sound speed at small scales $\lambda \ll \lambda_J$;
- $\lambda > \lambda_J$ (or $k < k_J$), the perturbation has two solutions, with opposite modes. There is an increasing mode and a decreasing mode. The increasing case is the interesting one from the cosmological point of view because it allows the fluctuation to grow exponentially and permit collapse. This phenomenon is called *Jeans instability*.

We can therefore conclude that small perturbations evolve according to their scale, and when they can grow, they do it exponentially.

2.1.2 Jeans instability in expanding universe

Let us now attempt to investigate a more realistic context. We want to analyze the evolution of perturbations in an expanding universe. In this framework, the background density is a function of time, $\rho_0 = \rho_0(t)$, so the continuity Eq. (2.4) becomes

$$\dot{\rho_0} + 3H(t)\rho_0 = 0\ . \tag{2.16}$$

The velocity is no longer zero but follows Hubble's law $\vec{v_0} = H(t)\vec{x}$. so we deal with a two-component velocity:

$$\vec{u} = \dot{\vec{x}} = H(t)\vec{x} + v_p \tag{2.17}$$

where v_p indicates the *peculiar velocity* of the fluid. Introducing a perturbation for the equation set (2.4), (2.5) and (2.6) in the same way as for a static universe, it is possible to obtain a new form of the dispersion relation, from which we obtain a solution for each Fourier mode of the type $\delta = \delta_k(t)\exp(i\vec{k}\vec{x})$:

$$\ddot{\delta_k} + 2H(t)\dot{\delta_k} + (k^2 c_s^2 - 4\pi G\rho_0)\delta_k = 0\ . \tag{2.18}$$

The term $2H(t)\dot{\delta_k}$ is known as the *Hubble friction*, while $k^2 c_s^2 \delta_k$ takes into account the characteristic velocity field of the fluid. Both terms oppose perturbation growth. This is a second order differential equation for $\delta(\vec{x}, t)$. Again, we can separate the nature of the solutions according to the value of λ relative to λ_j (Eq. (2.15)):

- $\lambda < \lambda_J$ (or $k > k_J$), the perturbation propagates as a stationary wave, exactly as in the case of a static universe;
- $\lambda > \lambda_J$ (or $k < k_J$), two solutions exist, with growing and decaying mode solutions:

$$\delta(\vec{x}, t) = A(\vec{x})\delta_+(t) + B(\vec{x})\delta_-(t) \ , \tag{2.19}$$

  where A and B are functions of the comoving coordinate and δ_+ and δ_- represent respectively the growing and decaying modes and are functions of the time.

We can rewrite the Eq. (2.18) for δ as an equation for the growth factor, with the time dependence replaced by the redshift dependence:

$$\ddot{\delta} + 2H(z)\dot{\delta} - \frac{3}{2}\Omega_{m,0}H_0^2(1+z)^3\delta = 0 \ . \tag{2.20}$$

Consider now a flat universe of only matter ($\Omega_m = 1$). We can apply the relations described in Table 1.1. The following dependencies are obtained:

$$\delta_+(t) \propto t^{2/3} \propto a \ , \tag{2.21}$$

$$\delta_-(t) \propto t^{-1} \propto a^{-3/2} \ . \tag{2.22}$$

It is immediate to understand how the δ_- solution decays with time while the δ_+ solution leads to gravitational instability. We are only interested in the growing solution, as it is the one that leads to the development of structures. For a generic universe, the growing solution has an integral form given by the following equation:

$$\delta_+(z) = H(z)\int_z^\infty \frac{(1+z')}{H^3(z')}dz' \ , \tag{2.23}$$

which has no analytical solution. There is, however, a pseudo-analytical relation that allows us to simplify the calculation:

$$f \equiv \frac{\mathrm{d}\log\delta_+}{\mathrm{d}\log a} \simeq \Omega_m^\gamma + \frac{\Omega_\Lambda}{70}\left(1 + \frac{1}{2}\Omega_m\right) \ . \tag{2.24}$$

This relation is very sensitive to the variation of Ω_m while less to that of Ω_Λ: perturbations are influenced more by matter than by the cosmological constant. The exponent γ is predicted to have a value approximately of 0.545 in GR (Coles & Lucchin, 2002). Its measure can be used as a test for the theory itself.

2.1.3 Statistical properties of the Universe

Until now, we analysed the linear evolution of a single perturbation of the density field, whose growth is defined by $\delta(\vec{x}, t) = \delta_+(t)\delta(\vec{x})$. However, the real evolution of the structures is given by the superposition of density fluctuations at different scales. Having described the perturbations in Fourier space is very useful, as it allows us to represent this situation as a superposition of plane waves independent of each other.

Let us introduce the *spatial Fourier transform* of $\delta(\vec{x})$:

$$\delta(\vec{k}) = \frac{1}{(2\pi)^3}\int \delta(\vec{x})e^{-i\vec{k}\cdot\vec{x}}d^3\vec{x} \ . \tag{2.25}$$

We can now define the *power spectrum* of the density field as the variance of the amplitudes at a given value of the wavenumber k:

$$\langle \delta(\vec{k})\delta(\vec{k'})\rangle = (2\pi)^3 P(k)\delta_D^{(3)}(\vec{k}-\vec{k'}) \,, \tag{2.26}$$

where $\delta_D^{(3)}$ represents the 3-dimensional Dirac delta function

$$\delta_D^{(3)}(\vec{k}) = \frac{1}{(2\pi)^3}\int e^{-i\vec{k}\cdot\vec{x}} d^3\vec{x} \,. \tag{2.27}$$

The power spectrum is related to the *two-point correlation function* (2PCF), $\xi(r)$, its analogue in real space, via the relationship provided by the *Wiener-Khintchine theorem*. It states that the autocorrelation function of a *wide-sense-stationary random process* has a spectral decomposition given by the power spectrum of that process. So the 2PCF can be defined as

$$\xi(r) = \frac{1}{(2\pi)^3}\int P(k) e^{i\vec{k}\cdot\vec{x}} d^3\vec{k} \,, \tag{2.28}$$

or, statistically as

$$\langle \delta(\vec{x})\delta(\vec{x'})\rangle = \xi(|\vec{x}-\vec{x'}|) = \xi(\vec{r}) = \xi(r) \,, \tag{2.29}$$

where r is the comoving distance between x and x', and $\xi(\vec{r}) = \xi(r)$ is due to the CP. We can interpret $\xi(r)$ as the probability excess (or defect) dP_{12} of finding a pair of objects separated by a comoving distance r, in two independent volume elements dV_1 and dV_2, with respect to a random uniform distribution:

$$dP_{12} = n^2[1+\xi(r)]dV_1 dV_2 \,. \tag{2.30}$$

When $\xi(r) = 0$, the measured distribution coincides with the random one.

According to inflationary theory, primordial perturbations are generated by stochastic quantum fluctuations of a scalar field (i.e., the inflaton) (Guth & Pi, 1982). The amplitude of the perturbations is described by a Gaussian distribution. It is also assumed that there is no preferential scale during the formation of the perturbations. It follows that the initial power spectrum can be described by a power law:

$$P(k) = Ak^n \,. \tag{2.31}$$

The spectral index, n, is generally assumed to be close to unity (Zeldovich, 1972). Unlike n, which is well fixed by theory, the amplitude A needs to be determined through observations. In particular, the most reliable and precise measure of A is obtained from the analysis of the temperature fluctuations in the CMB (Planck Collaboration, 2020).

From the assumption of Gaussianity of the perturbation amplitude distribution, it follows that the mean of it is formally equal to zero. However, the same cannot be assumed for the variance. The fluctuation amplitude variance σ^2 is defined by:

$$\sigma^2 = \langle|\delta(\vec{x})^2|\rangle = \sum_k \langle|\delta(\vec{k})^2|\rangle = \frac{1}{V_u}\sum_k \delta_k^2 \,, \tag{2.32}$$

where the average is taken over an ensemble of the Universe realisations of volume V_u. By assuming the validity of the CP and considering the limit $V_u \to \infty$, it follows that

$$\sigma^2 \to \frac{1}{2\pi^2} \int_0^\infty P(k) k^2 dk \ . \tag{2.33}$$

The expression for σ^2 obtained via Eq. (2.32) describes the point variance. In order to use this approach, it is required to know every point of the density field. Obviously, obtaining this reconstruction is generally very complex. Instead of using point variance, it is convenient to treat the fluctuation field as dependent on the scale R, which can be thought of as a filter function. With this approach we can recover the density fluctuation δ_M from a discrete distribution of tracers as

$$\delta_M = \frac{M - \langle M \rangle}{\langle M \rangle} \ , \tag{2.34}$$

where $\langle M \rangle$ is the mean mass present inside a spherical volume of radius R. δ_M can also be thought as the convolution between the point variance and a filter function, called *window function* W:

$$\delta_M(\vec{x}) = \delta(\vec{x}) \otimes W(\vec{x}, R) \ . \tag{2.35}$$

The combination of the definition (2.34) and Eq. (2.32) provides the so-called *mass variance*:

$$\sigma_M^2 = \langle \delta_M^2 \rangle = \frac{\langle (M - \langle M \rangle)^2 \rangle}{\langle M \rangle^2} \ . \tag{2.36}$$

Through the convolution theorem, setting us in the limit $V_u \to \infty$, it is possible to switch from Eq. (2.33) to

$$\sigma_M^2 = \frac{1}{(2\pi)^3} \int P(k) \hat{W}^2(\vec{k}, R) d^3\vec{k} \ , \tag{2.37}$$

where $\hat{W}$ is the Fourier-transform of the window function and is a function of R. 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 $\lambda \sim k^{-1} > R$.

As already explained, normalization of the power spectrum is not provided by theory. To estimate the fluctuations amplitude at present time, it is very common to use the parameter σ_8, defined as the value of the mass variance computed with a filtering of $R = 8\,h^{-1}$Mpc in the local universe:

$$\sigma_8^2 = \frac{1}{2\pi^2} \int P(k) k^2 \hat{W}^2(k, R = 8\,h^{-1}\text{Mpc}) dk \ . \tag{2.38}$$

The square root of this quantity, i.e. σ_8, besides representing the mass fluctuation in spheres with radius $8\,h^{-1}$Mpc is a free parameter that represent the normalization of the power spectrum and is essential in predicting the phenomenology of the local Universe.

2.2 Nonlinear theory

The structures we observe in today's universe are consequence of the gravitational collapse of the primordial fluid as a result of a nonlinear evolution of perturbations. The description of these objects can no longer be based on a linear theory since the values of the density contrast involved are much greater than unity and the approximation of small perturbations can no longer be applied. Moreover, even at the stage of semilinear evolution, the perturbation distribution function starts to evolve into a non-Gaussian shape. A further complication arises from the need to consider the evolution of baryonic matter as different from the evolution of DM. Baryonic matter also is subject to the laws of fluid dynamics and their effects, as star formation and evolution, AGN feedback, heating and cooling of gas. All of these phenomena make even more difficult the description of the whole scenario with a full and solid theory. At present, the most effective method of fully describing these phenomena and nonlinear evolution is to resort to numerical N-body simulations. However, some useful analytical solutions exist for specific cases, such as *spherical evolution* and the transition between the linear and nonlinear regimes, described through the *Zel'dovich approximation*.

2.2.1 Spherical evolution

The so-called *spherical evolution model* (Gunn & Gott, 1972) describe the isolated formation of spherical collapsed overdensities (i.e. DM haloes) and underdensities (i.e. cosmic voids). Consider an initially spherical top-hat perturbation, positive or negative. We can model it as a series of concentric, uniform shells. An important feature of this type of model is that, as predicted by Sheth & Van De Weygaert 2004, the evolution of the considered fluctuation is given uniquely by the total shell energy. Assuming the validity of the CP, we can suppose that each perturbation can be treated as an independent Friedmann universe until it evolves adiabatically. Based on these assumptions, we can describe the evolution of a spherical, isolated perturbation as a closed (overdensity) or open (underdensity) universe evolving in an EdS background from an initial time $t_i > t_{eq}$, where t_{eq} is the *matter-radiation equivalence time* (evolution of perturbations in the matter-dominated cosmic epoch).

The initial density distribution of our model can be expressed as follows:

$$\rho(r, t_i) = \rho_0(t_i) + \delta\rho(r, t_i) = \rho_0(t_i)[1 + \delta_i(r)], \tag{2.39}$$

where $\delta_i(r) = \delta_i(r, t_i)$ is the initial density contrast, which is a function of r. Since we are dealing with a spherical perturbation, it is convenient to make use of the proper radial coordinate $r = a(t)|\vec{x}|$, where $\vec{x}$ is the comoving coordinate and $a(t)$ the expansion factor. Consider an infinite shell located at distance r from the center. The motion of the matter contained in the shell is given by

$$\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = -\frac{GM}{r^2} = \frac{4\pi G}{3}\rho_0 a \ , \tag{2.40}$$

where:

$$M = \frac{4\pi G}{3} r\rho_0(1 + \Delta) \ , \tag{2.41}$$

$$\Delta = \frac{3}{r}\int_0^r \delta(r')r'^2 dr' \ . \tag{2.42}$$

Equation (2.40) is analogous to the First Friedmann Equation (1.25) for a single component universe:

$$\frac{\mathrm{d}^2 a}{\mathrm{d}t^2} = -\frac{4\pi G}{3}\rho_0 a \ , \tag{2.43}$$

and this proves the hypothesis that the perturbation evolves as a one-component Friedmann universe.

Starting from Eq. (2.40), the first integral of motion can be written as

$$\frac{1}{2}\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 - \frac{4\pi G}{3}\rho_0(1-\Delta)r^2 = E \ , \tag{2.44}$$

where E is a constant of integration, namely, the *total energy of the perturbation*. The fate of the shell is determined by the sign of E. If $E < 0$, $\dot{r}$ decreases as r increases, until it becomes negative. As a consequence, the shell decouples from the Hubble flow and collapses. On the other hand, if $E > 0$, $\dot{r}$ never becomes zero and the shell expansion continues indefinitely. Analogously to what was found above, Eq. (2.44) is similar to the second Friedmann equation (1.26) for a single component fluid:

$$\frac{1}{2}\left(\frac{\mathrm{d}^2 a}{\mathrm{d}t^2}\right) - \frac{4\pi G}{3}\rho_0 a^2 = -\frac{K}{2} \ , \tag{2.45}$$

which describes the expansion rate of the Universe. Having established these similarities, we can procede to derive the parametric solutions of the Friedmann equations for curved universes. The parametric solutions related to a massive shell are the following:

- $\mathbf{E > 0}$:

$$r = A(\cosh\theta - 1), \qquad t = B(\sinh\theta - \theta); \tag{2.46}$$

- $\mathbf{E < 0}$:

$$r = A(1 - \cos\theta), \qquad t = B(\theta - \sin\theta); \tag{2.47}$$

where A and B are two constants related by

$$A^3 = GMB^2 \tag{2.48}$$

and θ, called *development angle*, parameterises all the physical relevant quantities relating to the mass shell. It is possible to parameterize in the same way the model describing the background. Choosing the expansion parameter $a(t) = r_0(t)$ so that it encloses the same mass M as in Eqs. (2.46) and (2.47). Therefore we obtain:

- $\mathbf{E > 0}$:

$$r_0 = A_0(\cosh\eta - 1), \qquad t_0 = B_0(\sinh\eta - \eta); \tag{2.49}$$

- $\mathbf{E < 0}$:

$$r_0 = A_0(1 - \cos\eta), \qquad t_0 = B_0(\eta - \sin\eta); \tag{2.50}$$

where, as before, A_0 and B_0 are two constants related by

$$A_0^3 = GMB_0^2 \tag{2.51}$$

We can then calculate the mean density inside each shell by applying the relation $\rho = 3M/(4\pi r^3)$ and substituting with the dependencies (2.46) and (2.47) in equation (2.48). We obtain:

$$\rho(r,t) = \begin{cases} \frac{3}{4\pi G t^2} \frac{(\sinh\xi - \xi)^2}{(\cosh\xi - 1)^3} & E > 0 \text{ (open)} \\ \frac{3}{4\pi G t^2} \frac{2}{9} & E = 0 \text{ (critical)} \\ \frac{3}{4\pi G t^2} \frac{(\xi - \sin\xi)^2}{(1 - \cos\xi)^3} & E < 0 \text{ (closed)} \end{cases} . \tag{2.52}$$

The density contrast of a perturbation in a mono-component Universe can be expressed in the most general form through the following expression:

$$1 + \Delta(r,t) = \frac{\rho(r,t)}{\rho_0(r,t)} = \frac{f(\theta)}{f(\eta)} , \tag{2.53}$$

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

$$f(\xi) = \begin{cases} \frac{(\sinh\xi - \xi)^2}{(\cosh\xi - 1)^3} & E > 0 \text{ (open)} \\ \frac{2}{9} & E = 0 \text{ (critical)} \\ \frac{(\xi - \sin\xi)^2}{(1 - \cos\xi)^3} & E < 0 \text{ (closed)} \end{cases} . \tag{2.54}$$

In addition, the expansion or contraction velocity of a spherical shell can also be expressed as a function of θ and η. Let us consider the *peculiar velocity* of a shell, v_p:

$$v_p(r,t) = v(r,t) - H(t)r(t) , \tag{2.55}$$

where $v(r,t)$ represents the total velocity of the shell and $H(t)$ the Hubble parameter of the background Universe. It is further possible to define a generic Hubble parameter, H_s, for an arbitrary shell:

$$H_s(t) = \frac{\dot{r}}{r} = \frac{1}{r}\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{1}{t} g(\xi) , \tag{2.56}$$

where $g(\xi)$, the *cosmic velocity function*, is given by:

$$g(\xi) = \begin{cases} \frac{\sinh\xi(\sinh\xi - \xi)}{(\cosh\xi - 1)^2} & E > 0 \text{ (open)} \\ \frac{2}{3} & E = 0 \text{ (critical)} \\ \frac{\sin\xi(\xi - \sin\xi)^2}{(1 - \cos\xi)^2} & E < 0 \text{ (closed)} \end{cases} . \tag{2.57}$$

Through $g(\xi)$ it is possible to reformulate equation (2.55) as

$$v_p(r,t) = H(t)r(t)\left[\frac{g(\theta)}{g(\eta)} - 1\right] . \tag{2.58}$$

Relations (2.53) and (2.58) provide us explicit expressions for the evolution of a spherical perturbation in FLRW backgrounds with no cosmological constant. We will focus now on the case of a spherical perturbation evolving in an EdS background Universe.

Overdensities

Let us consider an initial overdense shell in an EdS universe and study its evolution: we know that in a flat universe a matter perturbation evolves as $\delta_+(t) \propto t^{2/3}$ or $\delta_-(t) \propto t^{-1}$ (Eqs. 2.21, 2.22), according to whether the mode considered is increasing or decaying. The density contrast can be expressed as the combination of the two modes:

$$\delta(t) = \delta_+(t_i)\left(\frac{t}{t_i}\right)^{2/3} + \delta_-(t_i)\left(\frac{t}{t_i}\right)^{-1} . \tag{2.59}$$

Assuming a motion consistent with the hubble flow for the perturbations (a null initial velocity relative to the background), we can compute the derivative of the latter relation with respect to the time considering $t = t_i$, finding:

$$\frac{2}{3}\delta_+(t_i) - \delta_-(t_i) \quad \longrightarrow \quad \delta_-(t_i) = \frac{2}{3}\delta_+(t_i) . \tag{2.60}$$

Therefore, the equation (2.59), for $t = t_i$, can be rewritten as

$$\delta(t_i) = \frac{5}{3}\delta_+(t_i) . \tag{2.61}$$

Initially, 3/5 of the perturbation is represented by the growing modes and 2/5 by the decaying ones. However, these modes decline over time, gradually becoming less and less relevant.

Consider now the density parameter of the perturbation, Ω_p. Treating the perturbation as a closed universe, we can impose the relation $\Omega_p > 1$:

$$\Omega_p(t_i) = \frac{\rho_p(t_i)(1+\delta_i)}{\rho_c(t_i)} = \Omega(t_i)(1+\delta_i) > 1 , \tag{2.62}$$

where ρ_c is the critical density and $\delta_i = \delta(t_i)$. Therefore, in order for a spherical perturbation to collapse, it is necessary that

$$(1+\delta_i) > \Omega(t_i)^{-1} . \tag{2.63}$$

Now, knowing that for Friedmann universes the relation

$$\Omega(z) = \frac{\Omega_0(1+z)^{1+3w}}{(1-\Omega_0) + \Omega_0(1+z)^{1+3w}} \tag{2.64}$$

is valid and imposing w=0 (perturbation of only matter), we obtain:

$$\delta_+(t_i) = \frac{3}{5}\delta_i > \frac{3}{5}\frac{(1-\Omega(t_i))}{\Omega(t_i)} = \frac{3}{5}\frac{(1-\Omega_0)}{\Omega_0(1+z)} . \tag{2.65}$$

To allow the perturbation to collapse, this inequality must be respected. It is immediate to understand how it is always respected for overdensity in closed or Eds

universes ($\Omega_0 \geq 1$). The same is obviously not true in the case of open universes ($\Omega_0 < 1$) where expansion inhibits collapse for small δ values.

An overdense perturbation growing in our Universe evolves, in its initial stages, to expand more slowly than the universe, thus increasing its density contrast. This expansion slows down more and more until a maximum radius r_{max} is reached. This moment is called *turn-around point*. At this time the motion reverses and the collapse phase begins, finally leading to structure formation and virialization. It can be shown that the density of the perturbation at the turn around point ($t = t_{max}$) is:

$$\rho_p = \frac{3\pi}{32 G t_{max}^2} \, . \tag{2.66}$$

Recalling the relationship for the background density (Tab. 1.1) of an EdS universe, we can calculate the density contrast at the turn-around point:

$$\delta(t_{max}) \simeq \frac{\rho_p(t_{max})}{\rho_0(t_{max})} - 1 = \left(\frac{3\pi}{4}\right)^2 - 1 \simeq 4.6 \, . \tag{2.67}$$

Therefore, at this stage the system is already largely in the nonlinear regime. If we had applied linear theory we would have obtained instead:

$$\delta_L(t_{max}) = \delta_L(t_i) \left(\frac{t_{max}}{t_i}\right)^{2/3} \simeq 1.06 \, . \tag{2.68}$$

The evolution following the turn-around should lead to the collapse of the system into a singularity at $t_{col} = 2t_{max}$. This does not happen due to the intrinsic angular momentum of the perturbation components and to microphysics: these phenomena lead to the virialization of the system. It can be seen from the hydrodynamic simulations that equilibrium is reached at $t_{vir} = 3t_{max}$. The radius at which the perturbation becomes stable is the virial radius R_{vir}.

Assuming that the system at the final stage is virialized, we can consider the relation

$$2\mathcal{T} + \mathcal{V} = 0 \, , \tag{2.69}$$

where $\mathcal{T}$ represent the *kinetic energy* and $\mathcal{V}$ the *potential energy* of the system. Considering the potential energy of a self-gravitating sphere of mass M:

$$\mathcal{V} = -\frac{3}{5}\frac{GM^2}{R} \, . \tag{2.70}$$

We know from the Virial Theorem (Eq. (2.69)) that when a system is virialized, $\mathcal{T} = -\mathcal{V}/2$. So, in this phase, the total energy is:

$$E = \mathcal{T} + \mathcal{V} = \frac{1}{2}\mathcal{V} = -\frac{3}{10}\frac{GM^2}{R} \, . \tag{2.71}$$

By assuming the conservation of mass and energy ($E(t_{max}) = E(t_{vir})$), we can write

$$\begin{aligned} E(t_{max}) = E(t_{vir}) \;\; \rightarrow \;\; \mathcal{T}(t_{max}) + \mathcal{V}(t_{max}) = \frac{1}{2}\mathcal{V}(t_{vir}) \;\; \rightarrow \\ \rightarrow \;\; 0 - \frac{3}{5}\frac{GM^2}{R_{max}} = -\frac{3}{10}\frac{GM^2}{R_{vir}} \rightarrow R_{vir} = \frac{1}{2}R_{max} \, . \end{aligned} \tag{2.72}$$

So, it follows that

$$\rho_p(t_{vir}) = 8\rho_p(t_{max}) \ . \tag{2.73}$$

We now have sufficient elements to calculate the value of the density contrast at times t_{coll} and t_{vir}:

$$\delta(t_{coll}) = \frac{8\rho_p(t_{max})}{\rho_0(t_{max})} \left(\frac{t_{coll}}{t_{max}}\right)^2 \simeq 180 \ , \tag{2.74}$$

$$\delta(t_{vir}) = \frac{8\rho_p(t_{max})}{\rho_0(t_{max})} \left(\frac{t_{vir}}{t_{max}}\right)^2 \simeq 400 \ . \tag{2.75}$$

Again, the same quantities can be calculated in linear theory for comparison:

$$\delta_L(t_{coll}) = 1.06 \left(\frac{t_{coll}}{t_{max}}\right)^{2/3} \simeq 1.69 \ , \tag{2.76}$$

$$\delta_L(t_{vir}) = 1.06 \left(\frac{t_{vir}}{t_{max}}\right)^{2/3} \simeq 2.2 \ . \tag{2.77}$$

The nonlinear theory thus predicts values very close to those measured, unlike the linear theory, which differs by two orders of magnitude from the actual data.

Underdensities

The evolution of subdensities is very different than that of their overdense counterparts. These regions, commonly called *voids*, exhibit direct radial acceleration of shells from the inner to the outer zones of the perturbation. The decrease in density is physically limited to the asymptotic reaching of $\delta = -1$.

Let us consider an inverse top-hat spherically symmetric underdense perturbation as a set of concentric shells with respective radii r_i. Knowing that each shell is in a newtonian regime, we can determine the acceleration through the mass contained in each of them:

$$\ddot{a} = \frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = -\frac{GM}{r} = -\frac{4\pi G}{3}\rho_0(1+\Delta)r \ . \tag{2.78}$$

At the initial time, t_i, we have:

$$M(t_i) = \frac{4\pi}{3}\rho_0 r(t_i)^3(1+\Delta(t_i)) \ , \tag{2.79}$$

and $\Delta(t_i)$ is the one from the Eq. (2.42) calculated at $t = t_i$, i.e. the average value of $\delta(t_i)$ within r_i. Equation (2.78) can be solved analytically, for an EdS universe, by parameterising the evolution of the density in the following way:

$$1 + \Delta(r,t) = \frac{\rho(r,t)}{\rho_0(r,t)} = \frac{9}{2}\frac{(\sinh\theta - \theta)^2}{(\cosh\theta - 1)^3} \ . \tag{2.80}$$

The initial *density deficit*, $\Delta(r,t)$, can also be derived in linear theory:

$$\Delta_i^L = -\left(\frac{3}{4}\right)^{2/3}\frac{3}{5}(\sinh\theta - \theta)^{2/3} \ . \tag{2.81}$$

The outflow of matter from the centre of the perturbation causes the density to decrease asymptotically to $\delta = -1$. During this process, the various shells experience a force that depends on their density. The shells that initially are located in the central zones of the voids are less dense than the outermost shells and therefore acquire a greater velocity. Thanks to this phenomenon, these shells can reach and overcome the outer ones: this process is called *shell crossing* and contributes to the tendency of matter in the universe to accumulate in filaments and sheets. We can think of the shell-crossing moment in an ideal void, i.e. spherical, isolated and without substructures, as that moment when the evolutionary regime changes from quasi-linear to midly-nonlinear. From the shell-crossing on, the evolution of the void can be described by a self-similar outward moving shell (Suto et al., 1984). We can show that, at the shell-crossing event, the void has a precisely determined excess Hubble expansion rate (Sheth & Van De Weygaert, 2004)

$$H_{sc} = \frac{4}{3} H(t_{sh}) \ , \tag{2.82}$$

where $H(t_{sc})$ is the Hubble parameter of the background Universe.

The value of the density threshold at which shell-crossing occurs can be calculated via Eq. (2.53) in which the parameter θ_{sc} is substituted:

$$1 + \delta_{v,sc}^{NL} \simeq 0.205 \ , \tag{2.83}$$

in which the notation "NL"means nonlinear. It is straightforward to observe, therefore, that voids are structures in a regime only moderately nonlinear. Furthermore, we can calculate the density threshold in linear theory, which results to be

$$\delta_{v,sc}^{L} \simeq -2.71 \ . \tag{2.84}$$

It is possible to calculate the expansion factor for the radius of a underdensity that has expanded until reaching the shell-crossing phase: $(1 + \delta_{v,sh}^{NL})^{-1/3} \simeq 1.697$ in comoving radius. Note that these numbers do not depend on the size of the void. We can therefore conclude that voids are subdense structures that evolve in a semi nonlinear manner tending to expand and assume a spherical shape, unlike overdensities that tend to evolve in a highly nonlinear way, collapsing and aggregating into sheets and filaments.

2.2.2 The Zel'dovich approximation

The transition between linear and nonlinear regime for density perturbations is described by the *Zel'dovich theory* (Zel'dovich 1970, see Shandarin & Zeldovich, 1989, for an exhaustive review). In particular, the *Zel'dovich approximation* relates the comoving coordinates $\vec{r} = \vec{x}/a(t)$, where $a(t)$ is the expansion factor, at a certain time t, to the Lagrangian coordinates $\vec{q}$ a $t \to 0$ through the following relation:

$$\vec{r}(\vec{q}, t) = a(t)[\vec{q} + b(t)\vec{s}(\vec{q})] \ . \tag{2.85}$$

The first term, $a(t)\vec{q}$, describes the cosmological expansion while the second term, $b(t)\vec{s}(\vec{q})$, describes the evolution of the perturbation. In particular, the factors $a(t)$ is, as usual, the cosmological expansion factor and $b(t)$ is the *growth rate* of linear

density fuctuation in the expanding universe. The factor $\vec{s}(\vec{q})$, the *initial displacement field*, Ψ, can be expressed as

$$\vec{s}(\vec{q}) \equiv \Psi(\vec{q}) = -\nabla_q \Phi(\vec{q}) \ . \tag{2.86}$$

In other terms, the Zel'dovich approximations assumes that the initial density perturbations are described by the gradient of a potential vector field (i.e. an initial force). Particles are not subject to additional interactions at later times and this implies that they can cross each other without causing any deviation in their motion.

The linear Poisson equation links the value of the density contrast to the gravitational potential through the relation

$$\delta = -D\nabla^2\Phi \ , \tag{2.87}$$

where $D = b(t)/a(t)$ is called the *linear growth function*. Thus, the divergence of the linear Zel'dovich displacement field can be related to the density contrast (see Falck et al., 2012, and references therein):

$$\nabla \cdot \Psi = -\delta \ . \tag{2.88}$$

This relation will be fundamental in this Thesis work as it will be the base for the reconstruction of the density field.

2.3 Halo bias

A fundamental problem in Cosmology is to understand how the spatial distribution of tracers relates to the distribution of DM. The most commonly used mass tracers in Cosmology are galaxies, galaxy clusters and DM haloes. We know that DM makes a significantly higher contribution to the total mass with respect to baryons. Furthermore, DM and baryons have a different evolution as they decouple from radiation at different times and the former is not affected by microphysics.

A density contrast field can be defined from mass tracer (subscript "tr") counts in a volume V:

$$\delta_{tr} = \frac{N_{tr}(V) - \overline{N}_{tr}(V)}{\overline{N}_{tr}(V)} \ , \tag{2.89}$$

where $N_{tr}(V)$ and $\overline{N}_{tr}(V)$ are the number of tracers and the mean number of tracers, respectively. The easiest way to parameterise the relationship between tracers in mass and the total distribution of matter in the Universe (subscript "m") is through the linear, local, non-stochastic bias model proposed by Kaiser (1984):

$$\delta_{tr} = b\delta_m \ , \tag{2.90}$$

where b is the *linear bias factor*, which depends on the cosmological scenario, redshift, and on tracer properties such as type, luminosity and mass. This model provides a description in linear regime at large scales. At small scales it loses validity due to the nonlinear effects caused by microphysics.

A simple and useful definition of bias is based on the 2PCF. Specifically, the bias value is calculated by the square root of the ratio between the 2PCF of the tracers and the 2PCF of the total matter component:

$$b = \sqrt{\frac{\xi_{tr}}{\xi_m}} \, . \tag{2.91}$$

ξ_{tr} has to be inferred from the distribution of tracers, while ξ_m usually can be derived analytically from the theory.

2.4 N-body simulations

Unlike many other fields of Physics, Cosmology is characterized by the non replicability of the studied phenomena, because of the unique nature of the Cosmos. An important method of overcoming this issue and testing cosmological models is through N-body simulations.

An N-body simulation is defined as a simulation capable of solving the N-body problem. In the simplest but also most common case, only gravitational attraction is considered. However, for more realistic results, the hydrodynamic effects resulting from the presence of the baryonic matter have also to be taken into account. Simulations in which also the baryonic component is evolved are called *hydrodynamic simulations*. Generally, the cosmological parameters of the Universe to be simulated are fixed, and an initial configuration of particles, tracing the total mass distribution, is created. This initial condition is allowed to evolve freely and in a nonlinear regime by solving the following system of equations:

$$\begin{cases} \mathbf{F}_i = GM_i \sum_{i \neq j} \frac{M_j}{r_{ij}^2} \hat{\mathbf{r}}_{ij}^2 \\ \ddot{\mathbf{x}}_i = \frac{d\mathbf{v}_i}{dt} = \frac{\mathbf{F}_i}{M_i} \\ \dot{\mathbf{x}}_i = \frac{d\mathbf{x}_i}{dt} = \mathbf{v}_i \end{cases} \tag{2.92}$$

In these equations, for each i-th particle, position $\mathbf{x}_i$, velocity $\mathbf{v}_i$, mass $\mathbf{M}_i$ and gravitational force $\mathbf{F}_i$ is computed. r_{ij} represents the distance between the i-th particle and the j-th particle, while $\hat{\mathbf{r}}_{ij}$ is the related versor. Given the system (2.92), the Euler equation can be formulated as

$$\frac{d\mathbf{x}_i}{dt} + 2\frac{\dot{a}}{a}\mathbf{v}_i = -\frac{1}{a^2}\nabla\Phi = -\frac{G}{a^3}\sum_{i,j\neq i} M_j \frac{\mathbf{x}_i - \mathbf{x}_j}{|\mathbf{x}_i - \mathbf{x}_j|^3} = \frac{\mathbf{F}_i}{a^3} \, , \tag{2.93}$$

where a is the scale factor. Applying the second Friedman equation reported in (1.26), the Poisson Equation becomes:

$$\nabla^2\Phi = 4\pi G\bar{\rho}(t)a^2\delta = \frac{3}{2}H_0^2\Omega_0\frac{\delta}{a} \, , \tag{2.94}$$

where $\bar{\rho}(t)$ is the average non-relativistic matter density, δ the local density contrast, H_0 is the Hubble parameter and Ω_0 the non-relativistic matter density parameter.

N-body simulations, like all numerical simulations, are subject to the discretization of quantities. The ones included in the system (2.92) are therefore calculated over discrete time steps dt. There are several methods for calculating the force $\mathbf{F}_i$ acting on the i-th particle, which differ in CPU efficiency, implementation simplicity, mass and spatial resolution:

- **Particle-Particle (PP)**. The force acting on the i-th particle is calculated directly by adding up the contributions of all the other particles in the simulation. This method is the most accurate, as it returns the exact solution of the value of the forces, but it is also extremely expensive in terms of calculation resources, as it scales as N^2 (where N is the number of particles). Moreover, if two particles move too close together, the force diverges to infinity, so a force *softening* parameter is introducted.

- **Particle Mesh (PM)**. All quantities which can be assumed to be field quantities are treated as such. Potential, gravitational force and density are calculated by interpolating the quantities on the grid. For each point, the values are calculated as a weighted sum using a weight function called *kernel*, which can be of different types. Using a regular grid, equations can be simplified by a transformation in Fourier space. Even though this may be the fastest of the possible methods, its resolution is limited by the smothing scale of the grid.

- **Tree code (HT)**. These methods make use of so-called barycentric codes and the graph theory. A hierarchical tree is constructed: each region is decomposed into sub-regions until it contains at most one particle. At each level, each region is characterised by the total mass of the hosted particles and the position of their centre of mass. Arbitrarily distant regions from the particle under investigation are considered as a single particle with the abovementioned characteristics. In this way, the number of operations required for calculating the force acting on each particle is reduced. This method is the most widely used to date. In its standard version it scales efficiently ($N \log N$) and is easy to linearise, but requires a great amount of memory to store all the levels of the hierarchical tree.

The first studies involving numerical simulations (Von Hoerner 1960, Aarseth & Hoyle 1963) were based on solving the N-body problem for a few hundred particles. Thanks to the enormous progress in technology and computational techniques seen in recent decades, nowadays it is possible to run simulations with a number of particles on the order of trillions (Potter et al., 2017, an example can be seen in Fig. 2.1. In particular, the image shows a detail of the Uchuu simulation, to date the biggest cosmological simulation with 2.1 trillion of DM particles). Despite the incredible successes of this branch of research, some strong limitations in cosmological simulations still exist today, mainly due to the discretization of quantities and the lack of resolution caused by the computational limitations.

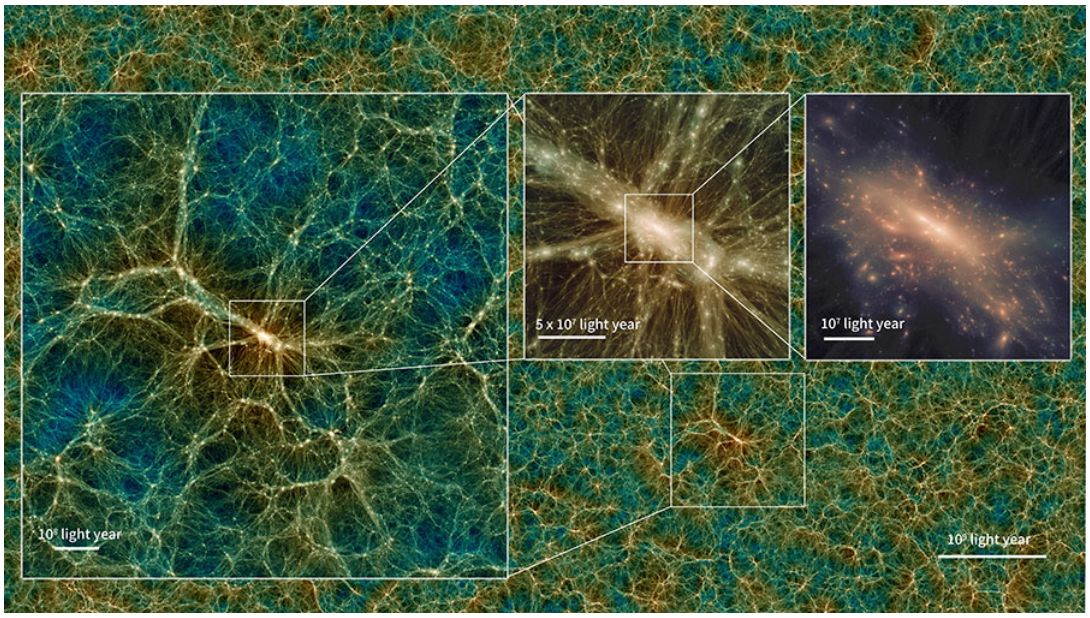

Figure 2.1: Distribution of DM in a Uchuu simulation snapshot, to date the cosmological simulation with the largest number of DM particles (2.1 trillion in a 2000 h^{-1}Mpc box) present in literature. The images show, in particular, the DM halo of the largest galaxy cluster formed in the simulation, at different magnifications. Credits: Ishiyama et al. (2021).

Chapter 3

Looking for cosmic voids

We can refer to cosmic voids as those wide, underdense structures that fill most of the volume of the Universe, setting between the filaments of the cosmic web. As seen in the previous chapter, voids are generated by the evolution of subdense primordial perturbations. In Fig. 3.1, we present the simulated evolution of a cosmic void from the primordial perturbation field to present time. During this evolution, cosmic voids shift from linear to mildly nonlinear regimes, increasing their dimension, becoming more spherical and moving mass from the central zones to the edges. In this way, cosmic voids contribute to the formation of the cosmic web.

Although the evolution history and the morphology of voids are well known, a widespread and unique definition of these objects has not been agreed yet. Despite this issue, voids provide powerful cosmological probes through number counts (Sec. 3.2.2) and density profiles (Sec. 3.2.3), analogously to their positive counterparts in the density field, i.e. galaxy clusters. These statistics are very important to constrain cosmological parameters and in particular to investigate the nature of dark energy and test cosmological models.

3.1 Void finding

Identification of voids is also non trivial. It is necessary to reconstruct their shape and identify the position of their centres from the positions of the luminous tracers, which are mostly arranged along the borders. Over time, several identification methods have been proposed by the scientific community. Following the scheme discussed by Lavaux & Wandelt (2009), we can classify these methods into three different classes, according to the detection criteria: *density*, *geometrical* and *dynamical*.

3.1.1 Density-based methods

The simplest and straightforward method for detecting voids consists in identifying empty spherical regions, or at most regions with densities below a certain fraction of the mean cosmic density (Kauffmann & Fairall 1991, Hoyle & Vogeley 2004, Foster & Nelson 2009, Elyiv et al. 2013, Micheletti et al. 2014). One of the most remarkable applications of this method can be found in the context of the VIMOS*Public Extragalactic Redshift Survey* (VIPERS, Garilli et al. 2014). VIPERS was an ESO

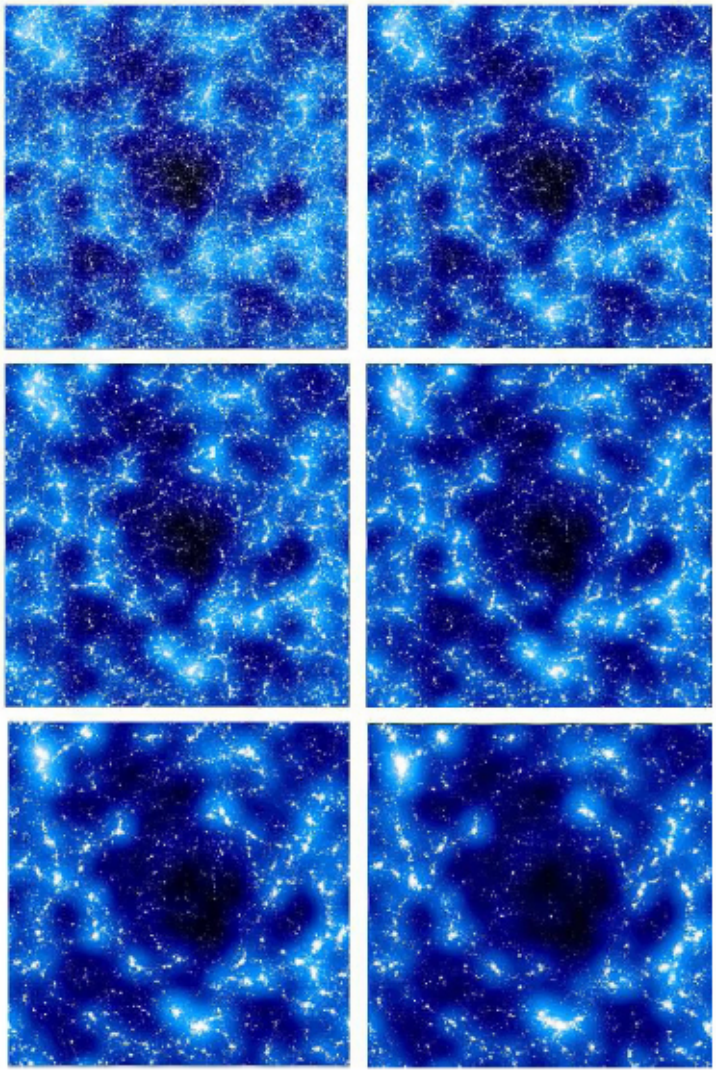

Figure 3.1: Simulation of an evolving void in a standard ΛCDM scenario. The slice is 50 h^{-1} Mpc wide and 10 h^{-1} Mpc thick. Particles and smoothed density field (smoothed on a scale of 4 h^{-1} Mpc) are shown at six different timesteps: $a = 0.05,\ 0.15,\ 0.35,\ 0.55,\ 0.75$ and 1.0, from top left to bottom right Credits: Van de Weygaert & Platen (2011).

programme with the purpose of examining the large scale structure (LSS) distribution of galaxies at $z = 0.5 - 1.2$, over total sky area of 24 deg^2. Micheletti et al. (2014) developed an algorithm for the detection of voids based on the identification of empty spheres. Voids found in this way are thus defined as regions devoid of galaxies with absolute magnitudes brighter than a specified limit (B-band brighter than $M_B = -19.8$). The method is based on the identification of empty spheres within a catalogue cleaned of isolated galaxies, i.e. galaxies located in underdense areas. This is done by calculating the distance between points of an equispaced grid and the nearest galaxy to each of these points. Then, the volumes of each sphere are calculated and the overlaps between voids removed. The main disadvantage of this method is that a single sphere will not be sufficient to reconstruct the real shape of the examined empty regions. Simply, more than one sphere will be found inside the same cosmic void. Furthermore, a simple analysis may be performed to define a posteriori the volumes that correspond to topologically connected spheres satisfying a specific density cut-off. Figure 3.2 shows the results of the application of the void detection algorithm on the VIPERS galaxies, which leads to a sample of voids, identified as maximal empty spheres.

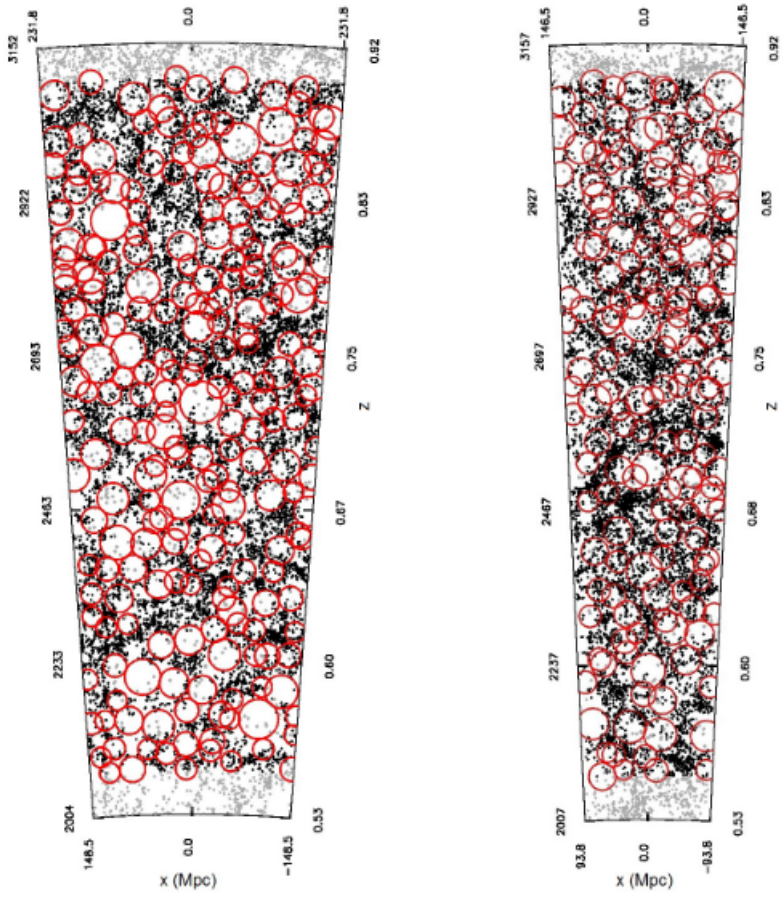

Figure 3.2: Maximal spheres (red circles) in the two different VIPERS samples; grey points are galaxies detected as isolated and galaxies outside the sample redshift range. The scales show comoving distance in Mpc and the corresponding redshift. Credits: Micheletti et al. (2014).

3.1.2 Geometrical-based methods

The class of geometric void finders bases the identification of voids on the reconstruction of the density field through the subdivision of the volume in cells (which can take various geometric forms). In each of these cells, the density value is evaluated. Voids are reconstructed as groups of underdense cells and their centres identified as local minima of the density field (Platen et al. 2007, Neyrinck 2008, Sutter et al. 2014a).

At the present day, the most popular algorithm exploiting this method is the *Void IDentification and Examination toolkit* (`VIDE`, Sutter et al. 2014a). `VIDE` is based on an enhanced version of the *ZOnes Bordering On Voidness* (`ZOBOV`, Neyrinck 2008) algorithm, a parameter-free method that does not require any assumptions about the void shape. Specifically designed to search for underdense zones, it is based on the original *VOronoi BOund Zones* (`VOBOZ`, Neyrinck et al. 2005) method, designed instead for the detection of overdensities.

The `VIDE` void finding procedure consists of three main steps:

1. As a first step, a *Voronoi tessellation* is performed on the tracer catalogue. This process allows the division of the volume into cells, called *Voronoi cells*, containing always only a single particle. Furthermore, each point within each individual cell is closer to the particle contained therein than to any other particle. At this point, assuming equal mass for each particle, the algorithm computes the density mean value of each cell, simply by calculating the inverse of the Voronoi cell volume. A continuous density field is thus obtained. An example of two-dimensional voronoi tessellation applied to a catalogue of galaxies, extracted from the work of Neyrinck (2008), is shown in Fig. 3.3.

2. The second step consists in identifying the local minima of the density field, represented by cells surrounded by neighbours with higher densities. Once this process has been accomplished, local underdensity basins, known as *zones*, are created. These basins are formed by merging cells of densities that are gradually higher than the previous ones. The process stops when a cell of lower density is encountered.

3. Finally, as last step, voids are shaped by joining the previously formed areas. This is performed through a process called *watershed*, devised by Platen et al. (2007). The word watershed refers to the analogy of a landscape being flooded by a rising level of water. Suppose we have a surface in the shape of a landscape (first image of Fig. 3.4). The surface is initially flooded at the location of each minima. As the water level rises, a growing fraction of the landscape will be flooded by the water in the expanding basins. Ultimately basins will meet at the ridges corresponding to saddle points in the density field. This intermediate step is plotted in the second image of Fig. 3.4. The ridges define the boundaries of the basins, enforced by means of a sufficiently high dam. The final result (see third image in Fig. 3.4) of the completely immersed landscape is a division of the landscape into individual cells, separated by the ridge dams. Furthermore, `VIDE` provides a void hierarchy, by using the identification of the various basins and ridges. In this substructure partitioning, a *parent void*

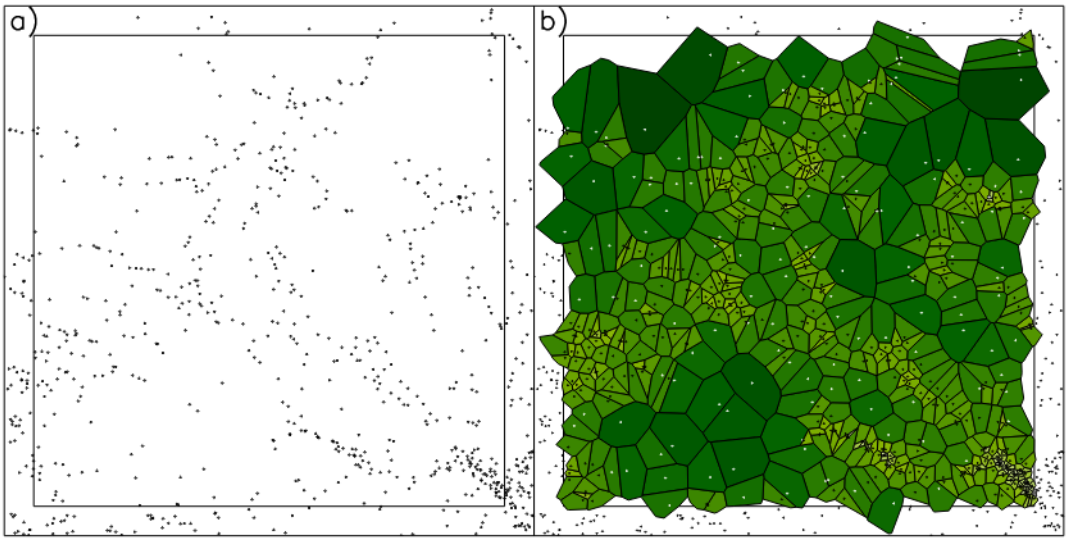

Figure 3.3: **(a)** Galaxies from a $40\times40\times5\,(h^{-1}\mathrm{Mpc})^3$ slice of the *Aspen-Amsterdam Void-Finder Comparison Project* (Colberg et al., 2008) region. The outer boundary is a $45\,h^{-1}$ Mpc square. **(b)** The 2D Voronoi tessellation of galaxies in this slice, with each particle Voronoi cell shaded according to its area. The galaxies outside the inner ($40\,h^{-1}$ Mpc) boundary are shown because they contribute to the tessellation. Credits: Neyrinck (2008).

holds a multitude of subvoids. Each void derives from one parent and could have several *child subvoids*, organized into different levels.

After this process, the centre of the voids identified by `VIDE` is redefined as the barycentre, $\vec{X}$, of the Voronoi cells constituting each void, weighted over the volume of each individual cell:

$$\vec{X} = \frac{\sum_i \vec{x}_i V_i}{\sum_i V_i} \,, \tag{3.1}$$

where $\vec{x}_i$ and V_i are the positions and Voronoi volumes of each tracer particle i, respectively. Finally, the effective radius, r_v, is calculated from the total volume of the Voronoi cells that constitute each void, $V_\mathrm{v} = \sum_i V_i$. The latter is defined as the radius of a sphere having volume V_v.

3.1.3 Dynamical-based methods

This third and final class of finders is very different from the previous two. Methods based on dynamic criteria do not use tracers to reconstruct the density distribution, but use them as test particles to sample the *velocity field* (Forero-Romero et al. 2008, Lavaux & Wandelt 2009, Elyiv et al. 2015). In these algorithms, voids are identified as regions from which there is an outflow of tracers, rather than as underdense zones. The dynamic void identification has the advantage of partially overcoming the problems introduced by the shot noise, i.e. by the sparsity of tracers. Moreover, it permits the actual reconstruction of continuous velocity and density fields.

Two of the most recent examples of this algorithm type are those presented by Elyiv et al. (2015). Both algorithms, the *Uncorrelating Void Finder* (`UVF`) and the

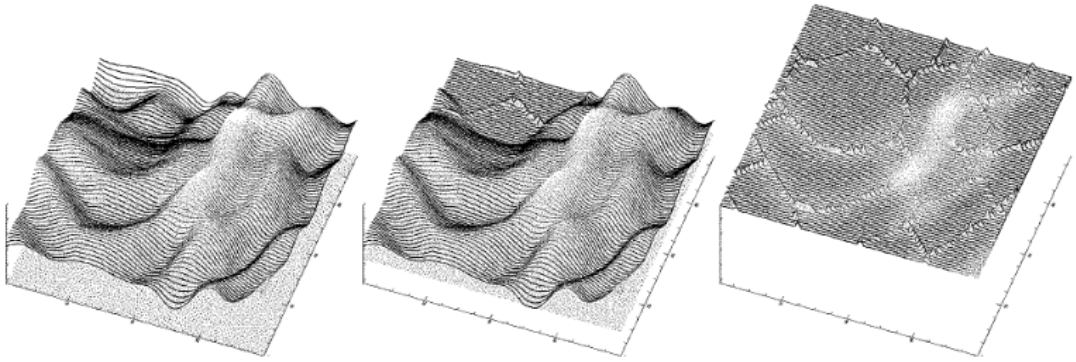

Figure 3.4: Three frames illustrating the principle of the watershed. The left-hand frame shows the surface to be segmented. Starting from the local minima the surrounding basins of the surface start to flood as the water level continues to rise (dotted plane initially below the surface). Where two basins meet up near a ridge of the density surface, a "dam" is erected (central frame). Ultimately, the entire surface is flooded, leaving a network of dams defines a segmented volume and delineates the corresponding cosmic web (right-hand frame). Credits: Platen et al. (2007).

Lagrangian Zel'dovich Void Finder (`LZVF`), can reconstruct the Lagrangian positions of galaxies, through the randomisation of the Eulerian ones. The basic idea behind the two dynamical void finders is rather simple. Let us consider a volume of the Universe characterized by large-scale structures in the DM component and probed by "particles" like haloes or galaxies. The goal is to use these as test particles and trace their orbits back in time to a homogeneous and isotropic initial distribution, i.e. to reconstruct their Lagrangian positions $\vec{q}$. This is done in two ways:

- `UVF`. The reconstruction is performed by exploiting the different correlation properties of the initial and final particle distributions. At the present epoch the spatial distribution of galaxies is highly clustered. Deviations from the homogeneity are conveniently and readily characterized by their 2PCF. On the other hand, at early epochs the distribution of matter is supposed to be highly homogeneous, with no spatial correlation. This suggests that a practical way to trace galaxy orbits back in time, at least in a statistical sense, is by relaxing their present spatial distribution to homogeneity, practically defined as a state in which the correlation function at all separations is zero. `UVF` uses an algorithm based on the Rintoul & Torquato (1997) method, effectively moving the system away from its minimum energy configuration;

- `LZVF`. The reconstruction, in this case, is based on the Zel'dovich approximation (see Eq. (2.2.2)) to the growth of density fluctuations. More specifically, the algorithm exploits the *path interchange Zel'dovich approximation* method of Croft & Gaztanaga (1997) to trace the orbit of the objects in a self-gravitating system by minimizing its action. Since objects have straight orbits, it is possible to simply connect their Eulerian positions to those of a randomly distributed sample. In each iteration performed by the algorithm, the connections are modified, hence setting new paths to different grid points, and accepting the result if the total path decreases. Since the total path is

proportional to the action, this system is relaxed to the correct dynamical configuration.

After reconstructing the Lagrangian positions $\vec{q}$, the *particle displacement field*, $\vec{\Psi}$ is obtained, by simply connecting them to the Eulerian positions, $\vec{v}$, hence assuming straight orbits: $\vec{q} - \vec{v} = \vec{\Psi}(q)$. The divergence of the displacement field is associated with the mass density field through the relation (2.88) and used to identify voids as sinks of mass streamlines in time reverse variables. The final void catalogue provides the comoving coordinates of the void centres, found as the position of a local minimum of the divergence field, and the related effective radii. An application of the two treated finders is shown in the Fig. 3.5. The top panels show the reconstruction of the displacement field, while the bottom panels show the position of the tracers and the divergence field obtained, which can be assumed to be the density contrast field. As can be seen, the arrows showing the back-in-time trajectories consistently point towards the centre of the underdense areas.

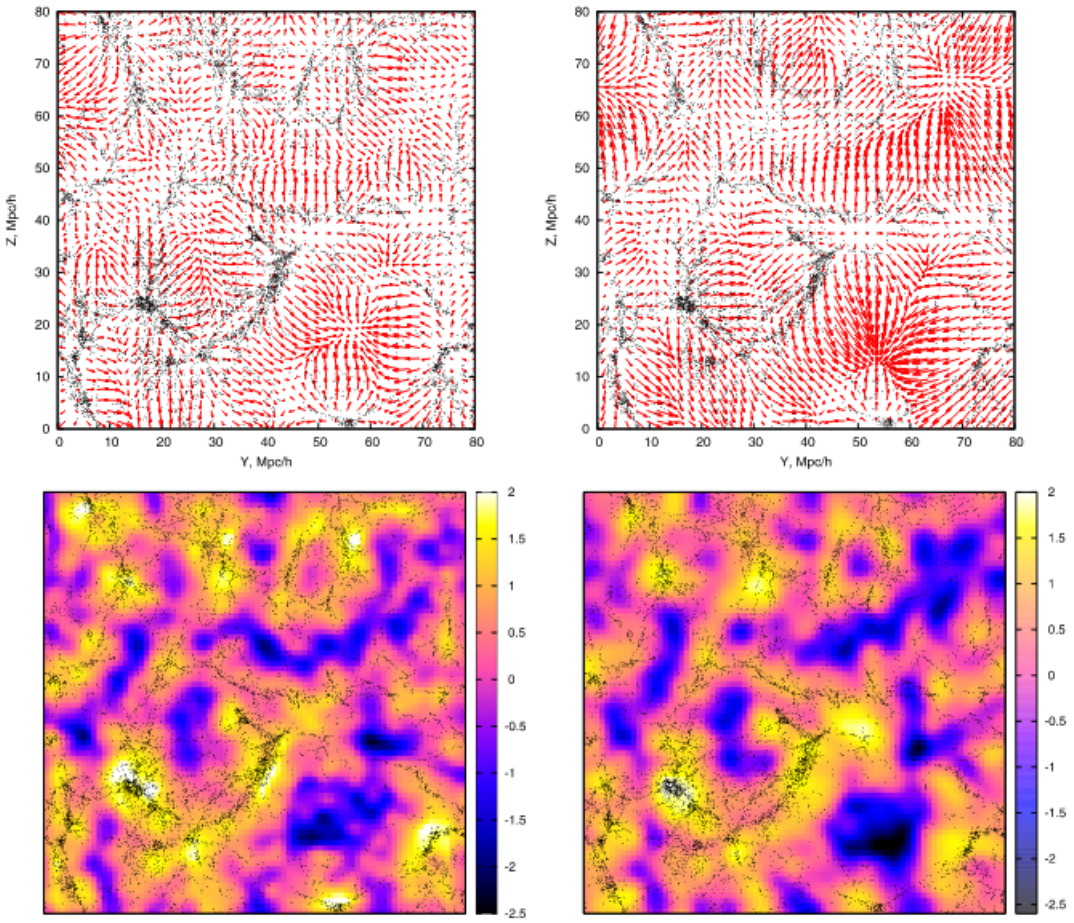

Figure 3.5: The reconstructed displacement field (top panels) and its divergence (bottom panels) obtained with the two void finders, the `UVF` (left-hand panels) and the `LZVF` (right-hand panels). The size of the displayed region is 80×80 h^{-1} Mpc, with a thickness of 5 h^{-1} Mpc. Black dots represent DM haloes. Credits: Elyiv et al. (2015).

3.2 Cosmic voids statistics

In the recent years, cosmic voids have emerged as an increasingly valuable cosmological probe. The development of large galaxy surveys has allowed us, and will increasingly do it in the future, to obtain information on a large number of tracers, distributed over wide volumes. Both of these characteristics are crucial for void statistics: wide volumes provide a significant number count even for the largest voids, while a high number of tracers allows to improve the spatial resolution and thus to study the smallest underdensities and the substructures of these systems.

The vast spatial extension of cosmic voids, in combination with their characteristic underdense nature, makes these objects the ideal environment for testing a wide variety of astrophysical and cosmological models. The quasi-absence of matter in the centre of the voids makes these regions particularly suitable for the study of *DE* and *modified gravity* models. For example, it is expected that possible deviations from general relativity will modify the density profiles (Cai et al. 2015, Barreira et al. 2015, Perico et al. 2019, Contarini et al. 2021). Furthermore, the shapes (Lee & Park 2009, Lavaux & Wandelt 2012) and sizes (Bos et al. 2012, Pisani et al. 2015, Verza et al. 2019, Contarini et al. 2022) of voids are sensitive to the DE equation of state. Voids are also particularly sensitive to the presence of *massive neutrinos*: the density fraction of neutrinos is indeed more prominent in underdense regions. Moreover, the typical void size spans the range of neutrino free-streaming scale, which depends on the neutrino mass. Indeed, both void density profiles and void abundances have been shown to have a great potential in constraining the total mass of neutrino species (Massara et al. 2015, Kreisch et al. 2021, Contarini et al. 2021, Contarini et al. 2022).

Another key feature of cosmic voids is their *complementarity* with standard overdense probes (Nadathur et al. 2020, Kreisch et al. 2021, Woodfinden et al. 2022, Contarini et al. 2022). The parallel and distinct evolution of subdensities with respect to their overdense counterparts has the effect of providing statistics independent of those of standard cosmic probes, resulting in constraints that are highly orthogonal to those provided by other cosmic structures. This feature is crucial, since with the joint analysis of cosmic voids and other probes at the same time, it is possible to resolve degeneracies in the estimates of cosmological parameters.

The most important cosmological contribution of void statistics comes from *number counts* and *density profiles*. In particular, concerning the first statistics, we underline that it is possible to predict the void counts as a function of their radius through the application of the so-called *excursion set formalism*, as we will see in the following section.

3.2.1 Excursion set formalism

The *excursion-set formalism* is an analytical framework to study the large scale structure of the Universe. This approach allows us to predict the number density of cosmic objects by relating the cosmological linear perturbation theory to its nonlinear counterpart at late time (Bond et al., 1991). Moreover, it is well known that, in combination with the spherical collapse, this approach provides insights into many aspects of halo formation and can be used to predict DM halo abundances and

clustering (see Zentner 2007, Jennings et al., 2013, and references therein). The excursion set formalism at its heart relies on the knowledge of the statistical properties of the linear density field. The linear density fluctuation field smoothed on a scale R, in Fourier space, is given by

$$\delta(\vec{x}, R) = \frac{1}{(2\pi)^3} \int \delta(\vec{k}) W(\vec{k}, R) e^{-i\vec{k}\cdot\vec{x}} d^3k \ , \tag{3.2}$$

where $\delta(\vec{k})$ is the Fourier transform of the density perturbation and $W(\vec{k}, R)$ is a filter function in Fourier space. The smoothing scale R is related to the corresponding variance of the linear density field:

$$\sigma^2(R) \equiv S(R) = \frac{1}{2\pi^2} \int k^2 P(k) |W(\vec{k}, R)|^2 dk \ , \tag{3.3}$$

where $P(k)$ is the matter power spectrum in linear theory. We can refer to a trajectory $\delta(x, S)$ as a sequence of overdensities given by successive increments ΔS in the smoothing scale. When a top-hat filter in k-space is used, then $\delta(x, S)$ executes a random walk. Given an underlying Gaussian distribution for the linear density field, the excursion-set formalism allows us to associate probabilities to random walks that satisfy a given set of criteria for the smoothing scale at which they cross various density thresholds.

The spherical evolution model, in combination with the excursion set, provides a good description of the statistics of DM halo for the collapse of perturbations. We can then use the excursion-set formalism to determine the fraction of trajectories that cross for the first time the barrier δ_c^L, i.e. the critical value for linear density fluctuation[1] (see Sec. 2.2.1), accounting also for the *cloud-in-cloud process*. The cloud-in-cloud process occurs when, during the formation of a structure, the trajectories cross the δ_c^L threshold several times. From a physical point of view, this occurs when one halo, when collapsing, envelops another halo. It is necessary to consider as halo only those objects that are not contained in larger objects, therefore considering only the smallest of the possible $\sigma(M)$ values measured at the various threshold crossings.

We can extend the model to predict the evolution of underdense regions in the initial density field. These are naturally associated with voids in the present day evolved density field. A key assumption in making the connection between the excursion set and the abundance of nonlinear objects is that each collapse occurs in isolation. This makes sense for collapsing objects since the comoving volume occupied shrinks. In contrast to the overdense regions which contract, voids expand. We shall see that this causes a problem for mapping the excursion-set predictions onto the statistics of voids. Nonetheless let us start with the simple spherical evolution model following Sheth & Van De Weygaert (2004). The critical density threshold is defined to be the shell-crossing density, $\delta_v = -2.71$. We can follow the excursion-set formalism to determine the fraction of random walks that pierce the barrier δ_v. Similar to the cloud-in-cloud process, the *void-in-void process* accounts for the fact that a void of a given size may be embedded in another underdense region on a larger

[1] This value lies in the range $\delta_c^L = [1.06, 1.69]$, density contrast of turn-around point and complete halo collapse, respectively.

scale. We thus define the first crossing distribution by associating the random walks with the smoothing scale for which they first cross the barrier $\delta_{\rm v}$. A second process, the *void-in-cloud scenario*, occurs when a void of a given size is embedded in an overdense region on a larger scale, which will eventually collapse to a halo and squash the void out of existence. Finally, the situation in which a large-underdense region embeds a small-overdense one, known as *cloud-in-void*, is irrelevant for the formation of collapsed structures, because clouds which condense in a large scale void are not torn apart as their parent void expands around them (Van de Weygaert, 2014). This asymmetry between the void-in-cloud and cloud-in-void processes leads to a symmetry breaking between the halo and void populations: they evolve out of the same symmetric Gaussian initial conditions, Though overdensities and underdensities are expected to evolve towards a distribution with different characteristics. In order to account for this asymmetry, Sheth & Van De Weygaert (2004) proposed that the excursion-set method applied to voids requires a second barrier, the threshold for collapse of overdense regions, $\delta_{\rm c}$. Therefore, when calculating the first crossing distribution, we need to determine the largest scale at which a trajectory crosses the barrier $\delta_{\rm v}$, given that it has not crossed $\delta_{\rm c}$ on any larger scale.

We show in Fig. 3.6 a summary of the four processes described by the excursion-set formalism. This approach provides the basics to the modelling of the theoretical void size function, which we describe in the next section.

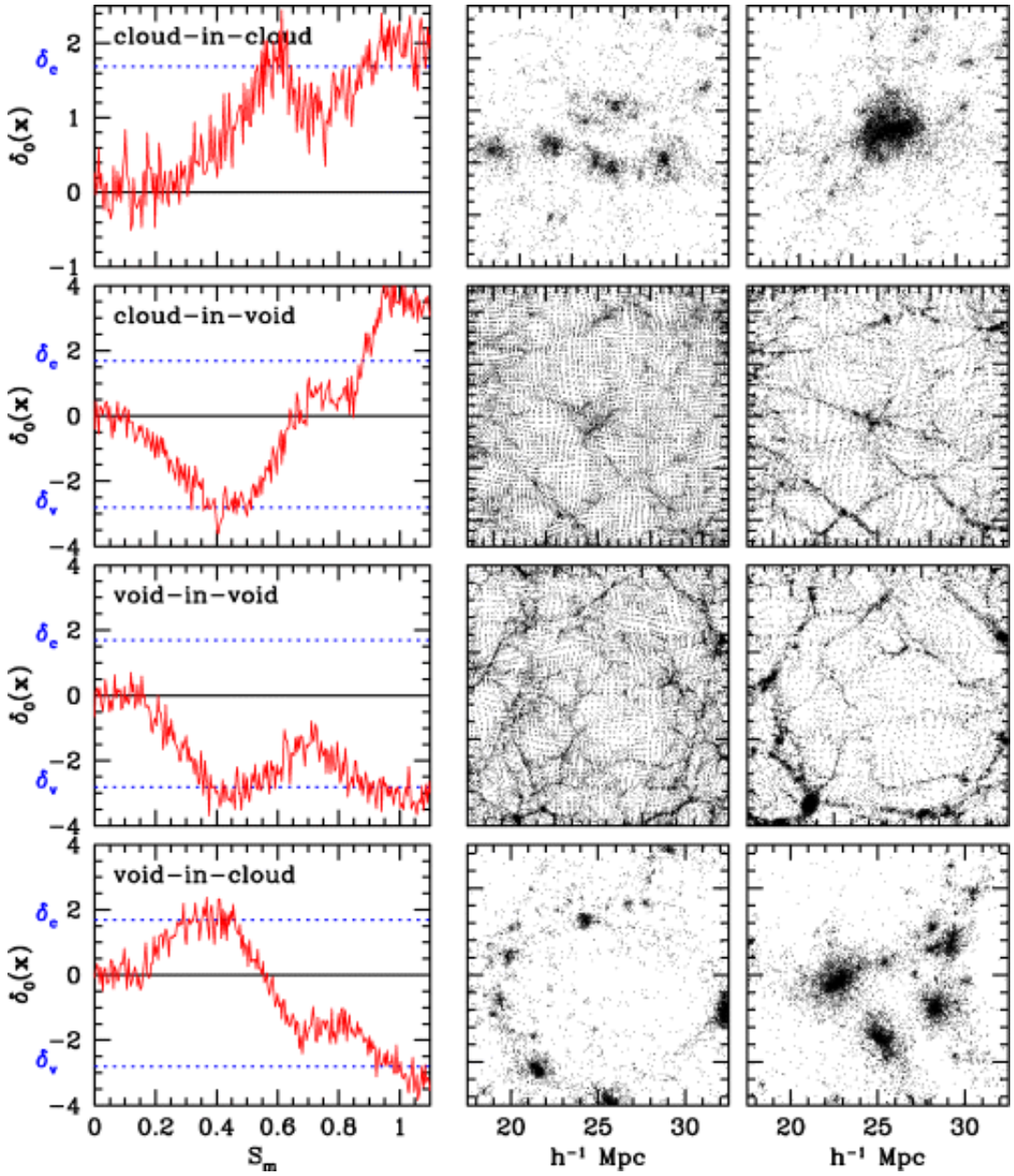

Figure 3.6: Four modes of the excursion-set formalism. Each row illustrates one of the four basic modes of hierarchical clustering: the cloud-in-cloud process, cloud-in-void process, void-in-void process and void-in-cloud process (from top to bottom). Each mode is illustrated using three frames. The leftmost panels show *random walks*: the local density perturbation $\delta(x)$ as a function of (mass) resolution scale S_m. In each graph, the dashed horizontal lines indicate the collapse barrier δ_c and the shell-crossing void barrier δ_v. The two frames on the right show how the associated particle distribution evolves, to an earlier (second column, central panels) and a later time (third column, rightmost panels). Credits: Sheth & Van De Weygaert (2004).

3.2.2 Void size function

Similarly to what happens with halo, the void overdense counterparts, a zero-order statistic can be defined for voids through object counts. However, while for overdensities the number counts are measured as a function of mass, obtaining, for example, the *halo mass function*, void counts are measured as a function of radius. Specifically, we define the *void size function*, VSF, as the comoving number density of cosmic voids as a function of their effective radii.

Sheth and van de Weygaert model

The VSF has been modelled for the first time by Sheth & Van De Weygaert (2004, hereafter the SvdW model), with the same excursion-set approach used to model the mass function of DM haloes (Press & Schechter 1974, Bond et al. 1991). As seen previously, the distribution of fluctuations that become voids with their evolution is obtained as the conditional first crossing distribution of the matter density contrast filtered at decreasing Lagrangian radius in a double barrier problem. A fluctuation can only be regarded as a void of radius $R_{\rm v}$ if the density contrast filtered on this scale has exceeded their threshold for void formation, $\delta_{\rm v}^{\rm L}$, without having exceeded the threshold for collapse $\delta_{\rm c}^{\rm L}$ at any larger scale. The assumption on which the model is based, i.e. spherical voids, allows us to overcome the problems associated with the fact that Sheth & Van De Weygaert (2004) multiplicity function is derived for spherical fluctuations in Lagrangian space. The initial density field evolved linearly at the epoch of interest, while the observed voids live in Eulerian space, in other words in the completely nonlinearly evolved density field.

The excursion-set theory applied to underdensities predicts a fraction of the Universe occupied by cosmic voids given by the *multiplicity function* $f_{\ln\sigma}$:

$$f_{\ln\sigma} = 2\sum_{j=1}^{\infty} j\pi x^2 \sin(j\pi\mathcal{D}) \exp\left[-\frac{j\pi x^2}{2}\right] , \tag{3.4}$$

where

$$\mathcal{D} \equiv \frac{|\delta_v^L|}{\delta_c^L + |\delta_v^L|} , \tag{3.5}$$

$$x \equiv \frac{\mathcal{D}}{|\delta_v^L|}\sigma . \tag{3.6}$$

The δ_c^L value is expected to be between $1.06 \leq \delta_c^L \leq 1.69$, i.e. a value between the DM turn-around and collapse density contrasts in linear theory, as both can be assumed as acceptable values. It is finally possible to derive the number density distribution of voids as a function of their size in linear theory, by exploiting the assumption that when moving from the linear to the nonlinear regime, the total number of voids is preserved:

$$\frac{\mathrm{d}n^L}{\mathrm{d}\ln r^L} = \frac{f_{\ln\sigma}(\sigma)}{V(r^L)}\frac{\mathrm{d}\ln\sigma^{-1}}{\mathrm{d}\ln r^L} \tag{3.7}$$

where $V(r^L) = \frac{4}{3}\pi(r^L)^3$ is the volume of the spherical fluctuation of radius r^L.

While reaching shell-crossing, underdensities are expected to have expanded by $a \propto (\delta_v^L)^{-3}$. A correction in radius is required. The void abundance therefore becomes:

$$\frac{\mathrm{d}n}{\mathrm{d}\ln r}\Big|_{\mathrm{SvdW}} = \frac{\mathrm{d}n}{\mathrm{d}\ln(ar^L)} \,. \tag{3.8}$$

However, Jennings et al. (2013) argued that this model is unphysical, as the conservation of the number of voids may not be valid, expecially for large voids, given their natural tendency to expand and overlap. Indeed, it is possible to calculate THE fraction of the total volume of the universe occupied by voids as:

$$\mathcal{F}(R) = \int_R^\infty V(r) \frac{\mathrm{d}n}{\mathrm{d}\ln r} \frac{\mathrm{d}r}{r} \,, \tag{3.9}$$

which leads to a $\mathcal{F}$ value greater than one.

Volume conserving model

In order to overcome the total void volume issues of the SvdW model, Jennings et al. (2013) proposed the so-called *volume conserving model* (Vdn, hereafter). This model is no longer based on the conservation of the total number of voids but assumes that the total volume of voids is conserved during the transition between linear and nonlinear regime. Specifically, if we define the *volume fraction in linear theory*, $\mathcal{F}_L$, as

$$\mathcal{F}_L(R) = \int_{R^L}^\infty V(r^L) \frac{\mathrm{d}n}{\mathrm{d}\ln r^L} \frac{\mathrm{d}r^L}{r^L} \,, \tag{3.10}$$

then this fraction is conserved if we define the nonlinear abundance as

$$V(r)\mathrm{d}n = V(r^L)\mathrm{d}n^L\big|_{r^L(r)} \,. \tag{3.11}$$

Therefore the void abundance becomes:

$$\frac{\mathrm{d}n}{\mathrm{d}\ln r}\Big|_{Vdn} = \frac{V(r^L)}{V(r)} \frac{\mathrm{d}n}{\mathrm{d}\ln r^L} \frac{\mathrm{d}\ln r^L}{\mathrm{d}\ln r}\Big|_{r^L(r)} \,. \tag{3.12}$$

Figure 3.7 shows a comparison of the volume fractions of the Universe occupied by the voids predicted for the various models considered. For the SvdW model, the fraction unphysically exceeds unity, while the Vdn model conserves the total fraction from the linear theory. Figure 3.8, on the other hand, shows the predicted VSF and the corrisponding void abundance measured in simulations. Void counts are in excellent agreement with the Vdn model, but not with the SvdW model, which systematically overestimates abundances.

Extension of the volume conserving model for biased tracers

Although the Vdn model provides an accurate prediction of void statistics, this is limited to voids evolving in the total density field. The distribution of matter in the Universe is represented in a first approximation by the DM distribution. By measuring voids through the density distributions of biased tracers, such as galaxies or DM haloes, the theory would fail to correctly predict the result of the

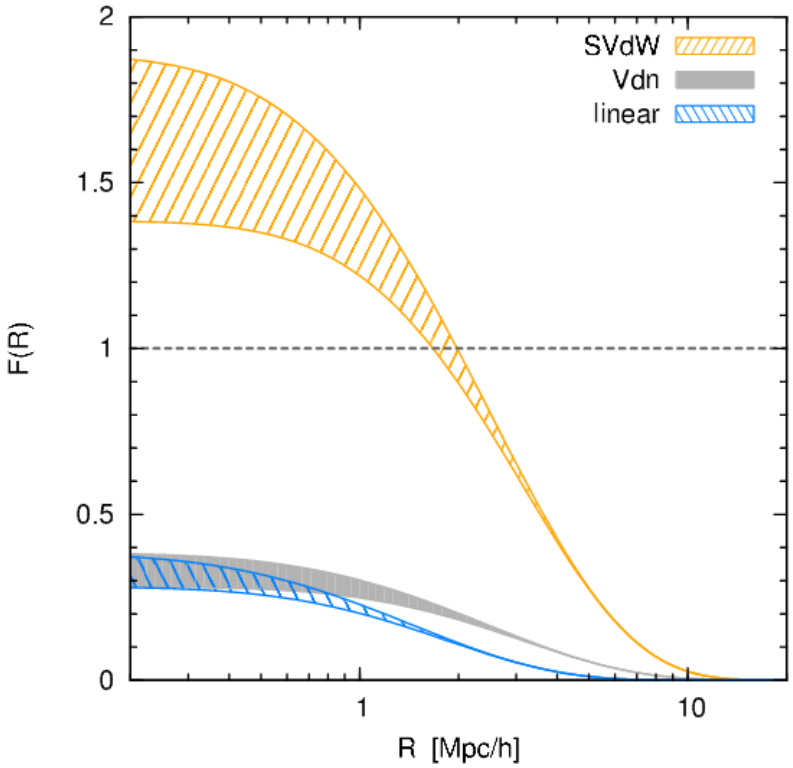

Figure 3.7: Cumulative volume fraction occupied by voids with radii larger than R for these theoretical models: linear theory (blue striped region, $R = r^L$), SvdW model (orange striped region, $R = r$) and Vdn model (grey shaded region, $R = r$). These regions correspond to the expected range of $1.06 \leq \delta_c \leq 1.69$ and with $\delta_v = -2.7$ throughout. For SvdW the fraction unphysically exceeds unity at $R \sim 2\,h^{-1}$ Mpc while for the Vdn model conserves the total fraction from the linear theory of $\mathcal{F}(0) \sim 0.3$. Credits: Jennings et al. (2013)

.

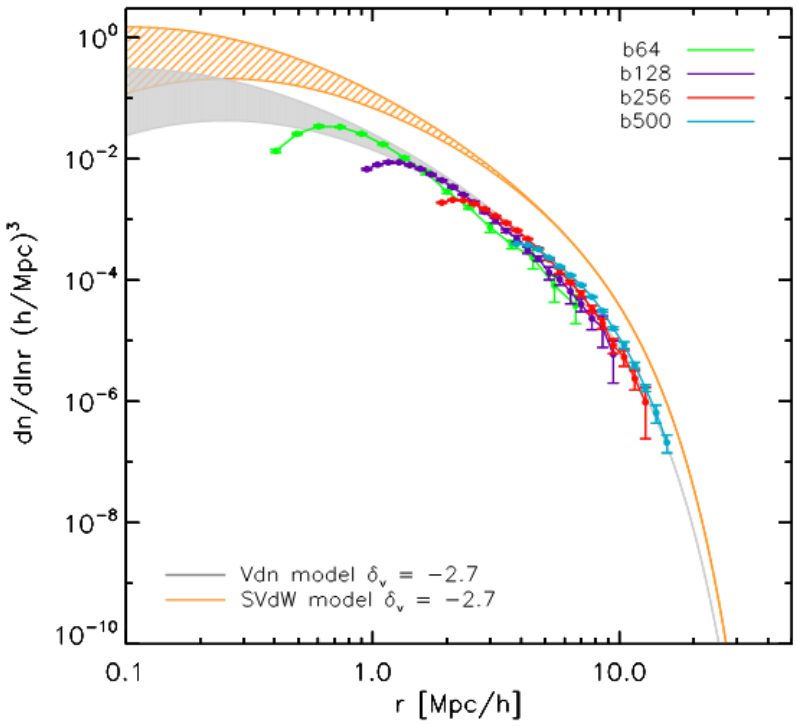

Figure 3.8: Void abundance in simulations compared to the model predictions. Voids have radius equivalent to the radius of a sphere with density $\rho_{\rm v} = 0.2\overline{\rho}_{\rm m}$ in the DM distribution of ΛCDM cosmology simulations, with box sizes of $64\,h^{-1}$ Mpc (green), $128\,h^{-1}$ Mpc (purple), $256\,h^{-1}$ Mpc (red) and $500\,h^{-1}$ Mpc (cyan). The error bars represent the scatter on the mean from eight different realisations of this cosmology in each box size. The range in predictions cover the parameter interval $\delta_c = [1.06, 1.69]$ with $\delta_v = -2.7$ and are consistent with simulations for Vdn (grey shaded) but not for SvdW models (orange hatched). Credits: Jennings et al. (2013).

measurement, due to the effect of the bias (see Sec. 2.3). The action of bias can also be measured in simulations through the trends of the mean profiles of voids in the various tracer distributions. The DM profiles are less deep and steep than those of the biased tracers. As a direct consequence, it follows that rescaling the sample of voids identified in different biased tracer field to the same density contrast, the corresponding mean effective radius increases with the minimum mass of the sample.

In order to be able to use real voids, i.e. measured by observation of biased tracers, as cosmological probes, it is necessary to relate the characteristic density threshold used in the theoretical model, $\delta_{\rm v}^{\rm L}$, to the corresponding one in the biased density field (Contarini et al., 2022). Based on Contarini et al. (2019) and Contarini et al. (2021) and as presented in Contarini et al. (2022), it is possible to show how the Vdn model can be extended to consider a linear relationship, $\mathcal{F}$, between tracer and matter density contrast in cosmic voids, $\delta_{\rm v,tr}^{\rm NL}$ and $\delta_{\rm v,DM}^{\rm NL}$, with a dependence only on the *large-scale effective bias* $b_{\rm eff}$:

$$\delta_{\rm v,DM}^{\rm NL} = \frac{\delta_{\rm v,tr}^{\rm NL}}{\mathcal{F}(b_{\rm eff})} \,, \tag{3.13}$$

where $\delta_{\rm v,DM}^{\rm NL}$ is the value of the threshold in the DM field to be used in the Vdn model, after its conversion in linear theory (Jennings et al., 2013). To use the rescaled threshold in the Vdn model, $\delta_{\rm v,DM}^{\rm NL}$ is converted to its linear counterpart through the fitting formula provided by Bernardeau (1994):

$$\delta_{\rm v,DM}^{\rm L} = \mathcal{C}[1 - (1 + \delta_{\rm v,DM}^{\rm NL})^{-1/\mathcal{C}}] \,, \tag{3.14}$$

where $\mathcal{C} = 1.594$. This equation is exact for models with null cosmological constant Λ, and is a good fit for any values of Λ, especially for the underdense regions.

We now introduce a convenient relation describing the punctual bias calculated at $r_{\rm eff}$, $b_{\rm punct}$:

$$b_{\rm punct} \equiv \left\langle \frac{\delta_{\rm v,tr}^{\rm NL}(r = r_{\rm eff})}{\delta_{\rm v,DM}^{\rm NL}(r = r_{\rm eff})} \right\rangle \tag{3.15}$$

Since in our analysis the value of $\delta_{\rm v,tr}^{\rm NL}$ is fixed (usually at -0.7), then $\delta_{\rm v,DM}^{\rm NL}(r = r_{\rm eff})$ is exactly the value we need to rescale the void size function model. The function $\mathcal{F}$, described in Eq. (3.13), is well modelled by the linear relationship (Contarini et al., 2019):

$$\mathcal{F}(b_{\rm eff}) = B_{\rm slope} b_{\rm eff} + B_{\rm offset} \,, \tag{3.16}$$

where $B_{\rm slope}$ and $B_{\rm offset}$ are the values of the first and second coefficients of the linear function, respectively. This relationship provides the necessary calibration for rescaling voids from the density field of the DM to the density field of the biassed tracers. The coefficients $B_{\rm slope}$ and $B_{\rm offset}$ are determined through a linear fit on the $b_{\rm punct} - b_{\rm eff}$ plane, i.e. on the values of $b_{\rm punct}$ calculated at various z.

3.2.3 Void density profile

Another important statistical tool related to voids is the density profile, studied in detail in the recent literature (see e.g. Sutter et al. 2014b, Hamaus et al. 2014). The void density profile is defined as the spherically averaged relative deviation of

mass density around a void centre from the Universe density mean value $\overline{\rho}$. Using the tracer particles, the density in a radial shell of thickness $2\delta r$ at distance r from a void centre can be estimated as

$$\rho_v(r) = \frac{3}{4\pi} \sum_i \frac{m_i(\vec{r}_i)\Theta(r_i)}{(r+\delta r)^3 - (r-\delta r)^3} \,, \tag{3.17}$$

where m_i is the mass of particle i, $\vec{r}_i$ represent its coordinate vector of length r_i, and $\Theta(r_i) \equiv \theta[r_i - (r-\delta r)]\theta[-r_i + (r+\delta r)]$ combines two Heaviside step functions θ to define the radial bin.

Hamaus et al. (2014) proposed a simple empirical formula to parametrise the profile:

$$\frac{\rho_v(r)}{\overline{\rho}} - 1 = \delta_c \frac{1-(r/r_s)^\alpha}{1+(r/r_v)^\beta} \,, \tag{3.18}$$

where δ_c is the central density contrast, r_s a scale radius at which $\rho_v = \overline{\rho}$, called *effective radius*, and α and β determine the inner and outer slope, respectively, of the density peak due to the overdense shell enveloping the void, called *compensation wall*. In Fig. 3.9 we show the best fits of this four-parameter model to the stacked void density found in a numerical simulation from Hamaus et al. (2014). Voids

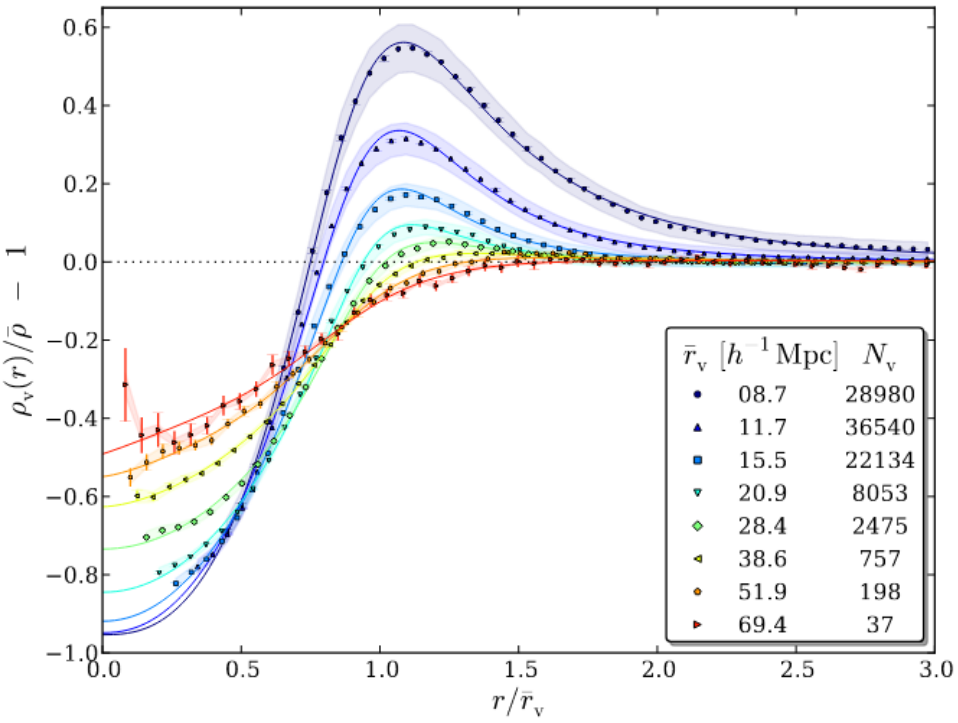

Figure 3.9: Stacked density profiles of voids at $z = 0$ in eight contiguous bins in void radius with mean values and void counts indicated in the inset. Shaded regions depict the standard deviation σ within each of the stacks (scaled down by 20 for visibility), while error bars show standard errors on the mean profile $\sigma/\sqrt{N_v}$. Solid lines represent the best-fit solutions from Equation (3.18). Credits: Hamaus et al. 2014.

are deeply underdense in their interiors, especially the smallest ones. All the profiles exhibit overdense compensation walls with a maximum located slightly outside

their effective radius, shifting outwards for larger voids. The height of the compensation wall decreases with void size, causing the inner profile slope to become less steep and the wall to spread. This trend divides all voids into the ones being *overcompensated* or *undercompensated*, depending on whether the total mass inside their compensation wall is respectively higher or lower than the mass missing in the centre. Ultimately, at sufficiently large distances to the void centre, all profiles approach the mean background density.

Another method for calculating density profiles consists in calculating the cross-correlation between centres of voids and tracers, using the two-point correlation function, $\xi_{\rm v,tr}(r)$. This estimator is associated to the probability of finding a tracer at a comoving distance r from a void centre (see Sec. 2.1.3). It can be expressed as the integrated void density contrast profile, computed in a sphere of radius r and volume V centred in voids:

$$\xi_{\rm v,tr}(r) = \frac{1}{3r^2}\frac{\mathrm{d}}{\mathrm{d}r}[r^3\Delta(r)] \ , \tag{3.19}$$

where

$$\Delta(r) = \frac{3}{r^3}\int_0^r \delta(r')r'^2 dr' \ . \tag{3.20}$$

A final method for calculating the density profile consists in using the relation between the displacement field divergence and the density contrast, described by Eq. (2.88) and, more in general, by the Zel'dovich theory (Sec. 2.2.2). According with this theory, the mean value of the divergence calculated at radius r from the centre of the voids assumes exactly the value of the density contrast. With respects to the previous measuring strategies, this method has the advantage that it does not depend directly on the tracer positions and thus overcomes the problems associated with mass discretisation. However, this relationship loses its validity in the central regions of voids, where the nonlinear regime starts to break down. Moreover, it requires the exact reconstruction of the linear displacement field. The application of this reconstruction of density profiles will be analysed in more details later in this Thesis work.

Chapter 4

Back in time void finder

The identification of voids represents the first and fundamental step towards their cosmological exploitation. As introduced in Sec. 3.1, the main identification methods, based on geometric and density criteria, suffer from problems related to shot noise and mass discretisation. To overcome these issues, in this Thesis work we develop a new void finder based on dynamic criteria, that we call `Back in time void finder` (`BitVF`), improving the idea proposed by Elyiv et al. (2015). `BitVF` allows us to reconstruct a continuous density field and provides the positions of voids without using mass tracers to sample the density field. The `BitVF` algorithm is presented in Sec 4.1. Moreover, it is possible to clean up the void catalogue obtained with `BitVF` in order to obtain void statistics in agreement with the Vdn model (Sec. 3.2.2), allowing the following cosmological exploitation of the acquired data. A first implementation of the cleaning algorithm, based on the work of Jennings et al. (2013), was presented in Ronconi & Marulli (2017). As part of this Thesis work we proposed a new improved version of this algorithm, characterised by higher stability and accuracy, presented in Sec. 4.2 Both `BitVF` and the cleaning algorithm are now included in the `CosmoBolognaLib` (`CBL`, see Sec. 4.4) library (Marulli et al., 2016), a large set of *free software* C++/Python libraries for cosmological computing. Finally, the results of the application of `BitVF` to various cosmological tracer samples from the DUSTGRAIN-pathfinder (Sec. 4.3) simulations are presented in Sec 4.5.

4.1 The void finder algorithm

The `BitVF` algorithm can be summarised in three main steps: reconstruction of the displacement field (Sec. 4.1.1), computation of the density field (Sec. 4.1.2), identification of voids (Sec. 4.1.3). In the first step, tracers are used to sample the displacement field between their current position and their position at the time of perturbation formation. The reconstruction can be achieved using two different methods. The first method, `Lagrangian Zel'dovich approximation for void finder algorithm` (`LaZeVo`) consists of a statistical reconstruction of the displacement field based on a minimum action minimisation. The second method, trivially called `exact reconstruction`, is intended for use in simulations and is based on the reconstruction of the displacement field using the *identification numbers* (IDs) of the various tracers. An exhaustive description of both methods is given in Sec. 4.1.1. Once the displacement field has been reconstructed, the density field is imme-

diately obtained through Eq. (2.88). At this point, the local minima of the density field, identified in the underdense regions, are defined as the centres of the voids. The position of the centres is corrected through a weighted mean of the local values of the density field. Finally, the void radius is computed as the radius of a sphere containing a certain fraction of the average density of the Universe. The various steps leading to the identification of voids are schematised in Fig. 4.1.

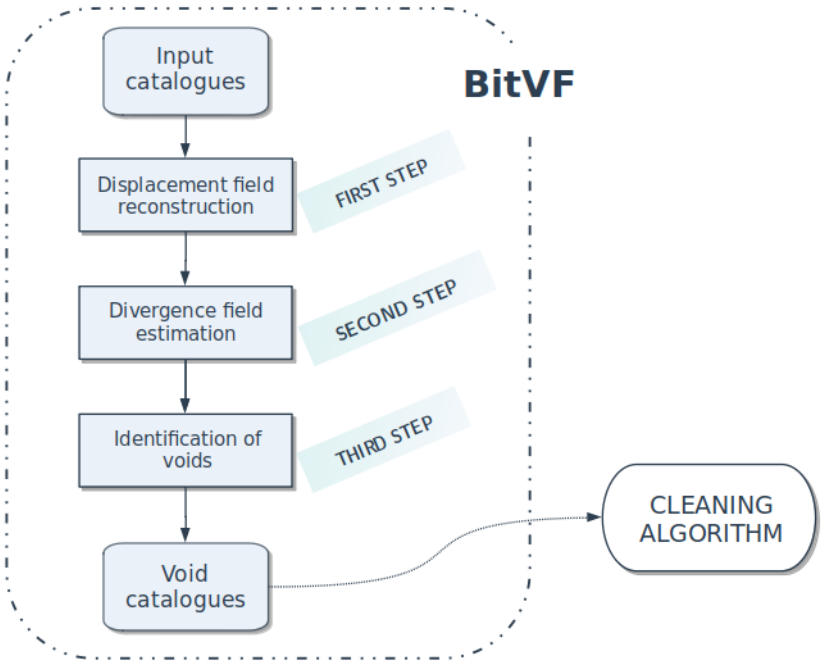

Figure 4.1: The diagram shows the various steps of the `BitVF` algorithm that permit the identification of voids from a catalogue of tracers.

`BitVF` requires two catalogues as input: one providing the comoving coordinates of the cosmic tracers and the other providing the comoving coordinates of the tracers at the time of the formation of the perturbations (i.e. a catalogue of random positions). The algorithm, after receiving as input a set with the fundamental cosmological parameters, is able to create a random catalogue in case it is not provided by the user (However, for future applications to real data, the catalogues may have complex geometries and therefore the random catalogue will have to be provided by the user). The coordinates of the random positions are randomly extracted from a uniform distribution in a volume equivalent to the one of the cosmic tracer catalogue.

At this stage, before moving on to the reconstruction of the displacement field, `BitVF` executes an algorithm whose purpose is to minimise the particle search time. This algorithm, called `ChainMesh`, is a "pixelisation"algorithm, i.e. it divides the space into cells. The volume is in fact divided into cubic cells, each of which is associated with the coordinates of the centre, a label and the index of the particles contained. Through an algorithm that allows cells to be selected around a selected position, the search volume is restricted to a region of arbitrary size, making the search for particles considerably more efficient. This is very useful for example, as we will discuss in Sec. 4.1.1, in the creation of the *tracer-random* pairs on which the `LaZeVo` method of displacement field reconstruction is based.

Figure 4.2 shows an example of the application of the `ChainMesh` algorithm for searching particles within a certain radial distance from a given point. This is just one of the possible applications of the algorithm, whose current version based on object-oriented programming, included in the `CosmoBolognaLib` libraries, was implemented as part of this Thesis work. In particular, a whole class of methods has been developed and included in the CBL, which allow particles to be searched and managed according to different criteria. For example, it is possible to search for the N particles closest to a certain point or to search for particles contained within a certain radial distance from a selected one, ensuring high performance efficiency.

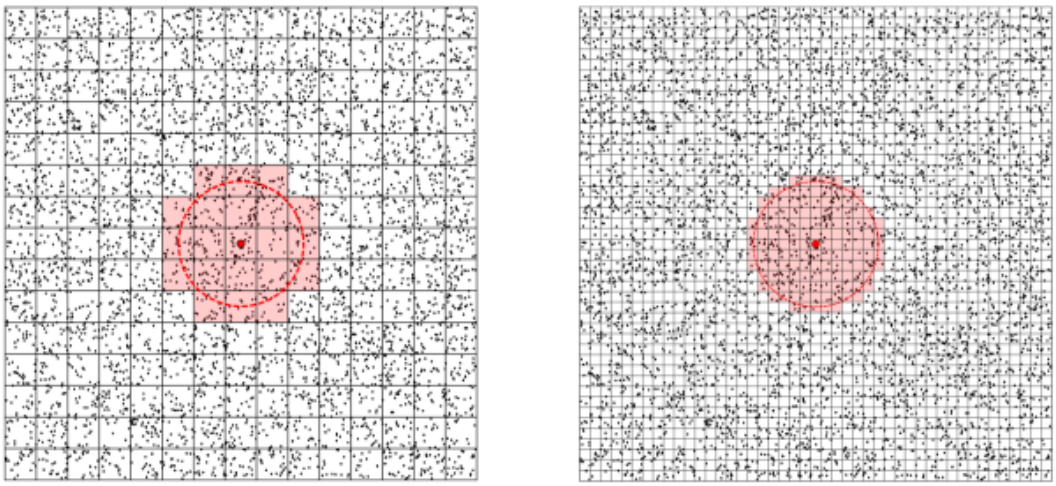

Figure 4.2: A schematic image of the `ChainMesh` technique for two different cell size values. In these applications, for a given point (the red dot), `ChainMesh` searches for objects within the region (marked by the dashed red circle) of radius $r_{\rm max}$. Cell size should be a compromise between RAM consumption and computational time. A larger cell number allows faster selection of the searched particles, at cost of higher resource consumption.

4.1.1 Displacement field reconstruction

Once the preliminary operations have been completed, `BitVF` is ready for the reconstruction of the displacement field, which is the cornerstone on which we based the search for voids through any considered method. Potentially, any method that allows us to statistically reconstruct straight-line trajectories between the current positions of the tracers and their positions at the Universe earlier time could be used for perform this reconstruction. Two displacement field reconstruction methods were applied in this Thesis work and included in `BitVF`. The first, `LaZeVO`, is based on a statistical reconstruction of trajectories performed by minimizing the total distances between tracer-random position pairs, a quantity directly proportional to the total action of the system. The second, on the other hand, called simply `Exact reconstruction`, is used for analyses of simulations. It allows the positions of the analyzed tracers to be linked with their positions at the initial time. However, it is crucial to remember that each method must reconstruct rectilinear trajectories,

as the entire density field reconstruction procedure is based on exploiting the displacement field through Zel'dovich theory (see Section 2.2.2), which involves the linearisation of trajectories.

LaZeVo reconstruction

The algorithm underlying the `LaZeVo` reconstruction is subdivided into two main phases: the *setting phase* and the *swapping phase*. The setting phase provides a preliminary assignment of tracer-random position pairs. This step is crucial to reduce the calculation time and increase the accuracy of the reconstruction process. The idea behind the method is to facilitate the identification of the minimum action configuration (or, more precisely, the minimum total displacement) by attempting to eliminate very long distance displacements, which are evidently unphysical, already in the initial stages. The setting phase operates in the following way:

1. The comoving coordinates of a point in the volume occupied by the tracers are randomly extracted;

2. Tracers located within a certain distance (set at four times the mean particle separation, mps) from the randomly extracted point are found;

3. The N random positions closest to the extracted point are identified, where N is the minimum between the number of tracers found in the previous step and $N_{max} = 100$. This limit is imposed to maintain a certain number of unpaired tracers during the first iterations of the algorithm, in order to make the initial displacement range more uniform. Without this expedient, each iteration might create a "bubble"of particles paired with neighbouring random positions, but without establishing interconnections between the various bubbles. This leads to a loss of efficiency in minimising the total displacement distance;

4. N tracer-random pairs are randomly created between the selected positions;

5. The algorithm is iterated until each tracer has been assigned to random objects. The result of a first run of the four steps of the algorithm, applied to a 2D toy model, is shown in Fig. 4.3. In particular, the process of particle selection and pairing is visualized.

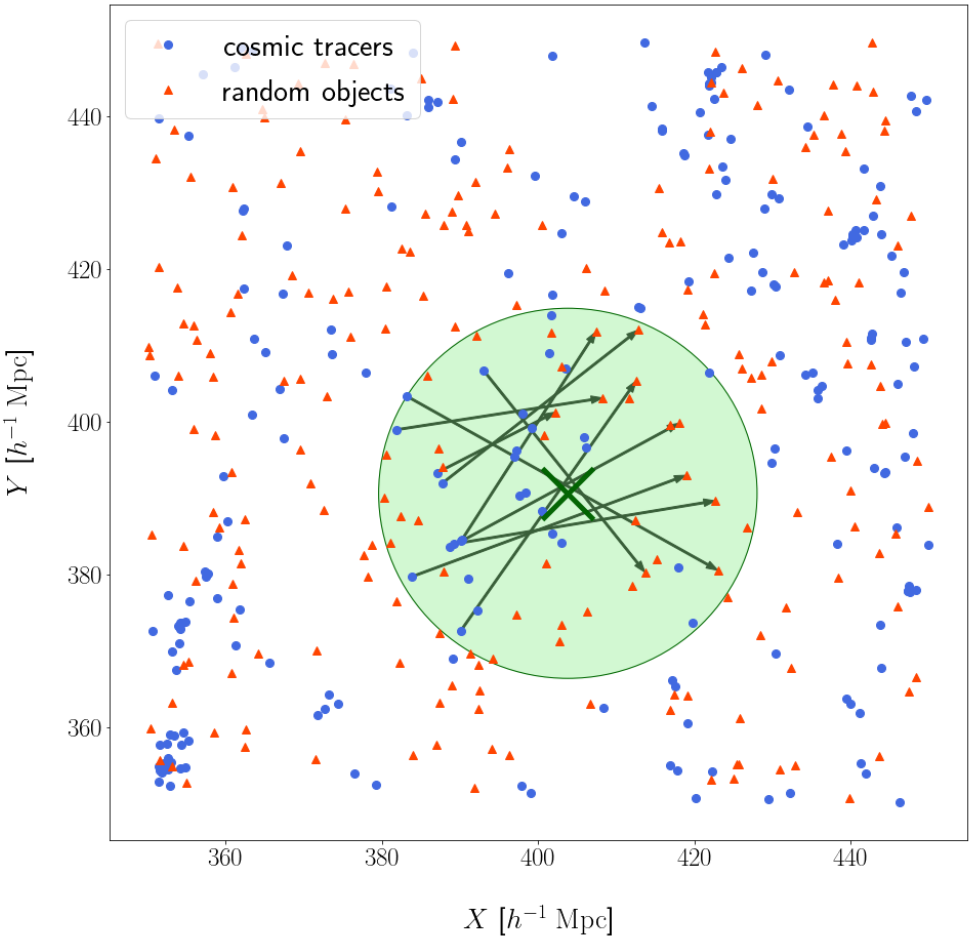

Figure 4.3: The figure shows the result of the first iteration of the setting algorithm of the `LaZeVo` method in a 2D toy model. In particular, the tracer positions taken from a very thin slice (5 h^{-1}Mpc) of the central zones (100×100 h^{-1}Mpc) of a 0.4% subsample of the DUSTGRAIN-*pathfinder* simulations (Giocoli et al., 2018) are shown (in blue), in addition to the random catalogue positions (in red). The green circle of radius 4 mps indicates the area around the randomly extracted position (dark green cross) where position pairing is performed. Finally, the arrows show the N matches performed. In this example, $N_{max} = 10$.

Once the setting process is complete, the actual reconstruction of the displacement field can be performed. The fundamental idea, as previously discussed, is to minimise the total distance of the displacements of the particles composing the system, in order to minimise the total action. The minimum solution can be calculated, from a theoretical point of view, by solving the combinatorial calculus *assignment problem*. This problem can be resolved exactly by evaluating all the possible permutations of the system, $N!$, where N is the number of particles. Dealing with systems with a very large number of particles ($N > 10^6$), it is simply impossible to solve the problem in this way. Although algorithms that greatly improve the scalability of the problem exist (Kuhn 1955, Orlin et al. 1993), the only possible method of dealing with such a large number of tracers is through the research of an approximate solution. The approximate solution is sought through a "swapping" process, starting with the configuration described above. A point-by-point description of the

implemented algorithm is given below.

1. A tracer is randomly selected.

2. Three other tracers are selected from the $N = 112$ closest to the selected one. N represents the number of particles expected in a sphere of radius equal to 3 mps, with density equivalent to the average density of the catalogue (excluding the selected particle). It is given by:

$$\overline{n}(r < R) = N(r < R)V(R) \quad \rightarrow \quad N(r < R) = \frac{\overline{n}(r < R)}{V(R)}, \tag{4.1}$$

   where $\overline{n}$ is the numerical density, R a certain radius and V the volume of a sphere of radius R. Knowing that mps is defined as:

$$\text{mps} = \overline{n}^{-1/3} \ , \tag{4.2}$$

   It follows that, if we consider $R = 3\text{mps}$:

$$N(r < 3\text{mps}) = \text{mps}^{-3} \cdot \frac{4}{3}\pi(3\text{mps})^3 \simeq 112 \ . \tag{4.3}$$

3. All possible 4!=24 permutations of the selected tracers are evaluated by swapping tracer-random pairs, using the method shown in the Fig. 4.4. As illustrated, the configuration that minimises the total distance of the system formed by the four particles is saved.

4. The process continues until a given threshold is reached. The threshold is evaluated as the number of tracer-random swapped pairs divided by the total number of pairs. This ratio is calculated at the end of each cycle, that is, any time all the tracers have been selected at least once.

Therefore, the displacement field reconstruction is based on solving assignment problems consisting of four tracer-random pairs. The choice of using quartets of pairs is the result of the requirement to maintain a low number of permutations for each individual problem, which scales factorially with respect to the number of particles, and the necessity of maximising the number of pairs swapped at each operation. Both of these features are required to achieve maximum efficiency in terms of computational time. Finally, the whole process can be repeated an arbitrary number of times to obtain different reconstructions of the displacement field. Averaging these reconstructions, the density field can be calculated in a more accurate way.

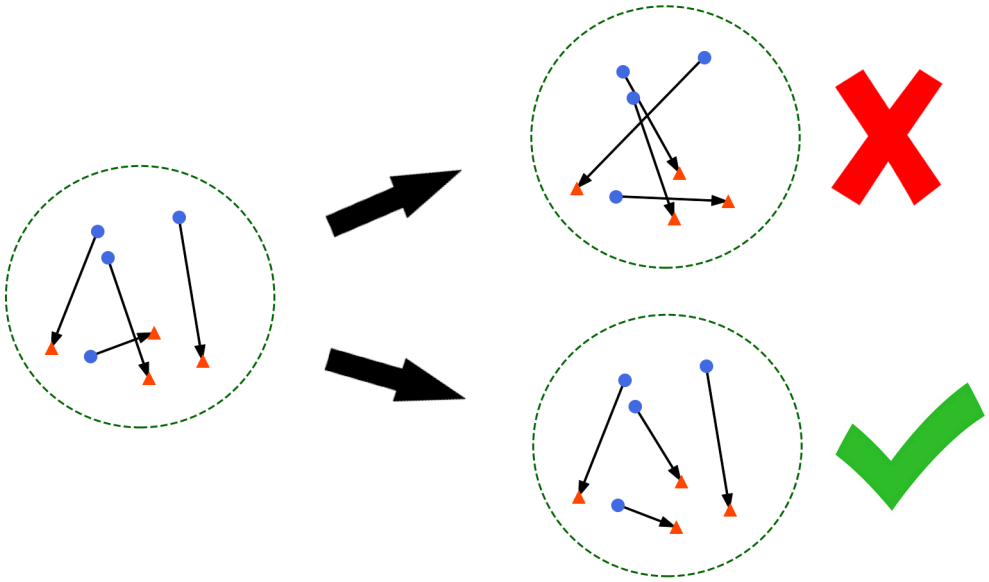

Figure 4.4: A schematic illustration of the swapping process. The assignment problem is solved for each possible permutation of the quartets. A possible case in which the result obtained from solving the problem for one of the permutations is discarded (top) and one in which it is accepted (bottom) are shown on the right.

Exact reconstruction

Unlike the `LaZeVo` approach, this method does not reconstruct the displacement field using statistical procedures but simply exploits the tracer IDs provided in the simulations snapshots. The displacement field is reconstructed by linking the position of the tracers in the snapshot where the voids should be identified with their positions in the initial snapshot. This method has the advantage of allowing the immediate reconstruction of the exact displacement field; however, it can only be exploited in simulations. This method can be used to test the efficiency of the displacement field reproduction through statistical methods by comparing the two reconstructions.

4.1.2 Density field calculation

Once the displacement field has been reconstructed, it is possible to compute the density field through Eq. (2.88), derived from the Zel'dovich theory. It is then necessary to calculate the divergence of the displacement field to obtain the value of the punctual density contrast. The displacement field can be thought as a continuous field sampled by the displacement vectors linking the positions of the tracers to the random positions. The divergence field can thus be calculated using the Gauss's theorem:

$$\int_V (\nabla \cdot \vec{F}) \mathrm{d}V = \oint_S (\vec{F} \cdot \hat{n}) \mathrm{d}S \,, \tag{4.4}$$

where $\vec{F}$ is the field whose divergence is sought, in this case represented by the tracer-random sampling, and $\hat{n}$ is the normal to the surface through which the flux

is computed. The volume is divided into cubic cells aligned with axes, in order to maintain $\hat{n}$ constant on each side x. This allows us to simplify Eq. (4.4) as follows:

$$\nabla \cdot \Psi = \sum_n \sum_{i=1}^{6} \frac{(\vec{F}_n \cdot \hat{e}_i)x^2}{x^3} = \frac{1}{x} \sum_n \sum_{i=1}^{6} (\vec{F}_n \cdot \hat{e}_i) \ , \tag{4.5}$$

where $\vec{F}_n$ is each sample and $\hat{e}$ is the normal to each side of the cell. $\hat{e}$ can assume six values: $[\pm 1, 0, 0]$, $[0, \pm 1, 0]$, $[0, 0, \pm 1]$. The divergence calculated in each cell is therefore reduced to the sum of the scalar product between each vector entering or leaving the cell and the normal to the surface. From a theoretical point of view, the value of the divergence of the displacement field should return the exact value of the density contrast. However, this statement is not always true, since, particularly in underdense areas, we are dealing with a small number of samples per cell. The cell size x, indeed, must represent a good compromise between the number of samples crossing its surface and spatial resolution. This parameter is set by default to $x \simeq 0.8$ mps. In the case where several reconstructions of the displacement field have been performed, each of them is considered for the computation of the divergence. The final value in each cell is then renormalised dividing it by the number of reconstructions.

Finally, the obtained divergence field requires a smoothing process. This is necessary in order to create a continuous vector field and eliminate the cells in which the divergence value is zero, i.e. those cells in which there are no samples of the displacement field. Smoothing is done through a *Gaussian weighting*, with the weight function centred in the middle of the cell under consideration and $\sigma = x$:

$$\Theta = \langle \nabla \cdot \Psi \rangle = \sum_i \frac{(\nabla \cdot \Psi)_i \exp\left(-d_i^2/2\sigma^2\right)}{\exp\left(-d_i^2/2\sigma^2\right)} \ , \tag{4.6}$$

where d_i is the Euclidean distance between the central and the i-th cell. It is important to note that only the considered cell and the immediately adjacent cells are taken into account by the function.

4.1.3 Identification of voids

The last step of the `BitVF` method consists in the identification of voids. The position of each void is theoretically defined as a local minimum of the density field. The algorithm then identifies the cells with a negative divergence value and bordering $3^3 - 1 = 26$ cells with a divergence value greater than the one of the central cell, but still less than zero. A schematic of the situation under consideration is shown in Fig. 4.5. All 27 cells possess negative divergence, however the central cell has a more negative divergence than the others. The cells found in this way constitute the core of the voids identified by the method. `BitVF` then assigns to each void the coordinates of its centre and its radius.

The position of the centre is assigned by a weighted mean for each individual coordinate. The mean is calculated over the $3^3 = 27$ adjacent cells and the divergence of each cell takes on the role of weight. In particular, the equation for each coordinate takes the form:

$$x_i = \frac{\sum_{j=1}^{27} \Theta_j d_j}{\sum_{j=1}^{27} \Theta_j} \ , \tag{4.7}$$

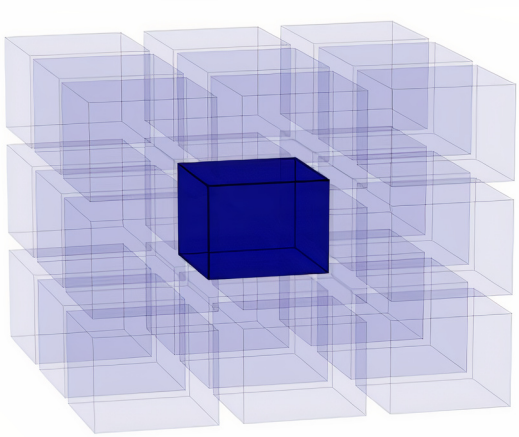

Figure 4.5: A schematic picture of a grid cell having a divergence value lower than the $3^3 - 1$ cubic cells surrounding it. However, all of them have negative divergence value.

where $x_i = [x, y, z]$ for $[i = 1, 2, 3]$, Θ_j represents the divergence computed within the j-th cell and d_j the coordinate difference between the centre of the inner and j-th cells. The position of the centre is thus corrected to reflect the value of the surrounding density field. In this way, the local minimum is identified in a more precise manner.

The void radius, instead, is assigned by reconstructing the density profile calculated in spheres of radius r from the centre. In particular, the radius of the void is considered to be that of a sphere with a density equal to $1/2$ the average density of the tracer catalogue. At this point, the void catalogue, including centre coordinates and radii, is ready.

4.2 Void cleaning

One of the main issues in exploiting cosmic voids as cosmological probes lies in the different adopted void definitions. In particular, the Vdn model, described in Sec. 3.2.2, defines the cosmic voids as underdense, spherical, non-overlapped regions that have gone through shell crossing. To extract cosmological information from void counts it is thus required to use the same definition when detecting voids, and to clean properly the void catalogues detected with standard methods. Therefore, in order to align the catalogues obtained not only through BitVF, but potentially through any void finder, a cleaning algorithm based on the work of Jennings et al. (2013) and Ronconi & Marulli (2017) was developed within the context of this Thesis. The original algorithm is based on three main steps, which can be summarised as:

1. Statistically irrelevant objects are removed. By statistically irrelevant objects are meant all underdensities with a radius outside a given range $[r_{\rm min}, r_{\rm max}]$ and those with a density contrast of more than a certain fraction of the average density of the Universe.

2. The radius of the voids is rescaled as the largest radius from the void centre which encloses an underdensity equal to the shell-cross density contrast δ_{sc}.

3. Overlaps are eliminated in order to avoid double counting. Two voids are considered to be overlapped when the distance between their centres is smaller than the sum of their radii. When two voids overlap, the void with the highest central density is removed.

However, this method presents issues due to the presence of free parameters. For example, the results are highly dependent on the radius at which central density and density contrast are calculated. Furthermore, the calculation of these quantities brings with it systematic issues related to the discretisation of the quantities and shot noise, in particular due to the presence of a small number of tracers in the central region of the voids. These difficulties lead to errors when removing overlaps, as the criterion is totally dependent on central density values, and to a strong dependence on $r_{\rm min}$. Finally, significant limitations related to the shape of the voids are present. The model assumes spherical voids; however the actual shape of the underdensities is irregular and elliptical. This, in combination with mass discretisation, leads to irregular spherical density profiles, which complicates the search for the threshold at which to set the radius. To overcome these difficulties, the algorithm was modified in the following way:

1. The central density is calculated as the density of a sphere of radius 2 mps from the centre. This distance represents the theoretical minimum spatial resolution for the identification of voids. Fixing this parameter independently of the considered void removes the dependence of the central density on the radius, improving the significance of the measurement. In particular, this distance turned out to be an excellent compromise between the necessity to minimise the error related to the discretisation of the mass and the requirement to consider the most central regions of the voids. A filter is thus applied to remove all underdensities with a central density below the threshold chosen for the radius rescaling. This filter has been proven to be crucial in removing spurious voids, eliminating most of the smaller underdensities, which represent the largest source of noise in the count statistics.

2. The density profiles of the underdensities are reconstructed and the radii of the voids, $r_{\rm v}$, are rescaled to coincide with the radius of the largest sphere of given density contrast (usually $\delta_{\rm v,sc}$ for DM). If no suitable profile point can be identified for radius identification (an increasing density profile section with values across the threshold is requested), the search region is expanded. In this case, the radius will be that of the smallest sphere found. Finally, if even in this case it is not possible to set the radius (e.g. no regions of density below the threshold are found), the void is removed.

3. An additional filter is applied to the profiles, repeating the rescaling process for a threshold $\delta_{\rm v} = -0.5$. If the threshold is not identified, the void is removed. This process removes voids that are badly centred or with very pronounced ellipticity.

4. The density contrast Δ between two spheres centred in the centre of the void, of radius 1.5 and 1 $r_{\rm v}$, respectively, is calculated. If Δ is lower than 1, the underdensity is considered as cloud in void and eliminated.

5. Voids with a radius outside the range $[r_{\rm min}, r_{\rm max}]$ are eliminated. $r_{\rm min}$ is typically set to $1.5 - 2$ mps. The method described in the previous steps results in a very stable statistic with respect to $r_{\rm min}$. Counts at $r > r_{\rm min}$ are poorly influenced by this parameter, differently than with the old algorithms.

6. Finally, overlaps are removed using central density as a selection criterion. When an overlap occurs, the void with the lower central density is preserved.

This new version of the algorithm has therefore the advantage of eliminating a number of free parameters and strongly reduces the dependence of the obtained statistic on the remaining ones. Moreover, it produces results in excellent agreement with the theoretical size function model. Fig. 4.6 shows the action of the various steps of the cleaning algorithm on the counts of voids found through the use of `BitVF`.

4.3 DUSTGRAIN-pathfinder simulations

The void finder algorithm presented in this Thesis work, BitVF, was tested on simulations with DM and biased tracers, at various redshifts and with various subsamples of the catalogue. For this purpose, we use a ΛCDM subset of the cosmological N-body simulations suite called DUSTGRAIN-pathfinder (Dark Universe Simulations to Test GRAvity In the presence of Neutrinos, Giocoli et al., 2018). These simulations have been specifically designed with the aim of investigating the degeneracies between $f(R)$ gravity models and massive neutrinos. The DUSTGRAIN-pathfinder simulations have been performed using `MG-GADGET` (Puchwein et al., 2013), a code based on an updated version of `GADGET2` (Springel, 2005). The code follows the evolution of a set of 2×768^3 particles of DM and massive neutrinos (768^3 for each popolution) within a periodic cosmological box of 750 h^{-1} Mpc per side. In the used ΛCDM simulation (i.e. the one characterised by GR and $M_\nu = 0$ eV) the CDM particle mass is equal to $M^P_{cdm} = 8.1 \times 10^{10}\ h^{-1}\ {\rm M}_\odot$. The cosmological parameters assumed in these simulations are consistent with the Planck 2015 constraints (Planck Collaboration, 2016): $\Omega_{\rm m} = \Omega_{\rm cdm} + \Omega_{\rm b} + \Omega_\nu = 0.31345$, $\Omega_\Lambda = 0.68655$, $h = 0.6731$, $A_s = 2.199 \times 10^{-9}$, $n_s = 0.9658$, which give $\sigma_8 = 0.842$. Among all the ΛCDM snapshots available for this project, we select the ones at the redshifts $z = 0$, 0.5, 1, 2, considering only CDM particles. Moreover, to reduce the computational time and test the effect of the subsample on the algorithm performance, we used 25%, 10%, 4% and 0.4% of the original particle sample.

The DM haloes, used to test the `BitVF` algorithm on biased tracers, have been identified for each snapshot following the Despali et al. (2016) approach. In particular, the halo catalogues have been obtained by applying the `Denhf` algorithm (Tormen et al. 2004, Giocoli et al. 2008) to the DM particle sample, finding DM haloes as gravitationally bound structures. In the analysis presented in this chapter we employ 200c halo catalogues, thus those derived imposing $\Delta_c = 200$, i.e. overdensities with densities of 200 times the critical density of the Universe. Moreover,

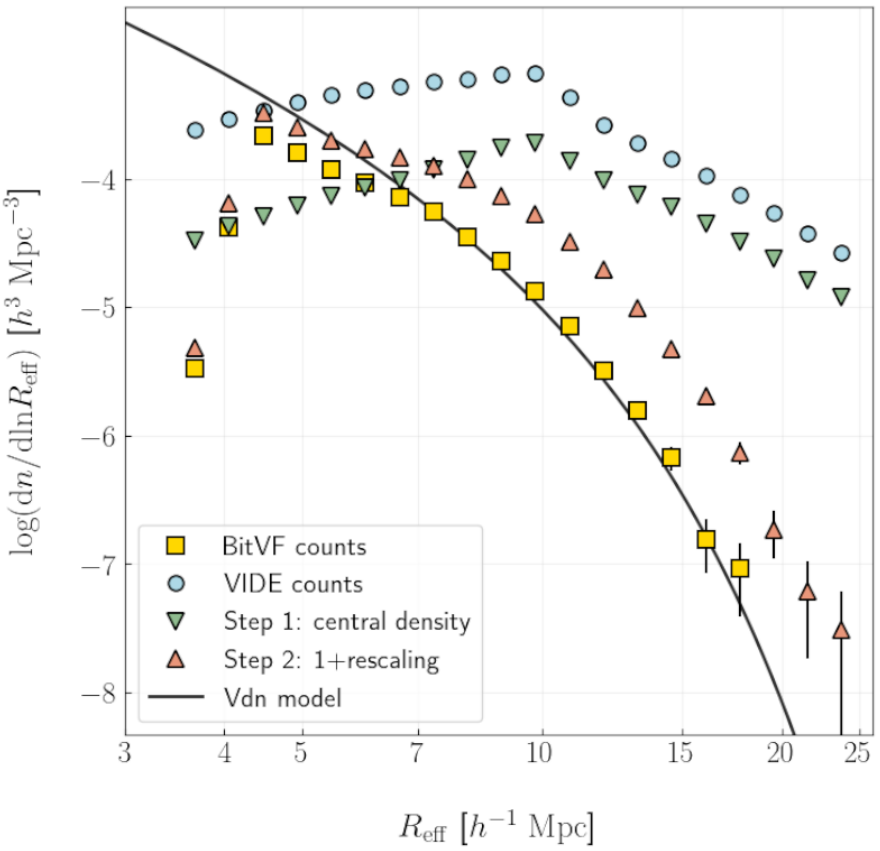

Figure 4.6: Void counts at various stages of the cleaning process, related to the theoretical Vdn model. Voids were identified by applying `BitVF` on DM, at $z = 0$, in a 10% subsample of DUSTGRAIN-*pathfinder* simulations. The light blue circles represent void counts before the action of the cleaner. Green triangles are the counts after the application of the filter on the density contrast. The red triangles show the counts after radius rescaling, profile and density contrast filtering. Finally, the yellow squares show the final counts, the result of applying all cleaning steps, including overlap removal. The continuous black line shows the prediction of the Vdn model. Error bars are obtained from the Poissonian error of the counts.

haloes with a number of embedded DM particles less than 30 are rejected, in order to avoid contamination by spurious density fluctuations and keep only statistically relevant objects. This mass cut corresponds to $M_{\rm min} = 2.43 \times 10^{12} h^{-1}$ M$_\odot$ has been chosen to select a complete, pure and dense sample of DM haloes, fundamental for identifying a statistically significant number of cosmic voids.

4.4 `CosmoBolognaLib`

The `CosmoBolognaLib`[1] is a large set of *free software* C++/Python libraries, that provide an efficient numerical environment for cosmological investigations of the large scale structure of the Universe (Marulli et al., 2016). This software is particularly suited to handle with real and simulated catalogues of astronomical objects. Thanks to the large amount of functions implemented, the `CosmoBolognaLib` offers all the required tools to analyse large data sets and to perform statistical analyses, with optimised performances. In particular, this Thesis project not only extensively exploited the `CosmoBolognaLib` (v5.3, v5.4, v5.5) but additionally integrated it with new codes, namely `BitVF`, the cleaning algorithm, and a number of functional codes for the development of the work, such as a new `ChainMesh` algorithm (see Sec. 4.1). The final version of the implemented algorithms will be available in the next CBL release.

4.5 `BitVF` application

In order to test the `BitVF` algorithm, we applied it to the DUSTGRAIN-pathfinder simulations (Sec. 4.3), both by searching for voids in DM and in biased cosmological tracers (DM haloes). In addition, we tested our void finder by applying it to tracer catalogues at different redshifts and subsample values. Finally, both displacement field reconstruction algorithms were tested, in particular using the results obtained from the `exact reconstruction` to validate the `LaZeVo` method.

The performed tests are listed below:

- DM particle catalogue, displacement field reconstruction through `exact reconstruction` method, $z = 0$, subsample=10%;
- DM particle catalogue, `exact reconstruction` comparison, $z = 0$, subsample=0.04%;
- DM halo catalogue, displacement field reconstruction through `LaZeVo` method, $z = 0$;
- DM particle catalogue, displacement field reconstruction through `exact reconstruction` method, comparison between $z = [0,\ 0.5,\ 1,\ 2]$, subsample=10%;
- DM halo catalogue, displacement field reconstruction through `LaZeVo` method, comparison between $z = [0,\ 0.5,\ 1,\ 2]$.

[1] https://federicomarulli.gitlab.io/CosmoBolognaLib

The first test is intended to check the accuracy of the reconstruction of the divergence field and the detection of the underdensities. This was performed through the exact reconstruction of the divergence field, exploiting the IDs of the DM particles. The result of this preliminary test is shown in the Fig. 4.7. A thin slice of the central zone of the DUSTGRAIN-pathfinder simulations is shown, illustrating the position of the tracers, the displacement field, the divergence field and the identified voids. We can see how the divergence field reconstructs very closely the density field. Furthermore, it can be observed how the arrows, representing the mean displacement field, depart from the overdense regions and point towards the underdense zones, in perfect agreement with what is expected. Finally, voids are correctly identified in underdense regions. In this work, except where otherwise indicated, voids subjected to the cleaning process (Sec. 4.2) and therefore in accordance with the Vdn model (Sec. 3.2.2) are shown: spherical, non-overlapping and with $\delta_{\rm v} = \delta_{\rm v,sc}$. Any overlaps observed in the image are due to projection effects.

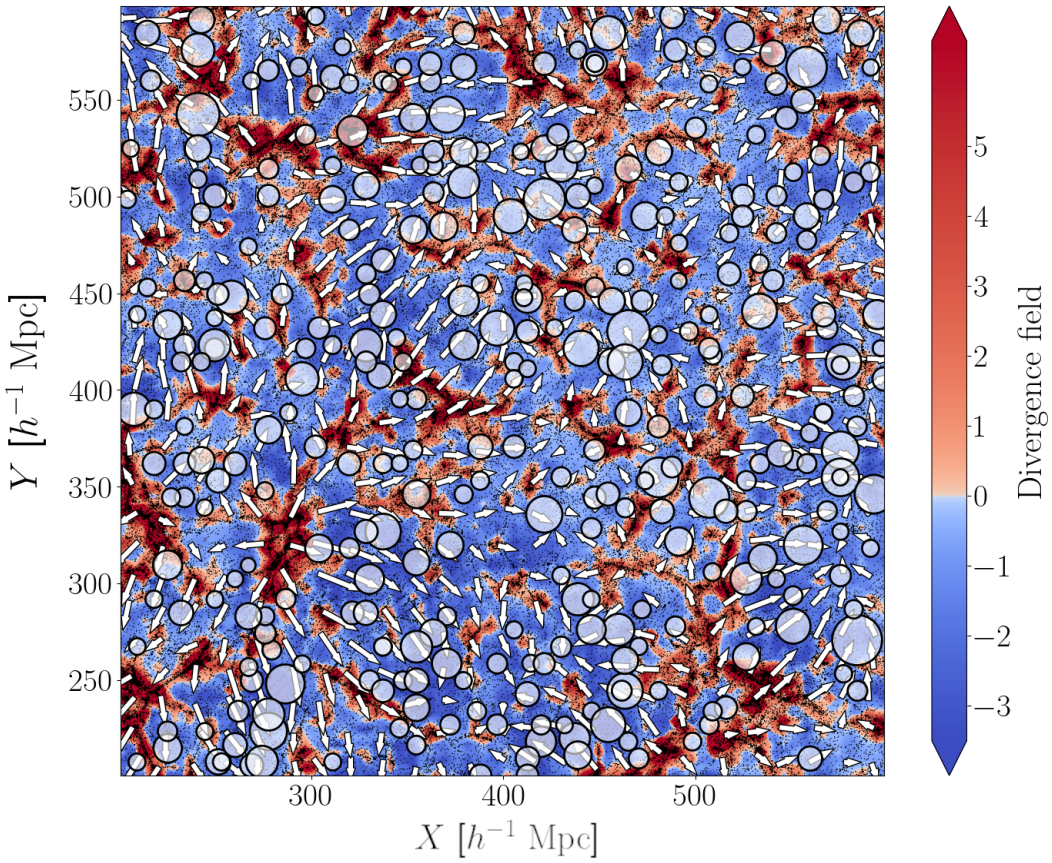

Figure 4.7: The image shows a thin slice (10 h^{-1} Mpc) of the central region (400×400 h^{-1} Mpc) of the $z = 0$ sample of the DUSTGRAIN-pathfinder simulations (10% subsample). In particular, the positions of the DM particles (black dots), the value of the reconstructed divergence field (underdense regions in blue, overdense regions in red), the mean back-in-time displacement field (through a grid of 25×25 arrows) and the voids identified in the sample (white circles) are shown. The reconstruction of the displacement field was performed exactly by exploiting the particle IDs, using tracers at $z = 99$ as a random catalogue.

Once tested the correct functioning of the algorithm for reconstructing the displacement field and identifying voids, we moved on to verify the functioning of the `LaZeVo` method. This method is fundamental for reconstructing the displacement field in all cases where particle IDs are not available, such as when using DM haloes, galaxies or more generally real data as tracers. The `LaZeVo` method was tested through the reconstruction of the divergence field and the identification of voids in DUSTGRAIN-pathfinder simulations at $z = 0$, using DM particles. To reduce the calculation time, a subsample of 0.4% was used. The test consists in comparing the results obtained through the application of the `LaZeVo` method with those obtained through the `exact reconstruction` of the displacement field. The result of this comparison is shown in Fig. 4.8 (the top subplot shows the result obtained via `LaZeVo`, the bottom subplot via `exact reconstruction`). It can be seen that both methods reconstruct the divergence field extremely accurately with respect to the position of the tracers, reproducing the overdensity and underdensity regions very closely. The displacement fields reconstructed with the two methods show the same average trends. However, the displacements obtained by the `LaZeVo` method have smaller amplitudes than the real ones. This effect is due to the minimisation of the action: it does not affect the effectiveness of the underdensity research but breaks the equivalence between divergence of the displacement field and density contrast, which becomes a simple proportionality. Voids are correctly identified in underdensities for both methods. There is a good agreement between the positions of the voids obtained through the two reconstructions, although not perfect. This slight discrepancy is probably the result of the strong subsampling. Based on what is shown, the method is considered validated and usable.

We now have all the necessary elements to test the algorithm on a catalogue of biased tracers. We run `BitVF` on a catalogue of 643590 DM haloes obtained through the application of the `Denhf` algorithm to the $z = 0$ snapshot of the DUSTGRAIN-pathfinder simulations (see Sec. 4.3 for further details). An example of the results obtained is shown in Fig. 4.9. The reconstruction of the displacement field was performed through the `LaZeVo` method. Once again, the displacement field correctly shows the back-in-time displacement of particles from overdense to underdense regions and the divergence field very accurately samples the density field. Voids are larger than those identified in DM due to the smaller number of tracers available, [2] are correctly identified in the most underdense areas. The possibility of exploiting this void finder on biased tracers is therefore confirmed: this aspect is very important as it opens the possibility of its future use with real data.

Finally, the capabilities of `BitVF` in finding voids at different redshifts were tested. The evolution of voids leads these cosmic structures to possess very different density profiles over time. An effective research algorithm must therefore be able to adapt to the context in which it is applied. From a theoretical point of view, `BitVF`, which does not directly exploit density for the identification of underdensities, has the perfect potential for such applications. Figures 4.10 and 4.11 show the results of `BitVF` application to four snapshots of DUSTGRAIN-pathfinder simulations, in an attempt to analyse the performance of the finder at various redshifts. In particular, Fig. 4.10 shows the results at decreasing redshift (z=2, 1, 0.5, 0), using

[2]resolution is inversely proportional to mps, which in turn is inversely proportional to the number of tracers. It follows that the resolution is directly proportional to the number of tracers

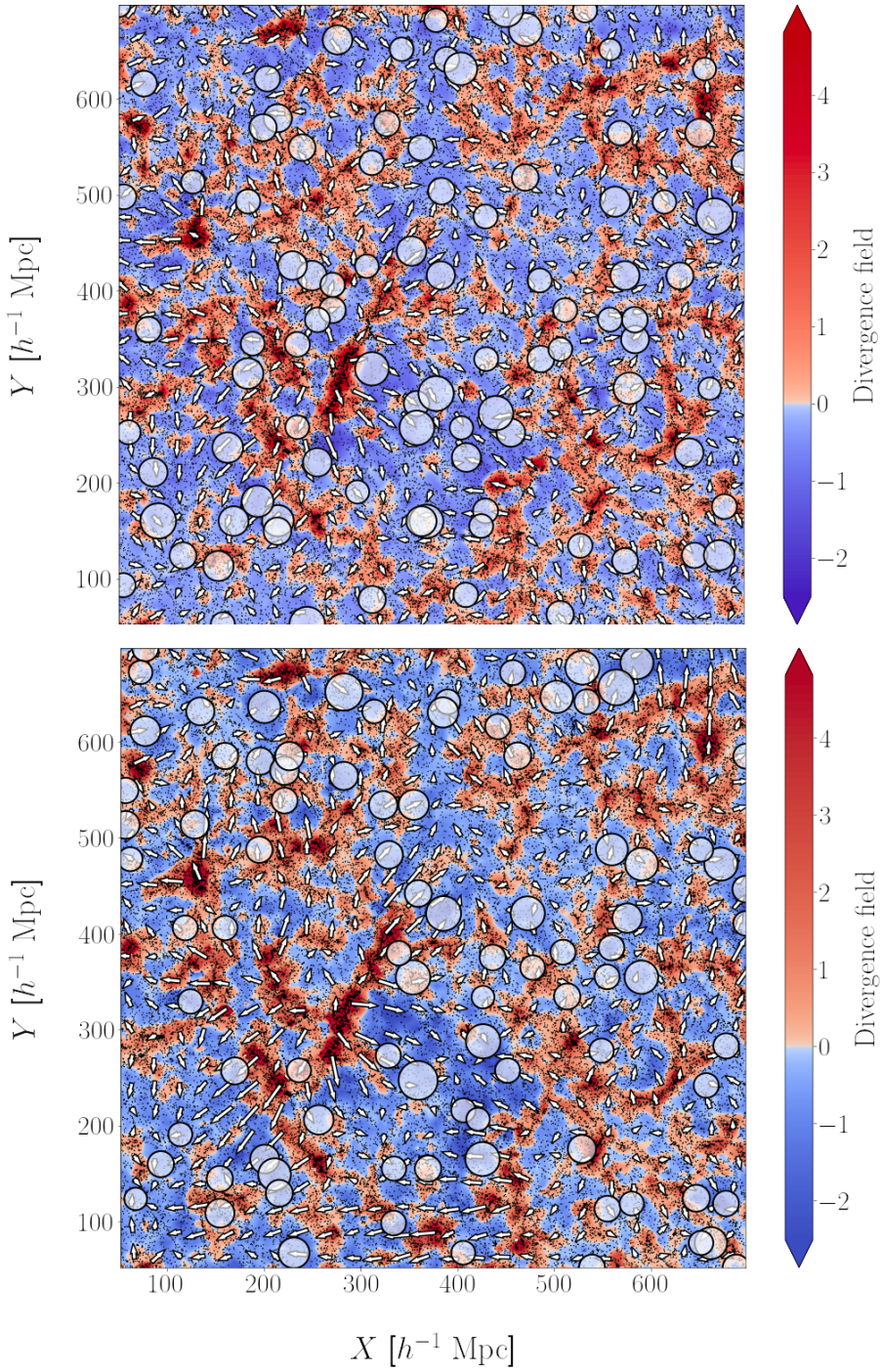

Figure 4.8: The image shows the comparison between the reconstruction of the displacement field through the `LaZeVo` method (top image) and the exact reconstruction of the same (bottom image). In particular, a slice of 35 h^{-1} Mpc of the central region (400×400 h^{-1} Mpc) of the $z = 0$ output of the DUSTGRAIN-pathfinder simulations (0.4% subsample) is shown. The following are displayed: positions of the DM particles (black dots), value of the reconstructed divergence field (underdense regions in blue, overdense regions in red as shown by the coloured scale on the right), mean back-in-time displacement field (through a grid of 25×25 arrows) and voids identified in the sample (white circles).

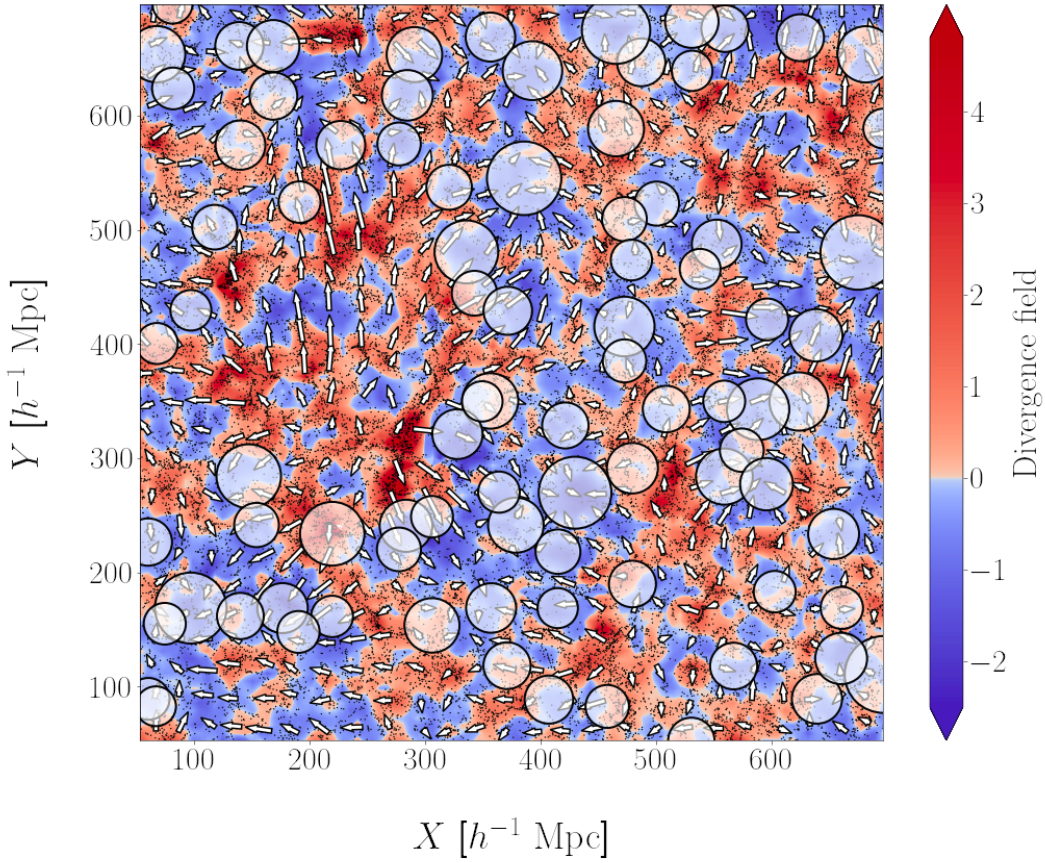

Figure 4.9: The image shows a slice (50 h^{-1} Mpc) of the DUSTGRAIN-pathfinder simulations where the 643590 DM haloes with $M > 2.43 \times 10^{12}$ h^{-1} $M_{\odot}$, identified through `Denhf` application, are exploited for the identification of cosmic voids. In particular, the positions of the DM haloes (black dots), the value of the reconstructed divergence field (underdense regions in blue, overdense regions in red as shown by the coloured scale on the right), the mean back-in-time displacement field (through a grid of 25×25 arrows) and the voids identified in the sample (white circles) are shown. The reconstruction of the displacement field was performed using the `LaZeVo` method.

DM particles as cosmological tracers. Figure 4.11 instead shows the results of the application at the same redshifts using DM haloes as tracers, identified according to the criteria used previously. It is important to note that, unlike DM particles, which remain constant in number over time, the haloes grow in number with time, as a result of the increasing matter clustering. In both cases, the displacement field and divergence field are reconstructed with great precision, effectively showing the evolution of large-scale structures as the redshift changes, starting from more homogeneous configurations and progressing to more clustered ones. With regard to the voids identified, we can notice that, as predicted by theory, at high redshift voids with $\delta_{\mathrm{v,sc}}$ are not detected, while they are detected in biased tracers (in this case $\delta_{\mathrm{v,h}} = -0.7$), which have already had time to cluster. As time progresses, voids emerge gradually from the density distributions of DM and DM haloes. However, the voids identified in the latter are gradually smaller in size as a result of the increase in the number of cosmic tracers.

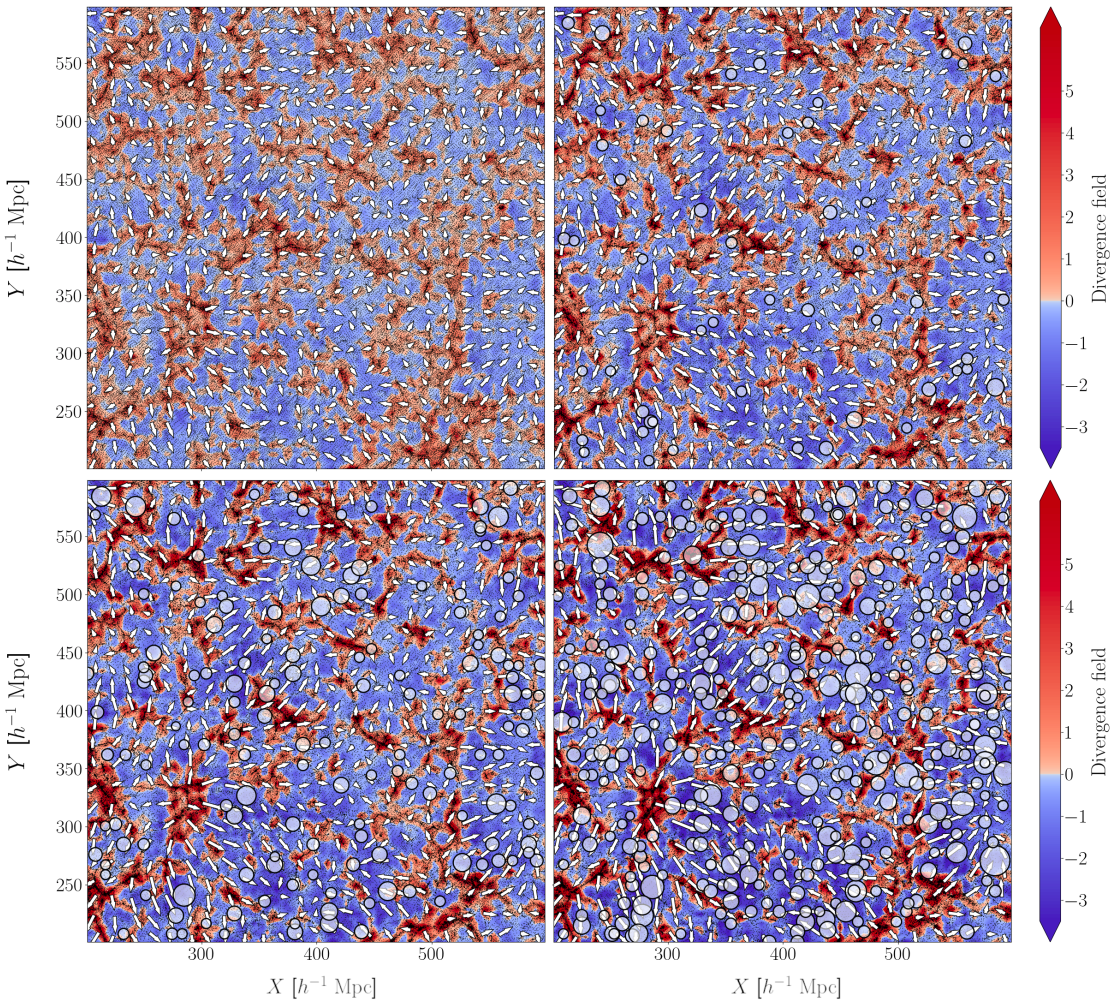

Figure 4.10: The image shows four thin slices (10 h^{-1} Mpc) of the DUSTGRAIN-pathfinder simulations with DM particles (10% subsample) in snapshot associated at different redshifts ($z = 2, 1, 0.5, 0$ from top left to bottom right). In particular, the positions of the DM particles (black dots), the value of the reconstructed divergence field (underdense regions in blue, overdense regions in red as shown by the coloured scale on the right), the mean back-in-time displacement field (through a grid of 25×25 arrows) and the voids identified in the sample (white circles) are shown. The displacement field was exactly reconstructed.

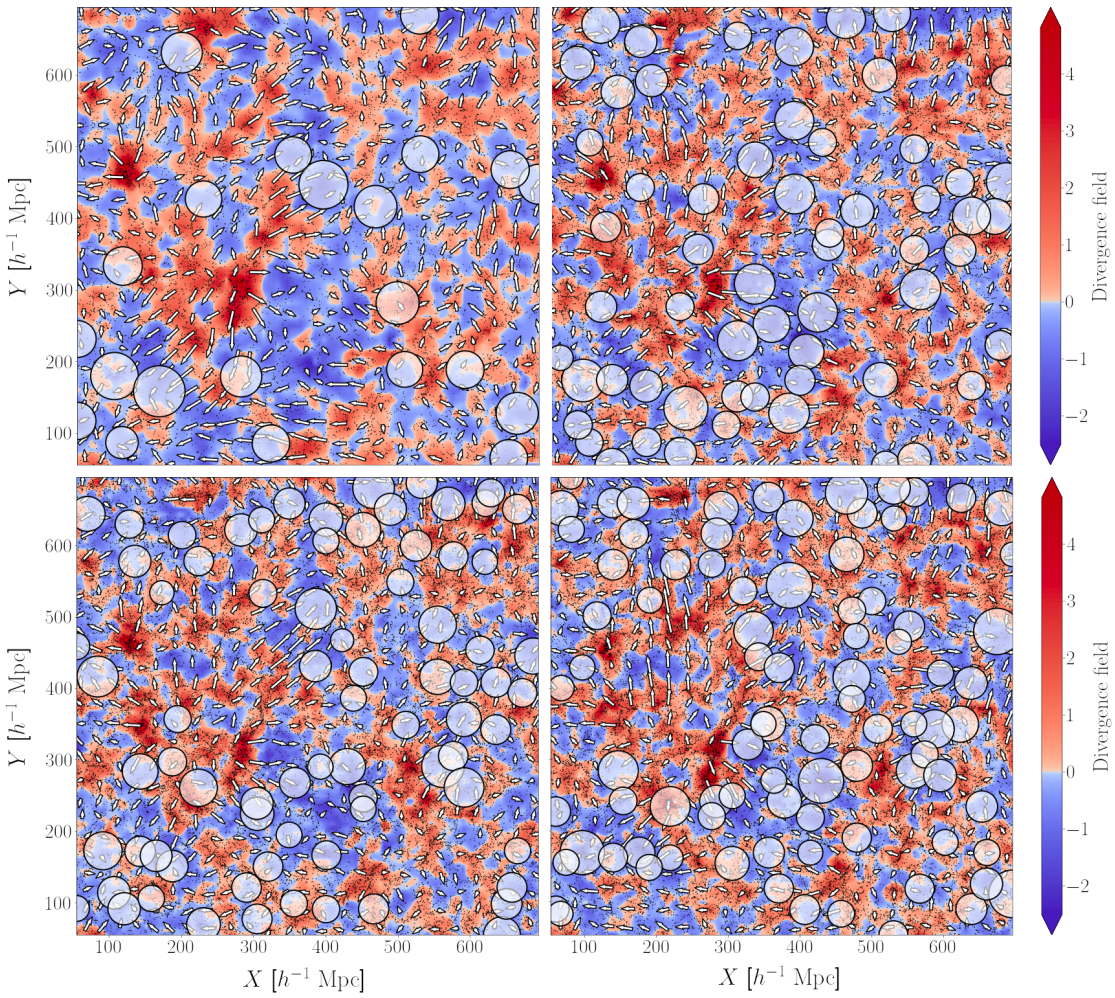

Figure 4.11: The same as Fig. 4.10 but for four slices thickness of (50 h^{-1} Mpc) with voids identified in DM haloes. The reconstruction of the displacement field was performed through the `LaZeVo` method.

4.5.1 Void profiles

According to the standard ΛCDM model, a self-similar phase during the cosmic void evolution is expected (Hamaus et al., 2014). This self-similar process ensures that the void structure can be characterised only by one parameter, the effective radius, $R_{\rm eff}$. Despite the wide range of void radii, it is possible to average the internal density distribution rescaling the radius to $R_{\rm eff}$, in order to investigate the common behaviour of different voids. Usually, in order to obtain a stacked density profile, the void sample is subdivided into groups, having similar effective radii. For this work, voids with $R_{\rm eff} > 2.75$ mps were considered. This value ensures we exclude any possible spurious voids. In addition, voids identified in the DM halo distribution were considered and their density profiles were measured in both this and DM distribution. Voids were matched to the Vdn model, assuming a threshold in the DM halo density field of $\delta_{\rm v,h} = -0.7$, and the profiles are computed through spheres of radius R centrered in voids. The measurement was repeated at different redshifts in order to obtain an estimate of the evolution of the profiles. In particular, void catalogues identified at z=[0, 0.5, 1, 2] were considered.

In Fig. 4.12 we can see that the average density profiles observed in the halo distribution are steeper than those measured in the DM, as predicted by the theory. Furthermore, again as expected, the bias between DM and haloes increases as the redshift increases. An indication of the good performance of the cleaner comes from the presence of the "loop"that forces the measured density profiles in the DM halo distribution to pass through the point $R/R_eff = 1 - \delta_{\rm v,h} = -0.7$. This indicates the effectiveness of the radius rescaling process of the voids at the chosen density threshold. Finally, the moderate scattering of the profiles around the median profile indicates how the centres of the voids are identified with good accuracy within the underdensities, minimising the effects due to void shape irregularity and tracer shot noise.

According to the Zel'dovich theory, in the linear regime an equality relation exists between divergence of the displacement field and the density contrast (Sec. 2.2.2). Fig. 4.13 shows the result of this measurement obtained on a sample of voids identified in the 0.4% DM distribution of the DUSTGRAIN-pathfinder simulations ($z = 0$, $\delta_{\rm v} = -0.7$) by interpolating the divergence field and integrating in R the values obtained, in a manner analogous to Eq. (3.20). It can be seen that there is a very close correspondence between the stacked profile of the density field and the stacked profile of the divergence, regarding the regions outside $R_{\rm v}$. In the central regions, the regime begins to move out of linearity and therefore the relationship is no longer valid. Using divergence profiles as a substitute of density profiles (e.g., in cleaning operations) has several advantages. Indeed, it can be shown that the divergence profiles are more regular than the density profiles, and are much less affected by mass discretisation. However, there are difficulties related to the construction of the displacement field. Any solution other than the exact solution possibly leads to different divergence values than the true density contrast. In the future, further studies will be needed to understand this issue more deeply and attempt to find a solution applicable to real data.

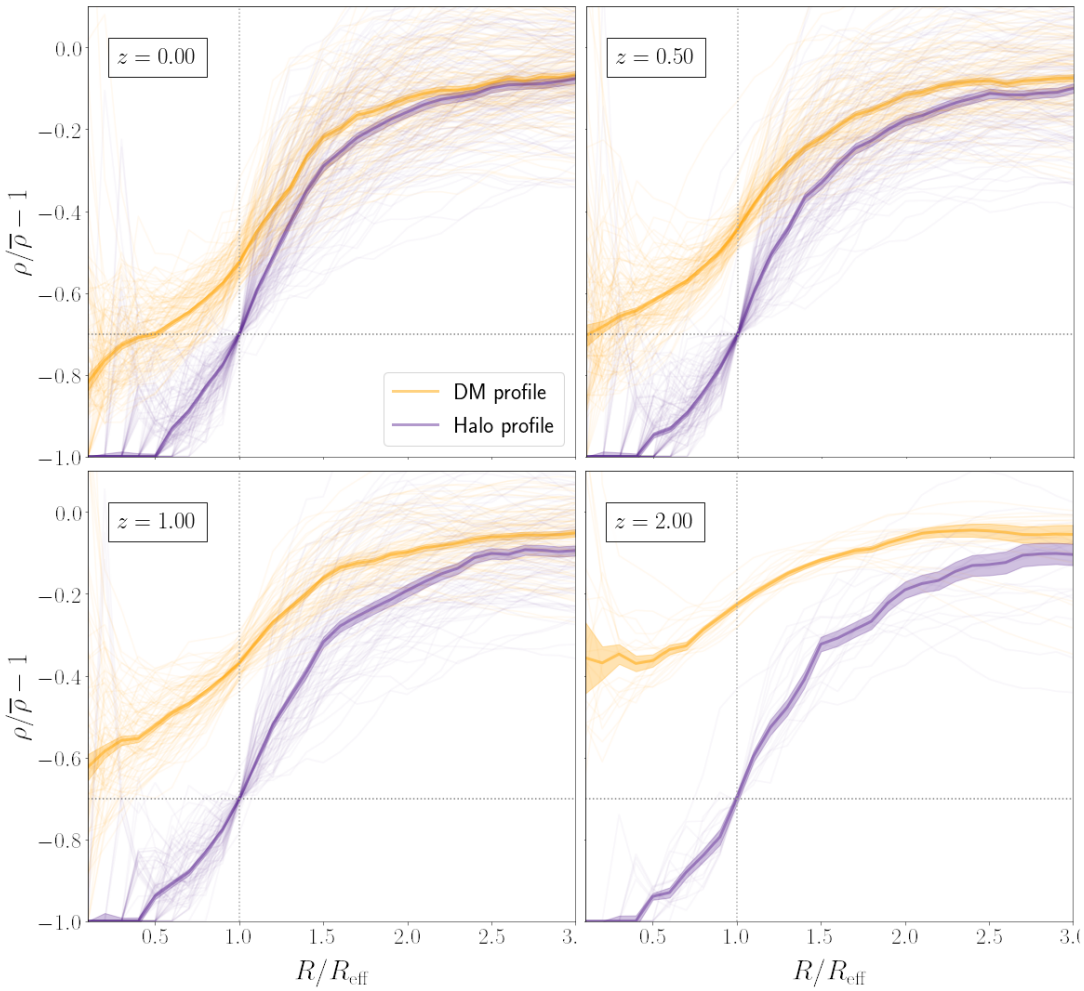

Figure 4.12: The density profiles of the voids identified in the distribution of DM haloes from the DUSTGRAIN-pathfinder simulation, at different redshifts ($z = 0,\ 0.5,\ 1,\ 2$ from top left to bottom right). We selected all voids with $R_{\rm v} > 2.75$ mps. The purple profiles are measured from the centres, identified in the manner previously described, in a 10% subsample of the DM distribution, while the orange profiles are measured in the DM haloes distribution. The thicker profiles represent the median values of the measured density profiles (stacked profiles).

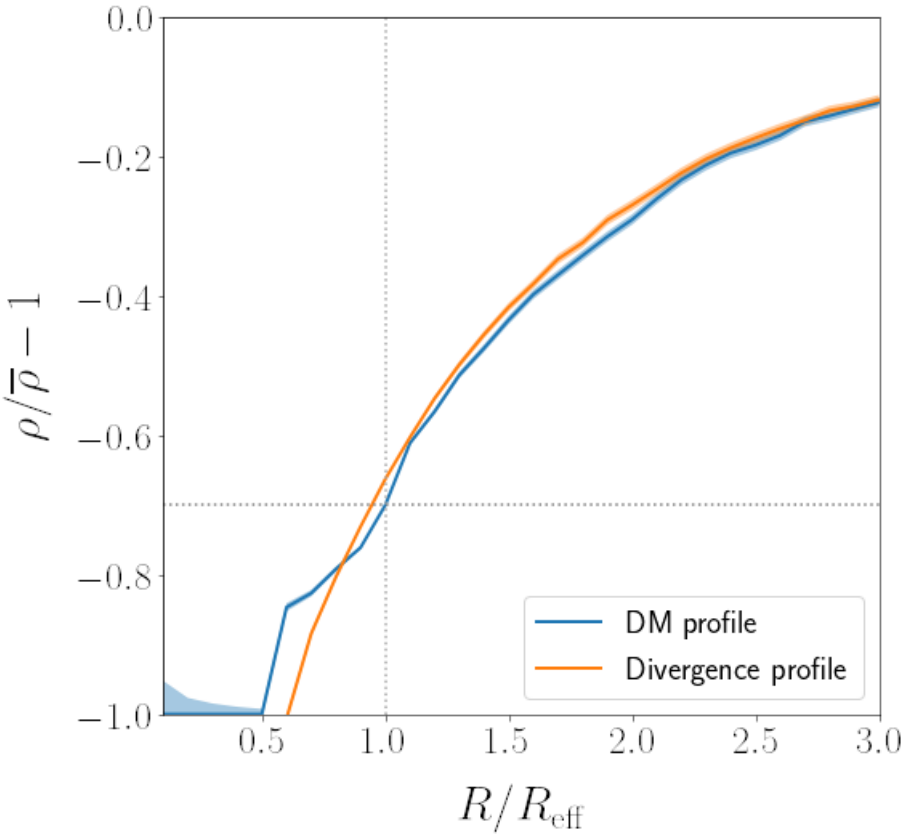

Figure 4.13: The density (blue) and the divergence (orange) stacked profiles of voids identified in the distribution of DM particles from a 0.4% subsample of the DUSTGRAIN-pathfinder simulations.

Chapter 5

BitVF-VIDE comparison

Once the proper functioning of BitVF and of the cleaning algorithm has been verified, we test the statistics of the void catalogue obtained with our algorithm against those derived through the use of VIDE, which is currently the void finder most widely employed in literature (see Sec. 3.1.2). The aim of the tests is to verify that our finder produces accurate enough void statistics, such as voids profiles (Sec. 5.1) and counts (Sec. 5.2), possibly better than those obtained through VIDE, in order to exploit the full potential of dynamic detection. This is mandatory to verify if BitVF can be used it for the exploitation of void statistics in constraining cosmological parameters (Sec. 5.3). As we will see, BitVF outperformes its competitor it in several cases, confirming the high potential of this algorithm.

5.1 Void profiles

Void density profiles represent one of the main statistics to describe these objects. Not only they provide information on the evolution of tracer distributions, when measured at various redshifts, but more importantly they can be exploited to constrain cosmological parameters. In particular, this is possible through the modelling of the dynamical distortions (Hamaus et al., 2020), which are induced by the tracer peculiar velocities changing the void apparent shapes, and also by applying the *Alcock-Paczynski test* (Alcock & Paczynski, 1979) to exploit these objects as standard spheres. Furthermore, the shape of the profiles gives us strong indications on the accuracy of the identification of the centre of the voids performed by the detection algorithms. Deeper and less scattered stacked profiles correspond to centres identified in the lower density regions.

Fig. 5.1 shows the stacked density profiles of (cleaned) voids identified at different redshifts ($z = 0,\ 0.5,\ 1,\ 2$) through the application of BitVF and VIDE. In particular, the profiles are rescaled to the effective radius of the voids and show the average density trend in the distributions of both DM haloes and DM particles. The profiles obtained through both algorithms match the theoretical expectations: they asymptotically reach $\delta = 0$ as the radius increases (we are dealing with profiles of average density, calculated in spheres centred on the centre of voids), they show the correct evolution of the bias between DM halo and DM particle distributions, and they cross the density threshold $\delta_{\rm v,h} = -0.7$ at which the effective radius was set. However, the profiles obtained through BitVF appear deeper than the VIDE ones,

especially in the regions outside the effective radius of the voids. Furthermore, the `BitVF` profiles appear less scattered in the central void regions. These two characteristics suggest that the centres of the voids identified by our new finder are located in more underdense regions than those identified by `VIDE`.

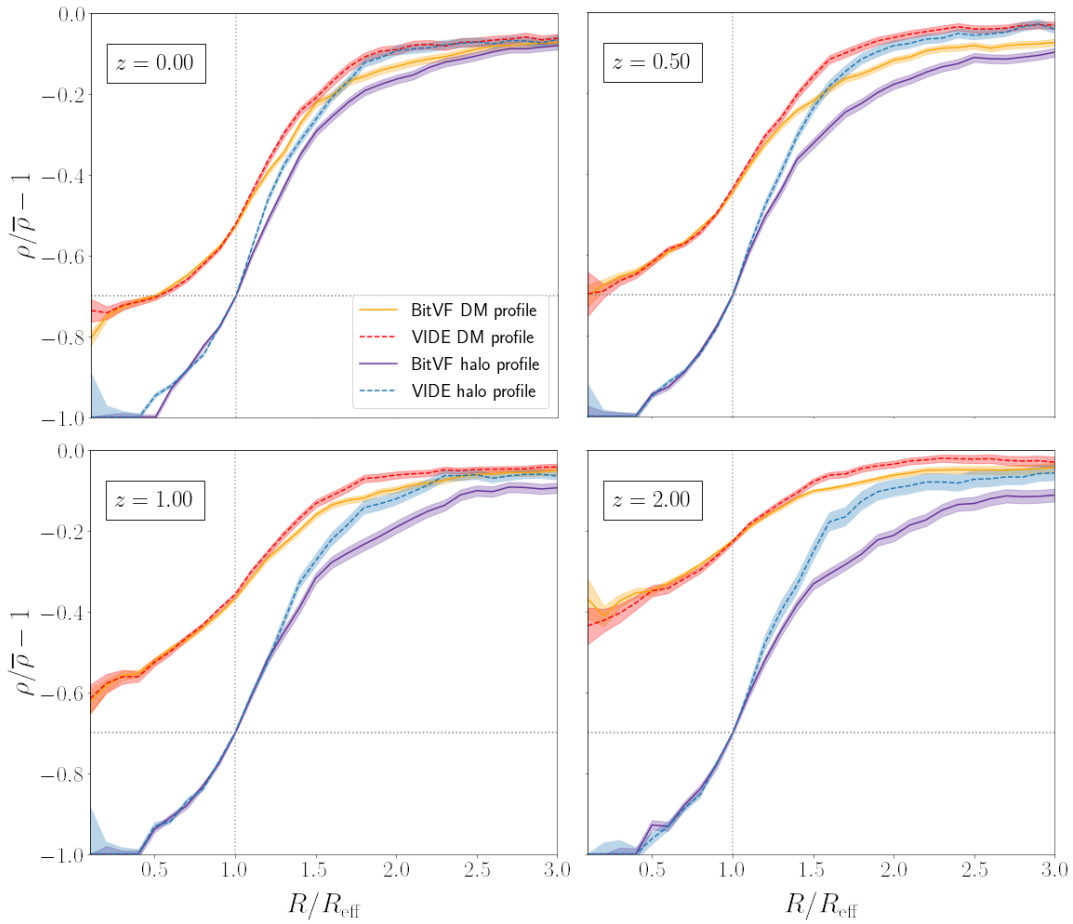

Figure 5.1: The stacked density profiles of the voids identified by `BitVF` and `VIDE` in the distribution of DM haloes and DM particles from the DUSTGRAIN-pathfinder simulation, at different redshifts ($z = 0,\ 0.5,\ 1,\ 2$ from top left to bottom right). We select all voids with $R_{\rm v} > 2.75$ mps. The red and the orange profiles are measured using a 10% subsample of the DM distribution, while the blue and the violet profiles are measured in the DM halo distribution. The shaded bands around the profiles represent the errors, calculated as the standard deviation of the density at each bin divided by the number of voids.

5.2 Void counts

Another powerful statistics to constrain cosmological parameters is provided by void counts, in particular the VSF. In addition, the VSF can be used to validate the accuracy of the void identification algorithms. In particular, it is possible to compare the theoretical VSF with the obtained counts, evaluating the residuals and the minimum radius below which the void catalogue is incomplete, a measure that gives us a precise indication of the spatial resolution.

Our void finder is tested by comparing the VSFs obtained in different contexts against `VIDE`. In particular, we test the response of both `BitVF` and `VIDE` to the identification of voids in DM distributions, at different redshifts and subsampling threshold, and in biased tracer distributions (DM haloes), again at different redshifts. Figure 5.2 shows the void counts obtained in the DM distribution at different subsampling values. Ensuring a stable response with increasing subsampling levels is important in order to guarantee good performances even in those cases where limited computing power is available, or in real cases where only data from a subpopulation of galaxies is provided. As can be seen, the radii at which the drop in the counts occurs are similar for both finders, indicating similar spatial resolution values for both the algorithms. However, `BitVF` appears more stable than `VIDE` as the subsample increases, and thus as the number of particles decreases, maintaining a consistent match between measured and theoretical VSF even at the highest subsample value considered (4% of total tracers). `VIDE`, on the other hand, shows a systematic overestimation of the counts in this case. Figure 5.3 shows the evolution of measured counts in a single DM subsample (10%) at different redshifts ($z = 0,\ 0.5,\ 1$). Although the void counts measured by `BitVF` are in better agreement with the theoretical model at large radii than the ones measured by `VIDE`, particularly for $z = 1$, it is noticeable that the spatial resolution worsens for our finder as redshift increases, which is not the case for `VIDE`. One justification for this behaviour might lie in the fact that the high redshift voids, particularly considering the DM distribution, are weakly underdense. `BitVF` operates ideally when pronounced underdensities are present, so that they can be more accurately identified as the points to which tracers flow as they move backward in time.

At this point, the analysis is extended to DM halo counts, to measure the difference in the response of the two finders when used on biased tracers, which more closely reproduce the real data than DM. Before this can be done, however, it is necessary to calibrate the relationship for bias between the distributions of DM and DM haloes, following the procedure outlined in Sec. 3.2.2. We expect different bias values for the two finders, particularly for the slope values, as the centres of the voids are identified in different ways. Figure 5.4 shows the calibration and the linear fits performed on the points in the $b_{\rm eff} - b_{\rm punct}$ plane, with the first value provided by theory and the second measured from the void profiles. As expected, almost identical values are obtained for the offset, but different values for the slope. Furthermore, the fact that the error on the fit measured on the `BitVF` data is lower than that obtained for the `VIDE` data, is a further confirmation of the fact that our finder identifies the centres of the voids with greater precision (since, trivially, lower errors are produced by a lower dispersion of the bias measured on each void, which is itself the result of a better positioning of the centres). The results obtained are then used to rescale

the theoretical VSF in order to predict the DM halo counts, according to the theory presented in Sec. 3.2.2. Finally, Fig. 5.5 shows the counts of voids identified in the DM halo distribution at different redshifts ($z = 0, 0.5, 1, 2$). The displacement field is reconstructed through the `LaZeVo` method. It is immediately noticeable how the counts obtained via `BitVF` are more consistent with the theoretical model and reach a better spatial resolution than the `VIDE` ones, for all redshifts considered. This test therefore shows how for biased tracers `BitVF` performs better than `VIDE`, both in terms of quality and quantity of the voids obtained. Furthermore, this confirms the goodness of the reconstruction of the displacement field using the `LaZeVo` method.

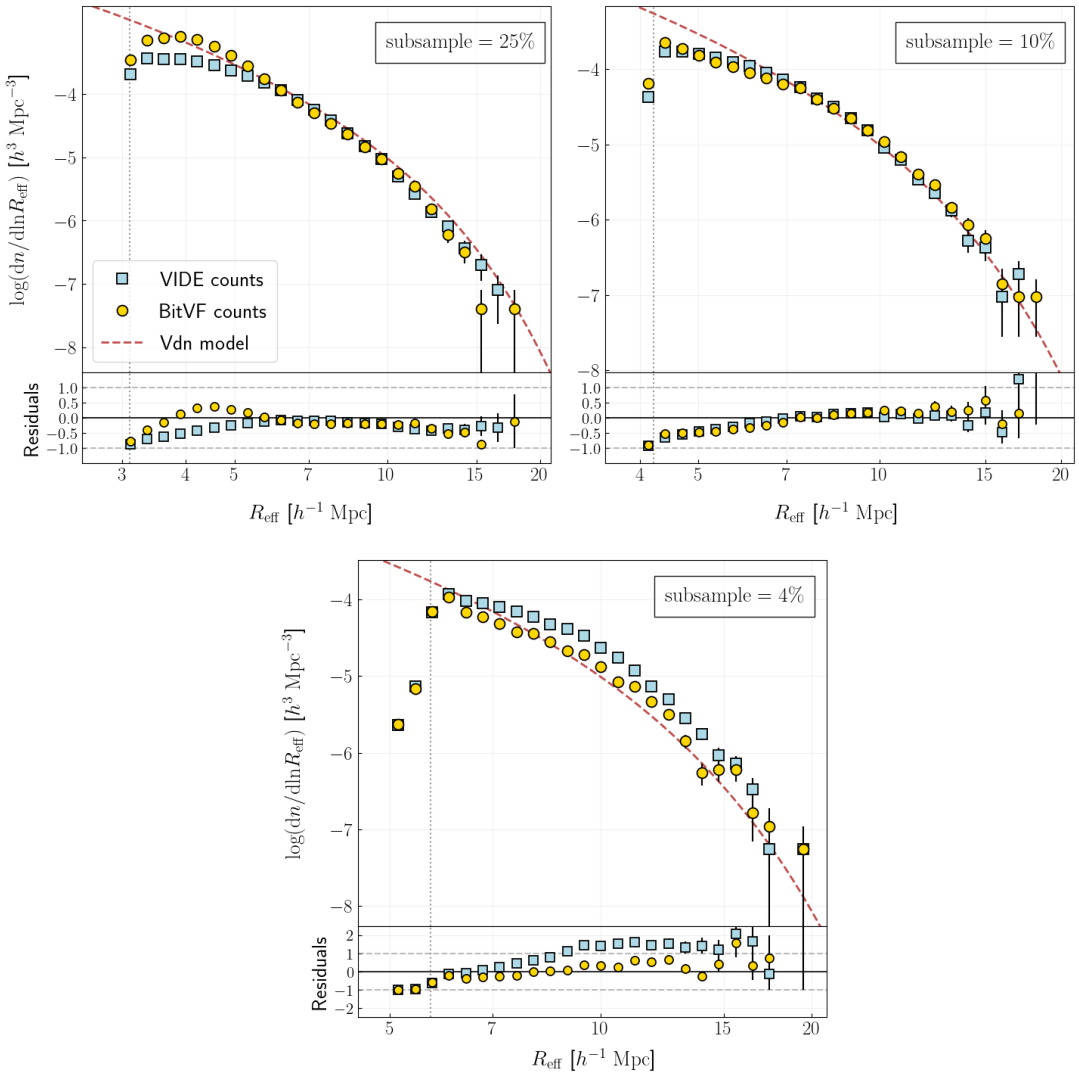

Figure 5.2: Comparison between the VSFs measured by `BitVF` (yellow circles) and `VIDE` (blue squares) with different subsamples (25%, 10%, 4% from top to bottom) of the DM distribution, from the DUSTGRAIN-pathfinder simulations, at $z = 0$, and the theoretical VSF (dashed red line) provided by the Vdn model. Voids are defined as in the standard Vdn model (thus $\delta_{\rm v} = \delta_{\rm sc}$). Errors on counts are assumed to be Poissonian and are shown by the black vertical bars. The lower part of each subplot shows the residuals, calculated as the ratio of the measure-model difference and the model. Finally, the vertical series of grey dots shows the theoretical minimum resolution, assumed to be equal to twice the mps of the tracers.

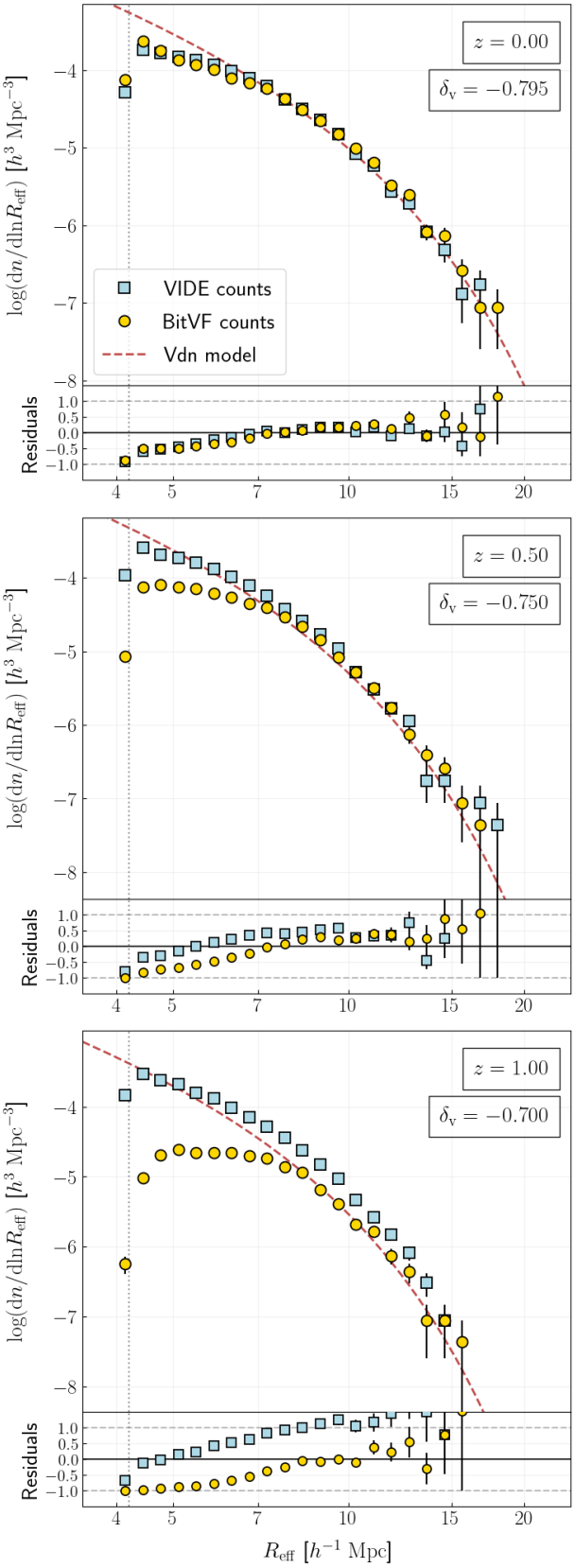

Figure 5.3: The same as in Fig. 5.2 but for voids identified at different redshifts ($z = [0,\ 0.5,\ 1]$ and thresholds, $\delta_v = [-0.795, -0.75, -0.7]$ from the top to the bottom) in the same DM subsample (10%).

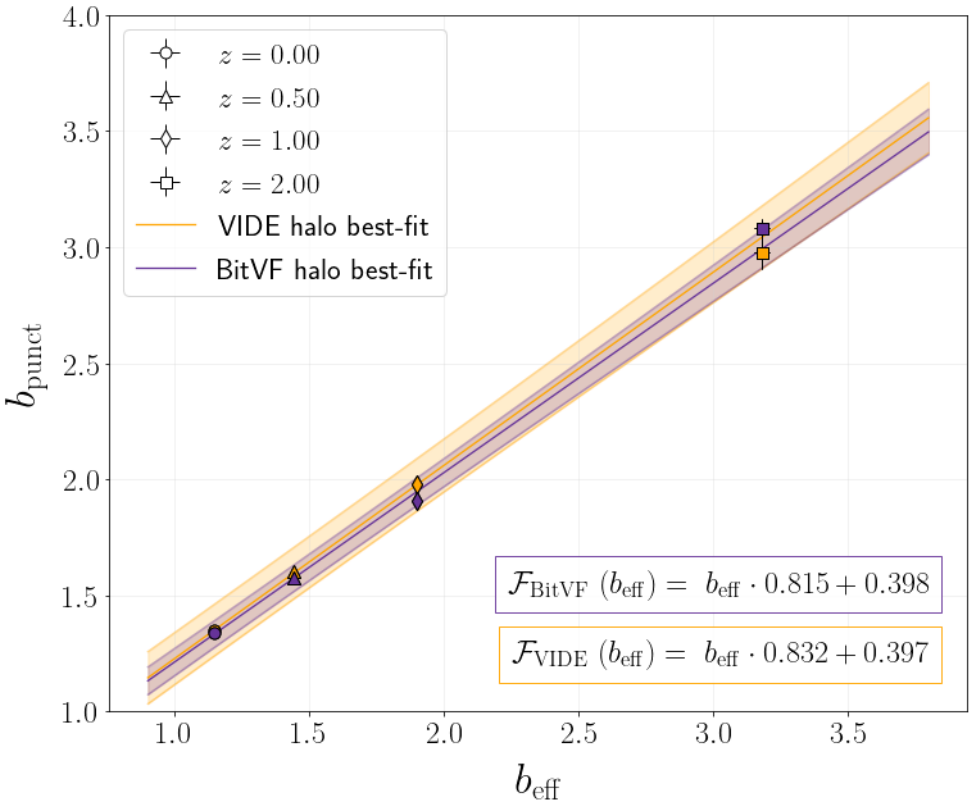

Figure 5.4: Calibration of the $\mathcal{F}(b_{\rm eff})$ relation, required for the conversion of the threshold, $\delta_{\rm v,tr} \to \delta_{\rm v,DM}$. The solid lines represent the best-fit linear relations $\mathcal{F}(b_{\rm eff})$ obtained with the calibrated coefficients $B_{\rm slope}$ and $B_{\rm offset}$, for `BitVF` (violet) and `VIDE` (orange) data. The shaded regions indicate an uncertainty of 1σ on the relationships. The markers represent the calibrations obtained for each bin of redshift, leaving $b_{\rm punct}$ as the only free parameter of the VSF model when fitting the measured counts. This calibration provides a value of $b_{\rm punct}$ for each redshift of the sample which is associated to the value of the effective bias, $b_{\rm eff}$, of the DUSTGRAIN-pathfinder galaxies at that specific redshift.

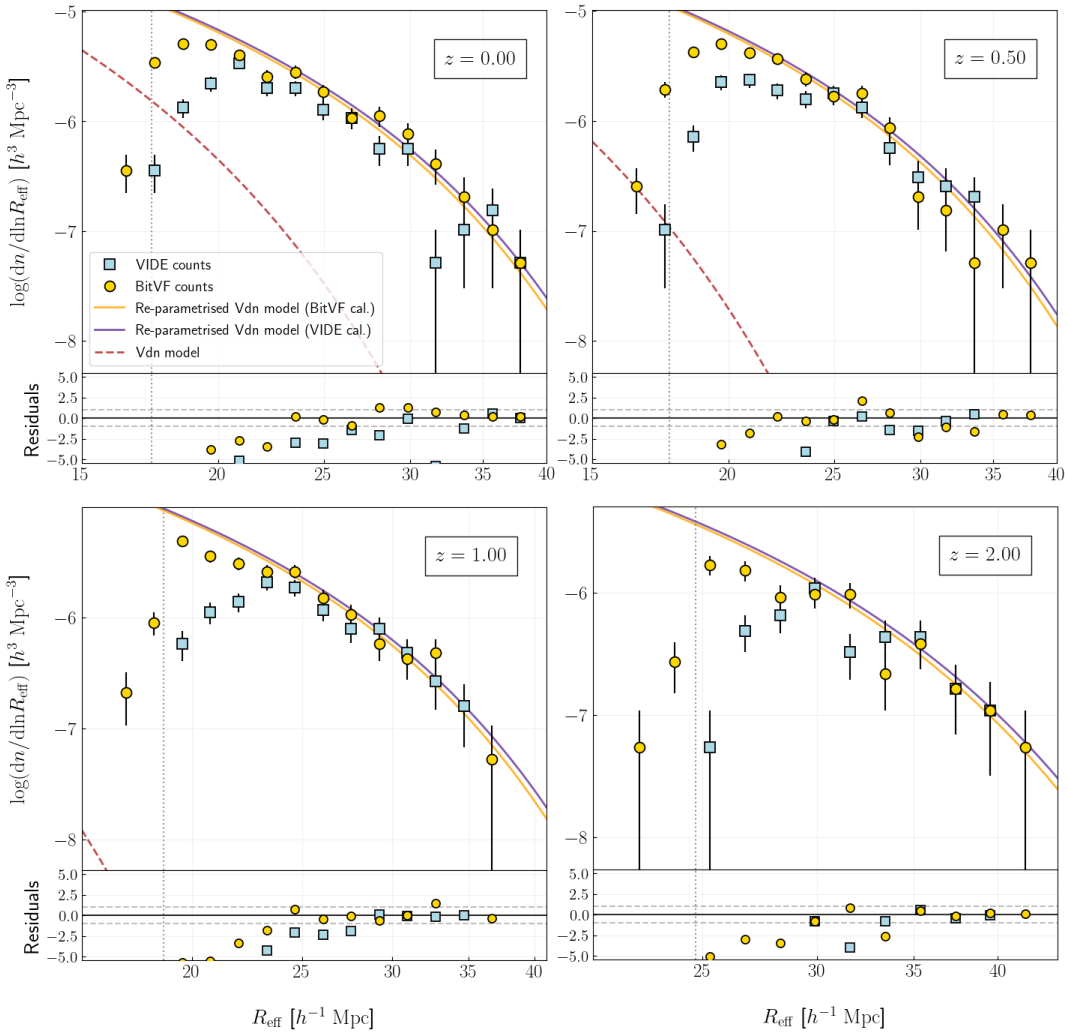

Figure 5.5: Comparison between the VSFs measured by `BitVF` (yellow circles) and `VIDE` (blue squares) at different redshifts ($z = 0,\ 0.5,\ 1,\ 2$ from top left to bottom right) of the DM haloes distribution derived from the DUSTGRAIN-pathfinder simulations, and the theoretical VSF (red dashed line) provided by the Vdn model, the rescaled theoretical VSFs calibrated for `BitVF` (orange) and `VIDE` (violet). Voids are rescaled at $\delta_{\mathrm{v}} = -0.7$. Errors on counts are assumed to be Poissonian and are shown by the black vertical bars. The lower part of each subplot shows the residuals, calculated as the ratio of the measure-model difference and the error. Finally, the vertical dotted grey lines show the theoretical minimum resolution, assumed to be equal to twice the mps of the tracers.

5.3 Cosmological parameter constraints

In this analysis, we want to test the reliability of the abundance of voids extracted with `BitVF` by computing the corresponding cosmological predictions attainable with the VSF modelling. Our goal is also to compare these results with those obtained with the samples of voids built by applying `VIDE` on the same DM halo catalogues, at the same redshifts. Therefore we consider both the sets of size function measured from `BitVF` and `VIDE` void samples, adopting the same radius range (based on the best resolution at each redshift), the same binning, and modelling these abundances with the corresponding extended Vdn model calibration (Contarini et al., 2019), presented in Fig. 5.4. We perform a Bayesian statistical MCMC analysis of the measured void size functions at different redshifts ($z = 0, 0.5, 1, 2$, see Fig. 5.5) by sampling the posterior distribution of the parameters σ_8 and Ω_m. We assume uniform prior distributions for σ_8 and Ω_m, and we fix the remaining cosmological parameters ($\Omega_\Lambda = 0.68655, h = 0.6731, A_\mathrm{s} = 2.199 \times 10^{-9}, n_\mathrm{s} = 0.9658$ and $z, b_\mathrm{eff}, B_\mathrm{slope}, B_\mathrm{offset}$ with values depending on the case in question) of the model to their true values. We also set with constant priors the model parameters b_eff, B_slope and B_offset. While b_eff is different for every redshift considered and its corresponding value is taken from Table 2 of Contarini et al. (2019), the parameters B_slope and B_offset take into account the redshift evolution of the VSF and are therefore unique for every sample of voids build with the same void finder. The results of this analysis are reported in Figs. 5.6, 5.7, 5.8, 5.9, at redshifts $z = 0,\ 0.5,\ 1,\ 2$, respectively. The values obtained, with their respective errors, are listed in Tab. 5.1. All four plots show how the real values of σ_8 and Ω_m fall within the 1σ confidence contours in the case of `BitVF`, whereas this is not the case for `VIDE`, especially for $z = 0$ and $z = 0.5$ (Fig. 5.6 and 5.7, respectively). This is not surprising, as `BitVF` exhibits a higher match between void counts and the theoretical VSF with respect to `VIDE` (see Sec. 5.2). It can be seen that the contours at $z = 0.5$ are offset for both finders. Although this is a significant anomaly, the contour obtained by `BitVF` in agreement within 2σ with the true value of the simulation. We will investigate in the future possible systematics not modelled in this analysis, related to the large-radius void defect observed in the second panel of the Fig. 5.5. In addition, we present the joint cosmological constraints obtained for both the void samples, by combining the results at all the analysed redshifts, and assuming the measured void abundances as independent (Figs. 5.10, 5.11 5.12). As shown, the point on the $\sigma_8 - \Omega_\mathrm{m}$ plane describing the real values of the two parameters is intercepted within the 68% contours obtained from the voids identified with `BitVF`. The contours obtained through `VIDE`, on the other hand, are slightly offset. This test shows that `BitVF` void statistics allow us to constrain the cosmological parameters more accurately than in the `VIDE` case, due to a more precise identification of the voids, resulting in a lower loss of counts at low scales. Recovering these scales is crucial for modelling VSF, and thus for having higher constraining power. Finally, in Fig. 5.13, the values of parameter S8 obtained by count analysis are shown. We can see that, again, the exact value falls within the 68% confidence range for BitVF, while the result obtained by VIDE is much more shifted. The `BitVF` algorithm is therefore a promising candidate for detecting voids to be used for constraining cosmological parameter values, especially in prevision of future use with real data.

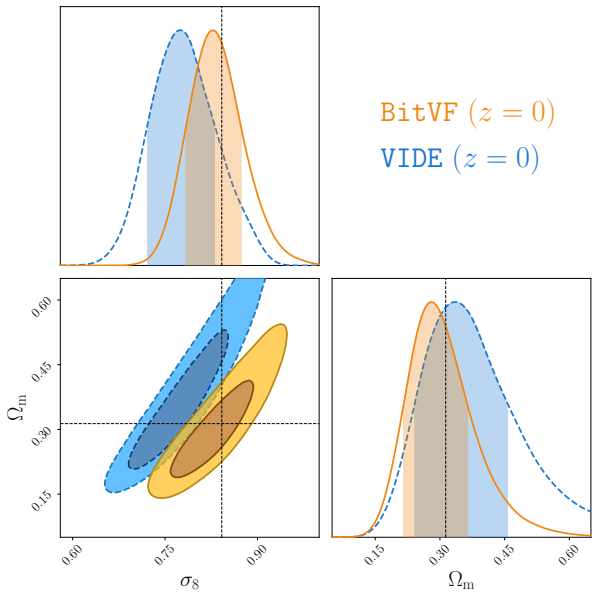

Figure 5.6: 68% and 95% confidence regions in the $\sigma_8 - \Omega_{\rm m}$ plane, for the halo catalogues with $M_{\rm min} = 2.43 \times 10^{12}\ h^{-1}\ {\rm M}_\odot$ from ΛCDM, $z = 0$, DUSTGRAIN-pathfinder snapshot. The colour corresponds to the different finders: orange for `BitVF`, blue for `VIDE`. The prior distributions for σ_8 and $\Omega_{\rm m}$ are uniform. The marginalised 1D posterior distributions of σ_8 and $\Omega_{\rm m}$ are shown in the top and bottom-right panels of each plot, respectively. The black lines represent the true DUSTGRAIN-pathfinder ΛCDM values ($\sigma_8 = 0.842$ and $\Omega_{\rm m} = 0.31345$).

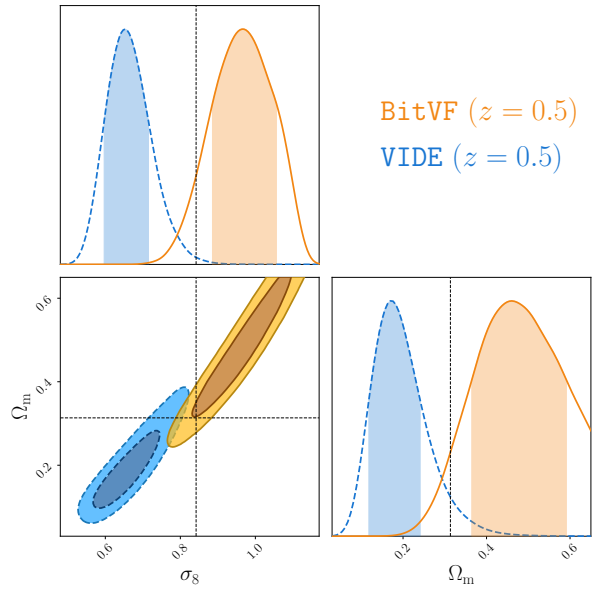

Figure 5.7: As Fig. 5.6, but for $z = 0.5$.

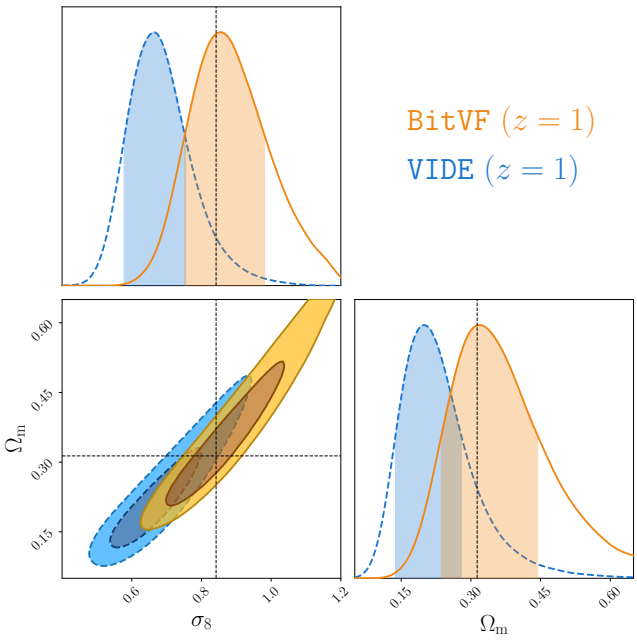

Figure 5.8: As Fig. 5.6, but for $z = 1$.

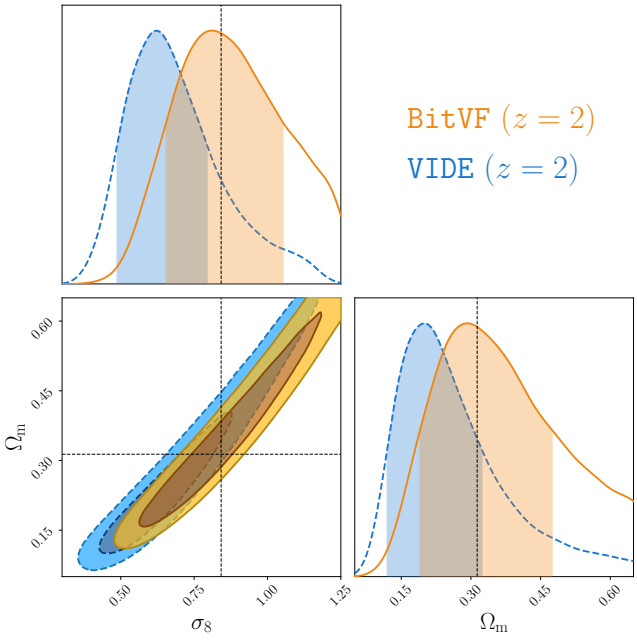

Figure 5.9: As Fig. 5.6, but for $z = 2$.

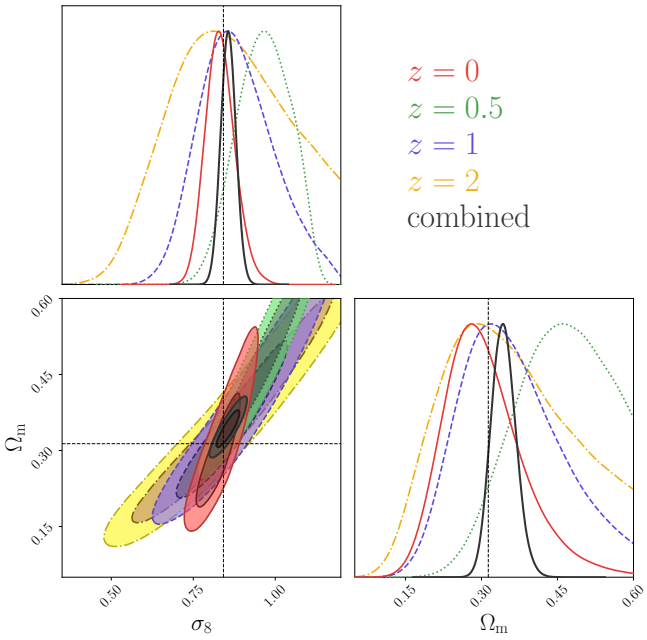

Figure 5.10: 68% and 95% confidence regions in the $\sigma_8 - \Omega_{\rm m}$ plane, for the voids identified in the halo catalogues with $M_{\rm min} = 2.43 \times 10^{12}\ h^{-1}\ {\rm M}_\odot$ from ΛCDM, DUSTGRAIN-pathfinder snapshot, with voids identified by `BitVF`. The colour of ellipses corresponds to different redshifts: red for $z = 0$, green for $z = 0.5$, blue for $z = 1$ and yellow for $z = 2$. The black ellipse corresponds to confidence levels for the combined distributions. The prior distributions for σ_8 and $\Omega_{\rm m}$ are uniform. The marginalised 1D posterior distributions of σ_8 and $\Omega_{\rm m}$, for each redshift, are shown in the top and bottom-right panels of each plot, respectively. The black lines represent the true DUSTGRAIN-pathfinder ΛCDM values ($\sigma_8 = 0.842$ and $\Omega_{\rm m} = 0.31345$).

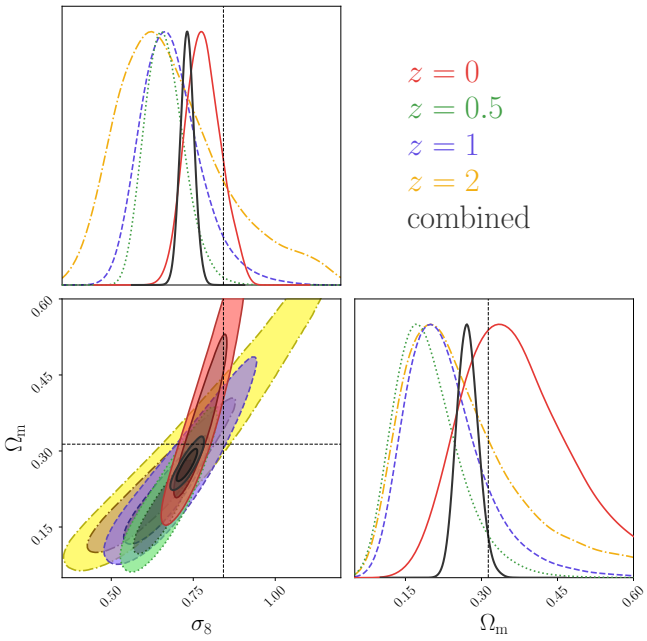

Figure 5.11: Same as Fig. 5.10, but for voids identified with `VIDE`.

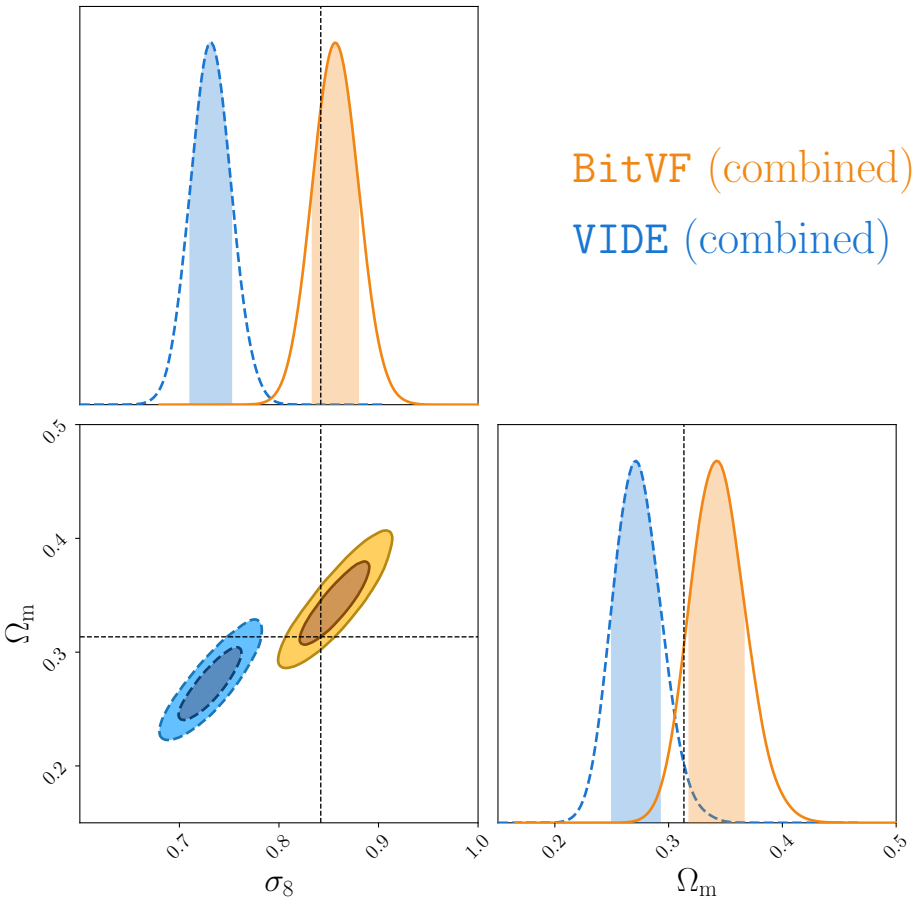

Figure 5.12: 68% and 95% confidence levels for the halo catalogues with $M_{\rm min} = 2.43 \times 10^{12}\ h^{-1}\ {\rm M}_\odot$ from ΛCDM, DUSTGRAIN-pathfinder snapshot, with voids identified by `BitVF`, obtained by combining the results at all the analysed redshifts ($z = 0,\ 0.5,\ 1,\ 2$) and assuming the measured void abundances as independent. The colour of ellipses corresponds to the different finders: orange for `BitVF`, blue for `VIDE`. The marginalised 1D posterior distributions of σ_8 and $\Omega_{\rm m}$ are shown in the top and bottom-right panels of each plot, respectively. The black lines represent the true DUSTGRAIN-pathfinder ΛCDM values ($\sigma_8 = 0.842$ and $\Omega_{\rm m} = 0.31345$).

z	BitVF		VIDE	
	σ_8	Ω_m	σ_8	Ω_m
0	0.83 ± 0.04	$0.28^{+0.08}_{-0.06}$	$0.77^{+0.06}_{-0.05}$	$0.33^{+0.12}_{-0.09}$
0.5	$0.97^{+0.09}_{-0.08}$	$0.46^{+0.13}_{-0.10}$	0.65 ± 0.06	$0.17^{+0.07}_{-0.05}$
1	$0.86^{+0.12}_{-0.11}$	$0.32^{+0.13}_{-0.08}$	0.66 ± 0.09	$0.20^{+0.08}_{-0.06}$
2	$0.81^{+0.25}_{-0.15}$	$0.29^{+0.19}_{-0.10}$	$0.62^{+0.17}_{-0.13}$	$0.20^{+0.12}_{-0.08}$
combined	0.86 ± 0.02	0.34 ± 0.02	0.73 ± 0.02	0.27 ± 0.02

Table 5.1: Summary of the cosmological constraints computed on the parameters σ_8 and Ω_m from the modelling of the size function of voids identified with 2 different finding codes: `BitVF` and `VIDE`. We report in the first column of the table the redshift of the DUSTGRAIN-pathfinder snapshot considered in the analysis. In the second and third columns we present the results obtained on σ_8 and Ω_m with `BitVF`, while in the fourth and the fifth columns the analogous results obtained with `VIDE`, respectively. In the last row of the table we report the cosmological predictions computed by combining the constraints at different redshifts, assuming each data set as independent.

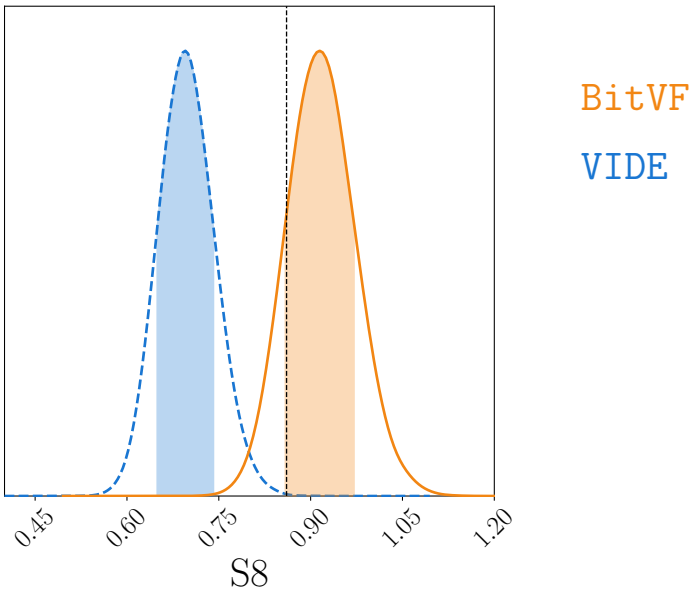

Figure 5.13: Cosmological constraints on the parameter $S8 = \sigma_8\sqrt{\Omega_m/0.3}$ computed from the abundance of voids identified in the DUSTGRAIN-pathfinder simulations at redshifts $z = 0, 0.5, 1, 2$. In orange we report the result obtained by modelling the number counts of voids identified with `BitVF` and in blue that obtained by using voids identified with `VIDE`. The shaded areas under the curves represent an uncertainty of 1σ on the cosmological predictions. We indicate with a black dashed line the true value of the simulation, $S8 = 0.86$.

Chapter 6

Conclusions and future perspectives

In this Chapter, we will summarise the scientific problem from which we started (Sec. 6.1), and then provide an overview of the work performed, showing the main advantages and issues of the method developed (Sec. 6.2). Finally, the future perspectives and new opportunities arising from this work will be outlined (Sec. 6.3).

6.1 The scientific problem

The current standard cosmological model, ΛCDM, is based on the solution of Einstein's field equations, assuming the validity of the Cosmological Principle, i.e. that on large scales the Universe is homogeneous and isotropic, and the validity of the Einstein's General Relativity (see Chapter 1). Furthermore, the model introduces two new cosmological components: the *cold dark matter*, a non-collisional and non-relativistic component that accounts a large majority of the matter, and the *cosmological constant* Λ, which dominates the Universe in terms of energy density and produces the present accelerated expansion observed through various probes, such as Type Ia supernovae (Riess et al. 1998, Schmidt et al. 1998, Perlmutter et al. 1999), the cosmic microwave background (Eisenstein et al., 2005) and of the large-scale structures of the Universe (Planck Collaboration 2016, Planck Collaboration 2020). At present, many open problems still exist, of both theoretical and observational nature, such as the exact nature of these two fundamental components and the tension between the current expansion rate of the Universe, H_0, when extrapolated from the early Universe or measured in the late Universe (Bernal et al. 2016, Moresco & Marulli 2017).

Cosmic voids have emerged in recent years as a viable alternative to traditional cosmological probes to investigate the validity of the ΛCDM model. The statistics of these objects, can help to break the degeneracies that arise in the determination of cosmological parameters by providing information on volumes of the Universe poorly sampled by the standard probes. However, as we explained in Sec. 3.1, the identification of voids is not trivial, due to their underdense nature and the lack of a formal definition. Over time, several identification methods have been proposed by the scientific community, as we discussed in Sec. 3.1. We can classify these methods into three different classes, according to the detection criteria: *density*, *geometrical* and *dynamical* (Lavaux & Wandelt, 2009) criteria. The first two classes

of methods directly reconstruct the density field from the position of the tracers. This leads to the problems of shot noise and the discretisation of the field sampled by the tracers. The dynamic class of methods, on the other hand, uses the tracers to reconstruct the velocity field, which is subsequently used to obtain the density field in an indirect manner through the calculation of its divergence. In this way, continuous and smoothed fields are obtained, which are less affected by the typical uncertainties associated with void identification, allowing a more precise detection and providing more accurate statistics, such as void counts and profiles (see Sec. 3.1 and Sec. 3.2).

6.2 Procedure and results

The main objective of this Thesis was the development of a new dynamic void finder, with the purpose of overcoming the most common difficulties associated with the void identification. Starting from the work of Elyiv et al. (2015), we have thus developed a dynamic void identification method, called `BitVF`, presented in Chapter 4. The `BitVF` algorithm aims at reconstructing the linear back-in-time displacement field of the tracers in order to obtain the density contrast field through the Zel'dovich theory (see Sec. 2.2.2). This our new algorithm identifies voids through the execution of three fundamental steps: the reconstruction of the displacement field (presented in Sec. 4.1.1), the computation of the density field (presented in Sec. 4.1.2) and the identification of voids (presented in Sec. 4.1.3). In the first step, tracers are used to sample the back-in-time displacement field between their current position and their position at the time when the perturbation formed. The reconstruction can be achieved using two different methods: `LaZeVo` and the `exact reconstruction`. The first method consists of a statistical reconstruction of the displacement field based on the action minimisation. The second method is intended for use in simulations only and provides the exact reconstruction of the displacement field using the identification numbers of the tracers. Once the displacement field has been reconstructed, the density contrast field is obtained through Equation (2.88). At this point the centres of the voids are identified as the local minima of the density field in the underdense regions. The position of the centres is corrected through a weighted mean of the local values of the density field. Finally, the void radius is computed as the radius of a sphere containing a certain fraction of the average density of the Universe. At this point, it is possible to clean up the void catalogue obtained with `BitVF` to allow the following cosmological exploitation of the acquired data (Sec. 3.2.2). A new version of the cleaning algorithm, developed by expanding and improving on Ronconi & Marulli (2017) work, was presented in Sec. 4.2.

The developed algorithms were tested on the DUSTGRAIN-pathfinder simulations (Sec. 4.3), with different distributions of cosmological tracers (DM particles and DM haloes) at different redshifts ($z = 0,\ 0.5,\ 1,\ 2$). The description of this work is given, together with the results obtained, in Sec. 4.5. As a first step, `BitVF` is applied to the distribution of DM with the aim of testing the algorithm through the exact reconstruction of the displacement field. The results obtained show that the reconstruction of the divergence field is in perfect agreement with the tracer field. In addition, voids are accurately identified in the underdensity regions. As a second

step, it was chosen to compare the LaZeVo method with the exact reconstruction of the displacement field in order to validate the method. The result of the application reveals a very good agreement between the two methods, justifying the future use of this statistical reconstruction. BitVF was then applied to the distribution of DM haloes at $z = 0$, reconstructing the displacement field using the LaZeVo method. Subsequently, the operation was repeated with all four available redshifts, both for the distributions of DM haloes and DM particles (in this case employing the exact reconstruction of the displacement field), with the aim of testing the algorithm response to more or less linear density distributions. In all the cases mentioned, it was shown that the algorithm can reconstruct the displacement and divergence fields with great accuracy, perfectly reconstructing the density field described by the cosmological tracers. In addition, it was proved that it correctly identify voids in underdense areas, reflecting the differences we would expect in the distributions of voids measured in the different distributions of tracers, at various redshifts. Finally, in Sec. 4.5.1, the density and divergence profiles of voids obtained from the BitVF application are shown. In particular, through the stacked profiles obtained, it is possible to test the correct functioning of the cleaning algorithm, as well as the quality of the void catalogue obtained.

Once we had verified the correct functioning of our void finder, we proceeded to compare it with VIDE, the most widely used void identification algorithm. The results of these tests are shown and discussed in Chapter 5. In particular, in Sec. 5.1, the stacked density profiles obtained by applying the two finders to the same sample of DM particles at different redshifts are compared. The profiles obtained from voids identified with BitVF are deeper and with an associated lower scatter, providing a strong indication that the void centres found are in more underdense positions than the void centres identified by VIDE. Section 5.2 shows the void counts obtained by applying the two finders to both tracer distributions, testing the effect of redshift and different level of subsampling (the latter tested on DM only). In addition, we discussed the calibration process required to extend the theoretical VSF provided by the Vdn model to predict void counts identified using biased tracers. The counts obtained through the application of the finders on the DM distribution show a better response of BitVF than VIDE as the subsampling increases, while showing a poorer spatial resolution at high redshifts. This is at present time the major issue of our finder, which arises because the underdensities in the DM distribution are not very pronounced at high redshifts, shortening the particle displacements and thus decreasing the finder resolving power. However, we point out that the void counts at higher redshifts also remain in better agreement with the theoretical predictions than the VIDE ones. Furthermore, it can be observed that the void counts obtained by applying BitVF to the DM halo distribution are far better than the VIDE ones, both in terms of quantity and spatial resolution, at all redshifts. This feature is extremely valuable, as it proves the great potential of this method when applied to biased tracers, i.e. tracers of the same type as we expect to have with real data. Finally, as a last test, in Sec. 5.3 we presented an attempt to derive the values of σ_8 and $\Omega_{\rm m}$ values of the DUSTGRAIN-pathfinders simulations by means of a Bayesian analysis, performed by running MCMC on the VSF counts measured in the halo distributions at various redshifts. For each redshift, BitVF shows a more accurate parameter estimate than VIDE. In particular, analysing the results obtained by combining the

distributions calculated at all redshifts, BitVF intercepts the true values within the 68% confidence contours, while the VIDE contours are offset. This result is due to the choice of setting common parameters for VSFs, such as bin number and minimum radius of the voids considered. BitVF, showing a better spatial resolution than VIDE, recovers information at smaller scales, constraining more effectively the VSF. The offset of the VIDE contours can be resolved by increasing the minimum radius of the voids considered, and thus decreasing the resolution, at the price of obtaining wider confidence contours and thus less precise estimates.

In conclusion, BitVF showed the potential to provide precisely identified void catalogues that can be used effectively for obtaining both void count statistics and density profiles. Furthermore, it has been proven to be extremely effective when applied to biased tracer distributions, recovering voids at better spatial resolutions than VIDE, which leads to more precise constraints on cosmological parameters. BitVF has therefore all the potential to be considered as a quality alternative to existing void detection methods, possibly providing a fundamental tool for future cosmological analysis.

6.3 Future perspective

The results obtained in this Thesis work demonstrate the efficacy of the BitVF algorithm when applied to simulations. However, this method is intended for application to real data, particularly on data from the upcoming large galaxy surveys. In order to make the algorithm suitable for any type of catalogue, we will work on the following aspects:

- we will make the algorithm more efficient from the point of view of the computational cost, in particular with regard to the statistical reconstruction of the displacement field. In fact, the LaZeVo method, unlike the exact reconstruction which has performances comparable to VIDE, can handle at most catalogues on the order of millions of tracers. In anticipation of the real data that will be produced in the coming years (for instance, observation of the order of more than tens of millions of galaxies with the Euclid survey), it is necessary to optimise the algorithm in order to exploit the available data in their entirety.

- Once the LaZeVo method has been optimised, it will be necessary to test it on larger simulations, again using both DM distributions and biased tracers.

- We will implement a version of the finder capable of handling light cones and $RA - Dec - z$ observed coordinate catalogues. The operating principles of BitVF will remain similar, but the redshift dependency of the tracer density must be managed.

- Once light cones can be handled, the use of masks allowing the reconstruction of complex observation fields and the management of the weights associated with the various observations will be implemented, in order to be able to manage the real data in every possible situation.

- Finally, it would be appropriate to further study the divergence profiles, both from a theoretical and a practical point of view. In the first case, by searching for a calibration that would allow these profiles to be aligned with the density profiles even in the central zones of the voids, in the second case by finding a statistical reconstruction of the displacement field that would maintain the reconstructed total distance of displacement of the tracers at the same value as the real one (see Sec. 4.5.1). Once this has been done, a new radius assignment and a new version of the cleaner could be implemented exploiting the divergence profiles instead of the density profiles, definitively eliminating any direct dependence on the position of the tracers.

The codes developed in this Thesis are included in the `CosmoBolognaLib` (Sec. 4.4) and will be made public with the next release.

Bibliography

Aarseth S. J., Hoyle F., 1963, Monthly Notices of the Royal Astronomical Society, 126, 223

Abdalla E., et al., 2022, arXiv e-prints

Alcock C., Paczynski B., 1979, Nature, 281, 358

Barreira A., Cautun M., Li B., Baugh C., Pascoli S., 2015, Journal of Cosmology and Astroparticle Physics, 2015

Bernal J. L., Verde L., Riess A. G., 2016, Journal of Cosmology and Astroparticle Physics, 2016, 019

Bernardeau F., 1994, The Astrophysical Journal, 427, 51

Bianchi E., Rovelli C., 2010

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, The Astrophysical Journal, 379, 440

Bos E. G. P., Weygaert R., Dolag K., Pettorino V., 2012, Monthly Notices of the Royal Astronomical Society, 426

Bosma A., 1999, in Merritt D. R., Valluri M., Sellwood J. A., eds, Astronomical Society of the Pacific Conference Series Vol. 182, Galaxy Dynamics - A Rutgers Symposium. p. 339

Cai Y.-C., Padilla N., Li B., 2015, Monthly Notices of the Royal Astronomical Society, 451, 1036

Colberg J. M., et al., 2008, Monthly Notices of the Royal Astronomical Society, 387, 933

Coles P., Lucchin F., 2002, Chichester: Wiley, —c1995

Collaboration L. D. E. S., 2012

Contarini S., Ronconi T., Marulli F., Moscardini L., Veropalumbo A., Baldi M., 2019, Monthly Notices of the Royal Astronomical Society, 488, 3526

Contarini S., Marulli F., Moscardini L., Veropalumbo A., Giocoli C., Baldi M., 2021, Monthly Notices of the Royal Astronomical Society, 504, 5021

Contarini S., et al., 2022

Crocce M., Scoccimarro R., 2008, Physical Review D, 77, 023533

Croft R., Gaztanaga E., 1997, Monthly Notices of the Royal Astronomical Society, 285, 793

Despali G., Giocoli C., Angulo R. E., Tormen G., Sheth R. K., Baso G., Moscardini L., 2016, The Monthly Notices of the Royal Astronomical Society Journal, 456, 2486

Di Valentino E., et al., 2021, Classical and Quantum Gravity, 38, 153001

Edwin H., 1929, Proceedings of the National Academy of Sciences, 15, 168

Einstein A., 1915, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), pp 778–786

Eisenstein D. J., et al., 2005, The Astrophysical Journal, 633, 560

Ellis R., 2010, Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, 368, 967

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, Monthly Notices of the Royal Astronomical Society, 448, 642

Falck B. L., Neyrinck M. C., Aragon-Calvo M. A., Lavaux G., Szalay A. S., 2012, The Astrophysical Journal, 745

Famaey B., McGaugh S. S., 2012, Living Reviews in Relativity, 15, 10

Forero-Romero J., Hoffman Y., Gottloeber S., Klypin A., Yepes G., 2008, Monthly Notices of the Royal Astronomical Society, 396

Foster C., Nelson L. A., 2009, The Astrophysical Journal, 699, 1252

Galli S., Iocco F., Bertone G., Melchiorri A., 2011, Physical Review D, 84, 027302

Garilli B., et al., 2014, Astronomy and Astrophysics, 562, A23

Gilman D., Birrer S., Treu T., Keeton C. R., Nierenberg A., 2018, Monthly Notices of the Royal Astronomical Society, 481, 819

Giocoli C., Tormen G., van den Bosch F. C., 2008, Monthly Notices of the Royal Astronomical Society, 386, 2135

Giocoli C., Baldi M., Moscardini L., 2018, Monthly Notices of the Royal Astronomical Society, 481, 2813

Green J., et al., 2012

Gunn J. E., Gott J. Richard I., 1972, The Astrophysical Journal, 176, 1

Guth A. H., Pi S. Y., 1982, Physical Review Letters, 49, 1110

Hamaus N., Sutter P. M., Wandelt B. D., 2014, Physical Review Letters, 112

Hamaus N., Pisani A., Choi J.-A., Lavaux G., Wandelt B. D., Weller J., 2020, Journal of Cosmology and Astroparticle Physics, 2020, 023

Hildebrandt H., et al., 2016, Monthly Notices of the Royal Astronomical Society, 465, 1454

Hoyle F., Vogeley M., 2004, The Astrophysical Journal, 607, 751

Icke V., 1984, The Monthly Notices of the Royal Astronomical Society Journal, 206, 1P

Ishiyama T., et al., 2021, Monthly Notices of the Royal Astronomical Society, 506

Jennings E., Li Y., Hu W., 2013, Monthly Notices of the Royal Astronomical Society, 434

Kaiser N., 1984, The Astrophysical Journal, Letters, 284, L9

Kauffmann G., Fairall A. P., 1991, Monthly Notices of the Royal Astronomical Society, 248, 313

Kreisch C., Pisani A., Villaescusa-Navarro F., Spergel D., Wandelt B., Hamaus N., Bayer A., 2021

Kuhn H. W., 1955, Naval Research Logistics Quarterly, 2, 83

Laureijs R., et al., 2011

Lavaux G., Wandelt B., 2009, Monthly Notices of the Royal Astronomical Society, 403

Lavaux G., Wandelt B. D., 2012, The Astrophysical Journal, 754, 109

Lee J., Park D., 2009, The Astrophysical Journal, 696, L10

Lemaître A. G., 1931, Monthly Notices of the Royal Astronomical Society, 91, 483

Marulli F., Veropalumbo A., Moresco M., 2016, Astronomy and Computing, 14, 35

Massara E., Villaescusa-Navarro F., Viel M., Sutter P., 2015, Journal of Cosmology and Astroparticle Physics, 2015, 018

Micheletti D., et al., 2014, Astronomy & Astrophysics, 570, A106

Moresco M., Marulli F., 2017, Monthly Notices of the Royal Astronomical Society: Letters, 471, L82

Mukherjee S., et al., 2020, arXiv e-prints, p. arXiv:2009.14199

Nadathur S., et al., 2020, Monthly Notices of the Royal Astronomical Society, 499, 4140

Neyrinck M. C., 2008, Monthly Notices of the Royal Astronomical Society, 386, 2101

Neyrinck M. C., Gnedin N. Y., Hamilton A. J. S., 2005, Monthly Notices of the Royal Astronomical Society, 356, 1222

Orlin J., B J., Lee Y., 1993, Massachusetts Institute of Technology (MIT), Sloan School of Management, Working papers

Perico E. L. D., Voivodic R., Lima M., Mota D., 2019, Astronomy & Astrophysics, 632

Perlmutter S., et al., 1999, The Astrophysical Journal, 517, 565

Pisani A., Sutter P. M., Hamaus N., Alizadeh E., Biswas R., Wandelt B. D., Hirata C. M., 2015, Physical Review D, 92, 083531

Pisani A., et al., 2019

Planck Collaboration 2016, Astronomy & Astrophysics, 594, A13

Planck Collaboration 2020, Astronomy and Astrophysics, 641, A6

Platen E., Van De Weygaert R., Jones B. J. T., 2007, Monthly Notices of the Royal Astronomical Society, 380, 551

Potter D., Stadel J., Teyssier R., 2017, Computational Astrophysics and Cosmology, 4, 1

Press W. H., Schechter P., 1974, The Astrophysical Journal, 187, 425

Puchwein E., Baldi M., Springel V., 2013, Monthly Notices of the Royal Astronomical Society, 436

Riess A. G., et al., 1998, Astronomical Journal, 116, 1009

Riess A. G., Casertano S., Yuan W., Macri L. M., Scolnic D., 2019, The Astrophysical Journal, 876, 85

Rindler-Daller T., 2020, European Journal of Physics, 41, 035602

Rintoul M. D., Torquato S., 1997, Journal of Colloid and Interface Science, 186, 467

Ronconi T., Marulli F., 2017, Astronomy & Astrophysics, 607

Schmidt B. P., et al., 1998, The Astrophysical Journal, 507, 46

Shandarin S. F., Zeldovich Y. B., 1989, Reviews of Modern Physics, 61, 185

Sheth R. K., Van De Weygaert R., 2004, Monthly Notices of the Royal Astronomical Society, 350, 517

Springel V., 2005, The Monthly Notices of the Royal Astronomical Society Journal, 364, 1105

Suto Y., Sato K., Sato H., 1984, Progress of Theoretical Physics, 71, 938

Sutter P., et al., 2014a, Astronomy and Computing, 9

Sutter P., Lavaux G., Hamaus N., Wandelt B. D., Weinberg D. H., Warren M. S., 2014b, Monthly Notices of the Royal Astronomical Society, 442, 462

Tormen G., Moscardini L., Yoshida N., 2004, Monthly Notices of the Royal Astronomical Society, 350, 1397

Van de Weygaert R., 2014, Proceedings of the International Astronomical Union, 11, 493

Van de Weygaert R., Platen E., 2011, International Journal of Modern Physics: Conference Series, 01, 41

Verde L., Treu T., Riess A., 2019, Nature Astronomy, p. 891–895

Verza G., Pisani A., Carbone C., Hamaus N., Guzzo L., 2019, Journal of Cosmology and Astroparticle Physics, 2019, 040

Von Hoerner S., 1960, Zeitschrift fur Astrophysik, 50

Weinberg S., 1989, Reviews of Modern Physics, 61, 1

Woodfinden A., Nadathur S., Percival W. J., Radinović S., Massara E., Winther H. A., 2022

Zel'dovich Y. B., 1970, Astronomy and Astrophysics, 5, 84

Zeldovich Y. B., 1972, Monthly Notices of the Royal Astronomical Society, 160, 1P

Zentner A. R., 2007, International Journal of Modern Physics D, 16, 763