SISSA

**Scuola
Internazionale
Superiore di
Studi Avanzati**

Physics Area - Ph.D. course in
Astrophysics and Cosmology

New avenues for investigating
the Large-Scale Structure of our Universe

Candidate:
Hasti Khoraminezhad

Advisors:
Carlo Baccigalupi
Matteo Viel

Academic Year 2020-21

Abstract

This thesis deals with characterizing the behavior of dark components (Dark energy and Dark matter relative to the baryons) in cosmology by following new avenues for investigating the Large-Scale Structure (LSS). Near future surveys would probe the structure formation with remarkable precision in order to constrain cosmological parameters and deepen our understanding of the nature of these components.

In the first part, with a general phenomenological view on the dark energy component of the Universe, we study the behavior of a perturbed Early dark Energy (EDE) model as an additional energy component in the early Universe involving the sound speed and anisotropic stress. We investigate the impact of EDE on cosmological observables such as the Cosmic Microwave Background (CMB) angular power spectrum as well as the linear matter power spectra. We mainly focus on the quantitative exploration of an extended parameter space including the mass of neutrinos and tensor to scalar ratio in the light of recently available data sets. As we will show, the constraints on the EDE parameters are remarkably stable even when Σm_ν, and r parameters are varied.

In the second part, we concentrate on the physics of the matter components of the Universe by a direct calculation of the coupling of baryons to the Cold Dark Matter (CDM) components. We perform 2-fluid gravity-only N-body simulations and assess the impact of relative baryon-CDM density perturbations in dark matter halo distribution which is usually neglected in LSS studies. Specifically, we focus on the baryon fraction in halos as a function of mass and large-scale baryon-CDM perturbations, which allows us to study the details of the nontrivial numerical setup required for such simulations as well. By quantifying the impact of such perturbations on halo-halo power spectra we found this effect can be degenerate with the one of massive neutrinos in near future and operating LSS surveys.

Finally, we investigate the statistics of various promising LSS probes in configuration space in gravity-only 2-fluid N-body simulations mentioned above. This allows us to study the impact of baryon-CDM perturbations on these statistics. Particularly, we focus on the statistics of the cosmic voids, as well as on the matter 2-point correlation function and Baryon Acoustic Oscillations (BAO) peak as a robust cosmological standard ruler. We find the impact of $1 - 2\%$ level at maximum on the void size function which is more prominent at higher redshifts, while the void density profile and void bias are roughly unaffected. Our results confirm the impact of baryon-CDM perturbations on cosmological constraints from the BAO feature in current and future galaxy surveys should be negligible at low redshift.

List of Publications

This thesis is the result of the research done during my Ph.D. in collaboration with Matteo Viel, Carlo Baccigalupi, Titouan Lazeyras, Pauline Vielzeuf, Maria Archidiacono, Raul E. Angulo and Oliver Hahn. It is based on the following papers:

- Hasti Khoraminezhad, Matteo Viel, Carlo Baccigalupi and Maria Archidiacono, *Constraints on the Spacetime Dynamics of an Early Dark Energy Component*, *JCAP* **07** (July, 2020) 039, [arXiv:2001.10252].

- Hasti Khoraminezhad, Titouan Lazeyras, Raul E. Angulo, Oliver Hahn and Matteo Viel, *Quantifying the impact of baryon-CDM perturbations on halo clustering and baryon fraction*, *JCAP* **03** (Mar., 2021) 023, [arXiv:2011.01037].

- Hasti Khoraminezhad, Pauline Vielzeuf, Titouan Lazeyras, Carlo Baccigalupi and Matteo Viel, *Cosmic voids and BAO with relative baryon-CDM perturbations*, submitted to *MNRAS*, [arXiv:2109.02949]

Acknowledgements

time amazed us with his wonderful pizza recipes. Giulio, with his always calm smile and a desk full of goaleadors. Natalia, it was fantastic to see her with a wonderful unicorn and an office full of delicious food. Anto, who didn't like using WhatsApp at all and I had to send him another message for the football call. Miguel, the great philosopher who played Risiko really well. Lotte, with the wonderful chocolate cakes, and Jon Pham, the amazing physicist.

My lovely Iranian friends in Trieste, Mina, Mandana, Nilufar, Sogand, Zainab, Sahra, Anahita, Najme, and Fereshte for being this much amazing, supportive, and kind to me.

Riccardo Iancer, for all his support since the time I arrive at SISSA up to now. Elenora, Lucia, Silvia, Patrizia, Gabriella, "Lucho", Andrea thanks for your kindness, your smiles, your help, thanks for the "Capo in B", "Nero", "Brioche" every single day. Thanks for your wonderful memory that could remember all our names. With you, I really felt I'm at home.

The "cosmology circle" wonderful friends, whom I learned a lot sharing time with them, back in time in Iran and now in different parts of the world, Daash-Abbas, who could make incredible visual perceptions of Gaussian curvature. Aghileh, who was always a pioneer in helping others. Amirnezam, that I could never distinguish if he is speaking or reciting poetry. Alireza, who could write a script in python to work/eat/sleep instead of him. Arash, the best god ever in online Mafia games in quarantine. Farbod, that was wonderful talking about science with him. Farida, that I was amazed by her wonderful memory. Farnik, I will never forget singing with him in Trieste during ICTP cosmology schools. Fateme, full of peace and calmness. Golshan, full of positive energy and happiness. Hajar, the wise friend that was joyful talking with her. Mahdi, that I will never forget the wonderful travel he arranged for us to Sangdeh. Mahmoodreza who could explain psychological theories in a simple way, gracious Mona, who was incredibly kind and peaceful to everyone. M.Reza, that was fantastic to have a scientific discussion with him. Sara, with always a kind smile and Sedighe with her always smart quotes in the group.

Our amazing football team, Sadaf, Diksha, Saby, Catlin, Digant, Gabri-Perfetto, Gabri, Andrea, Anto, Shrikanth, Elena, Andres, Cesar, Mina, Mandana, Ludy, Pier, Niki, Mattia, Viplove, Sebastian, Memo, Chile, Nicole, Miguel, and Regis.

Many thanks to Diletta, Nora, Davide, and little Ettore for being this kind. I would also like to thank nano-Mina for her always supportive and kind attitude, Tiago and Phoung, Elena, Andres and Cesar for all gatherings and happy moments.

Special thanks to Fariba Nameni, the best therapist ever, for supporting me and helping me all these years to know myself even if it was painful. I owe you this happy moment.

I would like to extend my sincere thanks to Clemencia, Alberto, Danitza, and Fernando, for their kindness, support, and love and for showing me how Colombia is beautiful.

Nobody has been more important to me in the pursuit of this thesis than the members of my family, my Mom, and Dad, the ones who were filling me with their unconditional love, support, and encouragement. My little brother, Fartash, my wonderful companion in life, that there is no way to repay his kindness. Thanks for always believing in me. I know how hard is the fact that I was never there during these years.

Finally, my sumercecito, Kevin, who provides me endless inspiration. Thanks for being an inexhaustible source of happiness and for daily reminding me how life is beautiful.

یادگاری که درین گنبد دوار بماند

Credit: By *Hafez*, A persian poet of 14th century, which literaly means:
"A relic that remains in the revolving dome!"

To Mom & Dad
who endured my absence
during these years

Table of Contents

List of Figures

List of Tables

1

Introduction

Understanding the fact that our Universe is expanding and it was much hotter and denser at early times compared to now, together with Einstein's discovery of General Relativity (GR), allowed us to go deeper through fundamental physical questions. Nowadays, we are facing a modern version of these fundamental questions, for instance: "how did the different elements of the Universe form?" or "what are the initial conditions of our Universe?", "why our Universe is so smooth at early times?" and "how can we have these clumpy non-linear structures from a smooth Universe?". These sort of questions could have quantitative answers in which can be tested against cosmological observations. Indeed during the last hundred years our understanding of the Universe has radically changed. In the beginning of the previous century, scientists were arguing whether the spiral nebulae were a part of our Galaxy or if they were distant "island Universes". But soon after it became clear that these galaxies were actually extra-Galactic systems. This evidence inspired the study by Edwin Hubble who measured the distance to Andromeda nebula using the period-luminosity relation of Cepheid variable stars in 1924 [16]. Only 5 years later, in 1929, Hubble made another essential contribution to cosmology by proving that galaxies are indeed receding from us by a velocity proportional to their distance [17]. The development of the General Theory of Relativity by Albert Einstein [18] in 1915 and 1916, enabled us for the first time in the history to establish a convincing and testable theory of the Universe linking the content of the Universe to its geometry via the Einstein field equations. A couple of years later, the solutions of these equations for a homogeneous Universe have been found by Aleksander Friedman [19, 20], describing the geometry and the expansion of the Universe as a function of the energy its content. (see chapter 2, Sec. 3.1 for more details.)

Cosmological Concordance Model

The expansion implies that our Universe at very early times must have been very dense and hot in such a way that radiation and the baryonic matter were coupled together via Compton scattering. The expansion caused the baryon-photon fluid to cool and the ionized Hydrogen to recombine and finally made the scattering inefficient and allowed the photons to free-stream through the Universe. This diffused background radiation, known as the Cosmic Microwave Background was predicted by Alpher & Herman in 1948 [21], they estimated the black-body radiation of CMB with a temperature roughly about $5\,\mathrm{K}$. CMB radiation has been detected by Penzias & Wilson in 1965 [22], which has confirmed the thermal history of the Universe. Further observations of the CMB established that the radiation was uniform over the sky with an accuracy of 10^{-5}, followed the Planck spectrum with a remarkable precision, with a black-body temperature of roughly $\sim 2.7\,\mathrm{K}$ [23].

Moreover, the prediction of the chemical abundances of helium and other light isotopes created during the primordial nucleosynthesis [24] was matched accurately by the measured abundances. At that time an upper limit was calculated for the mean baryon density which indeed was above the observed abundance of ^{2}H and ^{3}He that was generated in the Universe. This in turn could constrain the baryonic matter content of the Universe to $\lesssim 1/10$ of the required amount of the critical density would be needed for a flat cosmological geometry. The study of the gravitational collapse of perturbations in the homogeneous background has been the way to model the large amount of structures in the Universe. From the Nucleo-Synthesis constraints on the mean baryon density and due to the photon-baryon fluid pressure which can prevent the growth of baryonic perturbations

before recombination, together with the upper limits on the amplitude of the perturbations at the time of recombination using the CMB temperature fluctuations, we would face a problem to explain cosmological structures at present. The solution to solve this problem was to introduce a form of non-baryonic dark matter which only interacts gravitationally with the baryonic matter in the Universe [25]. Fluctuations in the dark matter distribution were able to grow during the radiation-dominated era and provided potential wells, into which the baryonic matter would fall into after recombination, lowering the expected temperature fluctuations in the CMB. Indeed, rotation curves of galaxies and cluster dynamics also proved the existence of dark matter as the dominating form of matter in the Universe and as the main driver for galaxy formation [26].

The determination of the specific luminosity profile of Type Ia supernovae [27] and the fact that they were standardized was instrumental in estimating cosmological distances paved the way to measure Hubble's law with higher accuracy and allowed the detection of the accelerated expansion of the Universe in 1998 [28, 29]. The accelerated expansion can be determined by the cosmological constant Λ in Einstein's equations which is also known as dark energy which is represented of the 70% of the energy content of the Universe today.
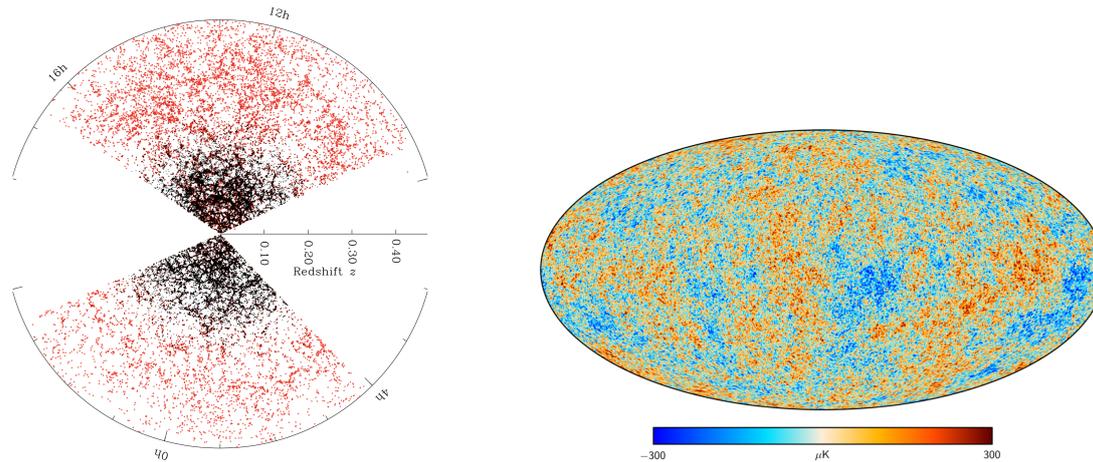
Figure 1.1: **Left panel:** The distribution of galaxies measured in the Sloan Digital Sky Survey (SDSS), more precisely, those within a slice of ± 3 deg of the celestial equator. Each dot depicts the position of a galaxy, different colors denote different galaxy samples. Black: magnitude-limited main sample. Red: Luminous red galaxy (LRG) sample. The total volume probed by the survey is 14.5 Gpc^3 in North Galactic Cap (NGC) and South Galactic Cap (SGC). (Credit: Michael Blanton and the SDSS Collaboration [1]). **Right panel:** A full sky map of the anisotropies of the CMB temperature observed by Planck satellite. The colored spots on the map correspond to photon energies at the time last scattering of photons by the electrons. The areas of higher energy are bluer, while the areas of low energy are redder. (Credit: image courtesy of ESA and Planck Collaboration [2])

Later, the accelerated cosmological expansion was also measured from the acoustic peaks in the CMB by the Wilkinson Microwave Anisotropy Probe (WMAP) [30], (moreover, other pre-WMAP suborbital CMB experiments can be found in LAMBDA NASA[1] website.) and the measurements of the clustering statistics from galaxy distribution of redshift surveys like 2dF Galaxy Redshift Survey [31] that in turn allowed to measure baryon acoustic oscillations at low redshift. The successful agreement between these results from totally independent cosmological probes has led to the establishment of the ΛCDM model as the standard model of the Universe, composite of a cosmological constant Λ, Cold Dark Matter, i.e. non-relativistic at decoupling time, with structures growing out of an initial Gaussian and scale invariant spectrum. After that, the improved measurements, such as the high precision measurements of the CMB anisotropy by Planck [2] (see Fig 1.1-right panel) or high precision measurements of the Baryon Acoustic Oscillations using clustering data with new galaxy catalogues by SDSS-BOSS/eBOSS [32, 33, 34, 35] (Fig 1.1-left panel), allowed us to measure the model parameters with high accuracy, which is sometimes would be referred to as the "era of precision cosmology".

---

[1]https://lambda.gsfc.nasa.gov/product/

Open problems

Despite the successes of the ΛCDM concordance model of cosmology, which provides us a remarkable fit to the available cosmological data, the picture of our Universe is still incomplete and there is a poor understanding the nature of the largest components (in terms of energy densities) of the Universe at a fundamental level: dark matter and dark energy. As we will discuss in chapter 2, Sec. 1, we assume the main ingredients of the ΛCDM concordance model (dark matter [36, 37], dark energy, as the cosmological constant (CC) which its energy density is constant in time and space, [28, 29] and inflation [38, 39, 40, 41, 42]) in their simplest ("vanilla") form. However, ΛCDM suffers from some problems. Among them, we can mention the cosmological constant problem which includes the fine-tuning problem and the cosmic coincidence problem. The fine-tuning problem refers to the fact that since the energy scale of the cosmological constant required for the cosmic acceleration is very small relative to that predicted by particle physics as the energy density of the vacuum (roughly about 10^{121} times), there must be some contribution from the fundamental theory which adds an additional contribution to the vacuum energy with the value incredibly close to the vacuum energy density. The coincidence problem states that the cosmological constant value is not only at odds with all possible fundamental energy scales and requires therefore fine-tuning, but also this particular value is of the same order of a totally unrelated number, the present matter energy density. On the other hand, by increasing the number and sensitivity of the experimental observations, discrepancies among some cosmological parameter emerged and if they are not due to the systematic errors, they could indicate a lack of success in the concordance model [43, 44, 45]. Among these tensions discussed extensively in the literature, are *i)* the tensions indicated between the Planck data in the context of the ΛCDM model [2] and the local determination of the Hubble constant for instance "Supernovae H0 for the Equation of State of dark energy" (SH0ES) Team: [46], which is the most statistically significant tension among the others (4σ to 6σ disagreements); *ii)* the weak lensing experiments [47, 48, 49, 50, 51] calculating the $S_8 \equiv \sigma_8\sqrt{\Omega_m/0.3}$ parameter[2] which is roughly in $\sim 3\sigma$ tension with Planck data [53]. *iii)* the Planck internal lensing anomalies linked to the extra amount of lensing in the temperature power spectrum, creating a tension between the cosmological parameters extracted in the high-ℓ and low-ℓ multipole ranges: $A_{\text{lens}} > 1$ [3] roughly about 2.8σ [2, 54]; *iv)* small-scale problems related to structure formation processes and the nature of dark matter or new physics [43].

Thesis content

It is a titanic work to face all open issues above, and in this thesis we focus on a few amongst the big ones, represented by the behaviour of the dark components, energy and matter in relation to baryons:

First, we perform a complete phenomenological study of the Dark Energy perturbation dynamics, investigating a *perturbed Early Dark Energy* set of models [55, 56] involving sound speed and anisotropic stress, in which the Dark Energy is dynamic and does not need to be negligibly small with respect to the other components in the early Universe, thus easing the fine-tuning problem [57, 58, 59]. Originally, this was the main motivation of introducing the EDE models, while recently some types of scalar field originated sets of this model has been recommended to resolve the Hubble tension as well [60]. We explore the impact of EDE on cosmological observables, more precisely, the effect of additional EDE component in the energy density of the early Universe and also perturbations on the CMB angular power spectrum as well as the linear matter power spectrum. We also single out the Integrated Sachs Wolfe (ISW) contribution of the temperature CMB power spectrum. Moreover, we study the effect of EDE components as well as the perturbations on CMB lensing potential. We mainly focus on a quantitative exploration of an extended cosmological parameter space including the mass of neutrinos and tensor-to-scalar ratio using a combination of recent data set available.

Second, we focus on the physics of the matter components of the Universe by performing a direct calculation of the gravitational coupling of baryons and CDM using N-body simulations and investigating the impact of the *relative baryon-CDM density and velocity perturbations*, δ_{bc}, θ_{bc}, on LSS tracers which are commonly neglected in the studies of structure formation. However, taking them into account might become very important in the era of high precision cosmology. The aforementioned perturbations are originated from the different evolution of baryon and CDM distributions due to the photon pressure before recombination. Since baryons were tightly

---

[2] S_8 is the principal-component parameter for weak gravitational lensing analyses (see for instance [52]), contains σ_8, which measures the amplitude of the linear matter power spectrum on the scale of $8\,\text{h}^{-1}\,\text{Mpc}$. and Ω_m, which measures the total matter density in the Universe. We will describe main cosmological parameters in chapter 3, Sec. 1.

[3] A_{lens} parameter is an "unphysical" parameter that rescales by hand the effects of gravitational lensing on the CMB angular power spectra and can be measured by smoothing of the peaks in the damping tail. For $A_{\text{lens}} = 0$, one has no lensing effect, while for $A_{\text{lens}} = 1$, one simply recovers the value expected in the cosmological model of choice. Interestingly, the Planck CMB power spectra show a preference for $A_{\text{lens}} > 1$ at more than two standard deviations.

coupled to photons prior to recombination, they begin to collapse on sub-horizon scales much later than the CDM. Gravitational evolution after recombination slowly erases this difference since baryons can then fall in CDM potential wells. This process is normally assumed to be over before redshift zero implying that the distributions of the two fluids become identical on large scales, with the power spectrum of the fluctuations of each fluid given by the total matter power spectrum. This assumption is however not exactly true, and recently substantial effort was put in correctly describing and simulating the evolution of the two fluids across cosmic history [11, 61, 62, 63, 64, 65, 10, 66, 67]. Performing gravity-only 2-fluid cosmological N-body simulations we assess the impact of such perturbations on the dark matter halos distribution. We generate the initial conditions for such simulations by introducing two different distinct transfer function for baryons and CDM. Since the two types of particles in our simulations, contains different mass, a high force resolution for the mass resolution would cause of appearing a spurious coupling between baryons and CDM particles. To alleviate it, we used the adaptive gravitational softening (AGS) [68] only for baryons. In particular, we focus on the baryon fraction in halos as a function of mass and large-scale δ_{bc}, which also allows us to study details of the nontrivial numerical setup required for such simulations. We further measure the cross-power spectrum between the halo field and δ_{bc} over a wide range of mass. This cross-correlation is nonzero and negative which shows that halo formation is affected by δ_{bc}. We measure the associated bias parameter $b_{\delta_{bc}}$ and compare it to recent results, finding good agreement. Finally, we quantify the impact of such perturbations on the halo-halo power spectrum and show that this effect can be degenerate with the one of massive neutrinos for surveys like "the Dark Energy Spectroscopic Instrument" (DESI)[4].

And third we consider the constraining potential of the most promising observables. We assess the impact of relative baryon-CDM perturbations on cosmic voids and on the real-space 2-point correlation functions of various fluid, in particular the position of the BAO peak. We do this using the aforementioned 2-fluid simulations, and compare the results with those obtained in a standard gravity-only 1-fluid simulation. We emphasis that we work in configuration space, in contrast with the second project where we worked in Fourier space.

Recently, cosmic voids became one of the most promising cosmological probes where to look at signatures of new physics. Because of the fact that they are almost empty regions, their evolution during cosmic history is at most weakly nonlinear and their properties could possibly be affected by the primordial density field from which they formed. This fact motivates us to investigate the effects of baryon-CDM relative perturbations on these regions and their statistics. In this project in particular, we investigate the void size function (VSF) using different tracers for finding voids (particles and halos), we measure the void-void and halo-void correlation functions, as well as the void density profile, and we compute the void bias in presence of baryon-CDM perturbations using 2-fluid simulations.

Measurements of the BAO feature in the distribution of galaxies are known as one of the most powerful tools for precision cosmology. Relative baryon-CDM perturbations could induce possible systematics for BAO measurements [69, 70, 71, 11, 72, 73], and could potentially bias the cosmological constraints as a systematic shift in $D_A(z)$, $H(z)$, and $f\sigma_8$ measurements. In this work we estimate this effect by comparing the real-space correlation function of various fluids in 2-fluid and 1-fluid simulations to explore the effect of such perturbations on the real-space matter and halo 2-point correlation functions (2PCF), and in particular we investigate the shift in the position of the BAO peak in the 2PCF of total matter, halos, CDM, baryon and the relative density δ_{bc}.

Outline

In Chapter 2 and Chapter 3, we begin with a review of the concordance model of cosmology, cosmological probes and numerical methods used in this work, where we highlight connections to techniques and questions addressed in later Chapters.

In Chapter 4, we brought our general picture of a perturbed EDE model which deviates from the cosmological constant, including a notable amount of energy in the early Universe and analyse the effects on CMB and matter spectra and mainly on the cosmological constraints.

In Chapter 5, we turn to dominant matter components in the early Universe, particularly the relative baryon-CDM perturbations. We present our numerical approach to add these perturbations in N-body simulations and study their impact on LSS observables in the Universe.

In Chapter 6, we focus on the impact of the aforementioned relative perturbations on the cosmic voids as a promising novel probe for cosmology as well as the shift in the position of the BAO peak in real-space 2-point correlation function as a most robust tool for precision cosmology.

---

[4]https://www.desi.lbl.gov

We summarize our main results and draw our conclusions including an outlook on future goals and projects in Chapter 7.

2

General Cosmology

As we anticipated, CMB and LSS observations led to the establishment of the concordance ΛCDM model which is based on the assumption of isotropy and homogeneity at any given time, the so-called *cosmological principle*. In this chapter we give a brief summary of the ΛCDM cosmology, focusing on the expansion, (Sec. 1 and Sec. 2), the main relevant equations (Sec. 3), the initial conditions (Sec. 4) as well as introducing the CMB (Sec. 5) and the the large-scale structures in the linear regime (Sec. 6). We refer to textbooks [74, 75, 76, 77, 5, 78, 79], as well as specific works as indicated.

1   The concordance model of cosmology

Expansion means that back in time the distance between us and a distant galaxy was smaller than it is today. Conveniently, we describe this effect by introducing the scale factor, a, whose present value by convention is set to 1. As a consequence, the comoving distance between two points that serve as a coordinate system, remains constant but the physical distance is proportional to that through the scale factor, so it grows as time evolves. That is why the physical wavelength of light emitted from a distant object is stretched out proportional to the scale factor. Conventionally this stretching factor is defined by the "redshift", z, as follows:

$$1 + z \equiv \frac{\lambda_{obs}}{\lambda_{emit}} = \frac{a_{obs}}{a_{emit}} = \frac{1}{a_{emit}} \,. \tag{2.1}$$

There are three possibilities concerning the geometry of the Universe, the Universe can be Euclidean (flat), open or closed. These different concepts can be understood by considering two freely traveling particles that start their journey parallel to each other. In Euclidean geometry, their trajectories remain parallel, in a closed Universe, these two initially parallel particles gradually converge, and in an open Universe, their path diverges. According to the GR, the total energy density in the Universe determines its geometry (see [80, 76] as comprehensive references). In a closed Universe the total energy density is higher than a critical value, $\rho_{cr} \approx 10^{-29}\,\mathrm{g cm}^{-3}$, while in an open one it is lower. In the Euclidean Universe the total energy density is precisely equal to the critical value. Observations up to now indicate that our Universe is Euclidean to within errors, and this is a precise indication for what concerns the early Universe, as we will see [2].
The scale factor, "a" evolves through cosmic time as follows: at early times, in which radiation dominates, $a \propto t^{1/2}$ while at later times, non-relativistic matter accounts for most of the energy density, it switches to $a \propto t^{2/3}$. It would be useful to define the Hubble rate to quantify the change in the scale factor: $H(t) \equiv (1/a)(da/dt)$, which measures how rapidly the scale factor changes. For example for a Euclidean matter-dominated Universe, in which $a \propto t^{2/3}$, then $H = (2/3)t^{-1}$. Overall, general relativity predicts that the scale factor is determined by the Friedmann equation which can be driven using the time-time $(0 - 0)$ component of Einstein equations, which reduces to

$$H^2(t) = \frac{8\pi G}{3} \left[\rho(t) + \frac{\rho_{cr} - \rho(t_0)}{a^2(t)} \right], \tag{2.2}$$

where G is the Newton's constant, $\rho(t)$ is the energy density in the Universe as a function of time, and $\rho(t_0)$ is its value today. ρ_{cr} is the critical density, mentioned above, which is constant, and given by $\rho_{cr} \equiv (3H_0^2/8\pi G)$. It is worthy to mention that Eq. 2.2 allows for the possibility that the Universe could be non-Euclidean. Since if it is Euclidean, the sum of the energy densities today would be equal to the critical one and the last term would vanish. Therefore if the Universe is not Euclidean, the curvature contribution scales as $1/a^2$ (since the second term in Eq. 2.2 is proportional to $1/a^2$). In order to make use of the Friedmann equation we need to know how the energy density of different species of the Universe evolves with time. In Fig. 2.1, we plotted the behavior of this evolution. Let us consider first the energy of the non-relativistic matter, which means the energy of the collection of some particles (equal to their own rest mass energy that would remain constant with time) times their "number density". Since number density is proportional inversely to the volume, it would be proportional to a^{-3} (as we see in cyan/solid in Fig. 2.1). The other components of the Universe are the photons, where traveled freely since the Universe was very young. Today their wavelength appears in the microwave regime and they constitute the CMB, a perfect black-body spectrum with a very well-measured temperature of $T_0 = 2.726 \pm 0.001\,\mathrm{K}$ today [81, 21, 22, 82] (we will briefly discuss it in Sec. 5). Considering the redshift relation Eq. 2.1, since $\lambda = c/\nu \propto a$, the frequency of the photon, ν, decays as $1/a$ with expansion. Since the black-body spectrum is a function of ν/T, so we can describe this effect equivalently by stating that the temperature of the radiation as a function of time is given by: $T(t) = T_0/a(t)$. The energy density of black-body radiation scales as $T^4 \propto a^{-4}$. (as we see in orange/dashed in Fig. 2.1).
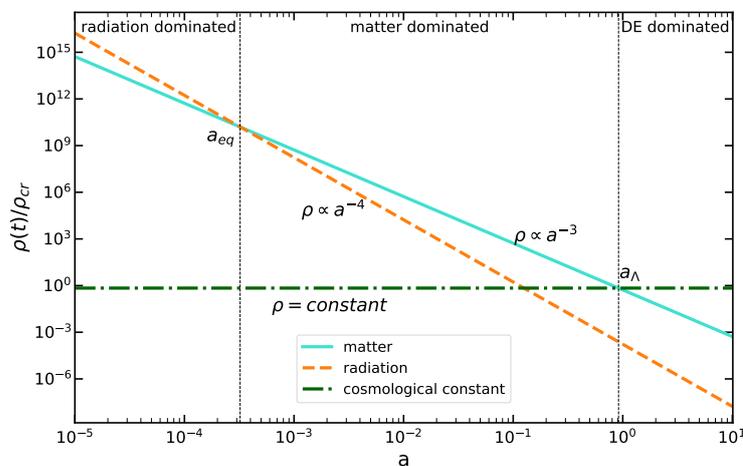
Figure 2.1: The behavior of the energy density as a function of scale factor for different components of the Universe in different epochs. For the Euclidean fiducial cosmology whose parameters are: $\Omega_m = 0.3106$ and $\Omega_\Lambda = 0.6894$ (we will describe cosmological parameters briefly in chapter 3, Sec. 1), For non-relativistic matter the energy density behaves as ($\propto a^{-3}$), and for radiation as $\propto (a^{-4})$. The cosmological constant as it appears from its name behaves as a constant energy density through all the epochs. All these components are plotted in the unit of critical density today. Radiation, matter, and the dark energy-dominated era are illustrated in the figure as well. The epoch in which the energy densities of matter and radiation are equal is known as a_{eq}, while the epoch in which the densities of matter and cosmological constant match is shown by a_Λ. This is the original motivation for introducing an amount of energy in the early Universe, the Early Dark Energy component, which we will describe in detail in chapter 4 to solve the fine-tuning problem which has been built on top of this question that why the amount of energy in the early Universe should have begun from an incredibly low initial amount in comparison with the other components of the Universe such as matter and radiation (compare the initial values of energy with matter and radiation at very small a).

Since the Universe is expanding, we should observe galaxies moving away from us. Edwin P. Hubble [17] in 1929 found that distant galaxies are in fact all apparently receding from us, i.e. they are "redshifted". He also suggested that velocity increases with distance, see Fig. 2.2. If we have two galaxies with physical distance $d = ax$, (x is the comoving distance), in the absence of any peculiar velocity, $\dot{x} \equiv dx/dt = 0$, the relative velocity v would be equal to:

$$v = \frac{d}{dt}(ax) = \dot{a}x = H_0 d \quad (v \ll c),$$
(2.3)
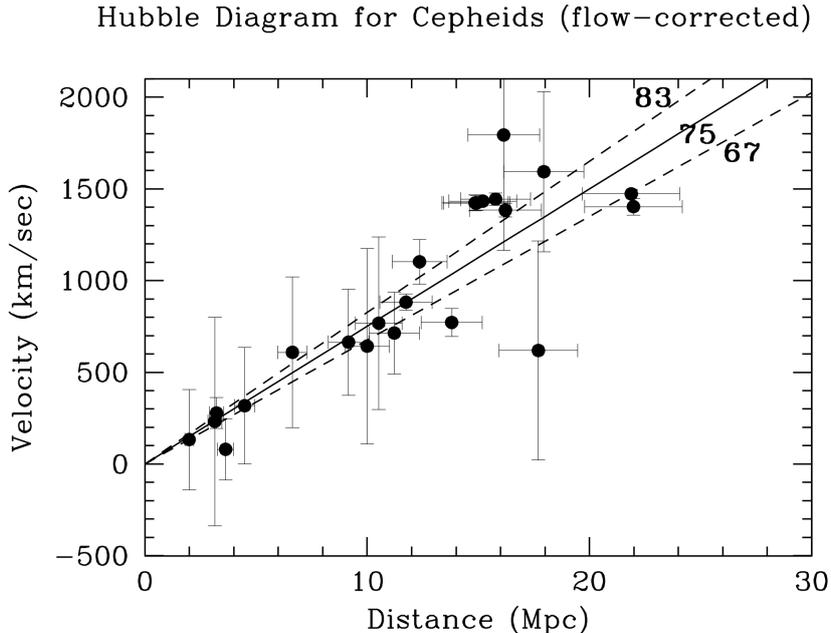
Figure 2.2: The Hubble diagram, (modern version) from Hubble Space Telescope Key project [3]. Each point corresponds to a galaxy whose distance has been estimated using pulsating stars known as "Cepheid" variables. The recession velocity for each galaxy is then corrected using a model for the peculiar velocity field in the neighborhood of the Milky Way. The lines show the prediction of the Hubble law with different values of H_0 (in kms^{-1}Mpc^{-1}), as indicated.

where the overdots indicated derivatives with respect to time t. By considering the fact that $z \simeq v/c$ (c is the speed of light, this relation shows that the apparent velocity should increase linearly with distance with a slope given by H_0, the "Hubble constant". Indeed Eq. 2.3 represents the "Hubble law" in which the value of the Hubble constant can be simply determined by measuring the slope of the line in the Hubble diagram shown in Fig. 2.2. We note that, indeed the distance-redshift relation depends on the energy content of the Universe. As we mentioned already, data indicate a current best-fit scenario that is Euclidean and contains about 70% of the energy in the form of a cosmological constant, or some other form of dark energy [2, 83].

The concordance model of cosmology can be summarized as follows: a Euclidean Universe that is dominated today by non-baryonic CDM and a cosmological constant, with initial perturbations generated at very early times though a process known as the Inflation, which we will describe later. Currently, all measurements are consistent with dark energy as a cosmological constant, Λ, that is why the concordance model of cosmology is usually called a (flat) "ΛCDM". In the following, we briefly discuss the terms that appeared in defining the concordance model.

Cold Dark Matter is not a part of the standard model of particle physics. The cold part of this term refers to the fact that these particles are indeed non-relativistic at the time in which they decouple by thermal equilibrium. The CMB anisotropy analysis from WMAP and Planck shows that around a fifth of the total matter is in the form that does not interact with ordinary matter or photons [84, 85]. Moreover, the theory of Big Bang nucleosynthesis (BBN), which accurately predicts the observed abundance of the chemical elements, indicates that the vast majority of dark matter in the Universe cannot be baryonic [86]. However, another evidence comes from the rotation curves of galaxies. Indeed, a mismatch between the matter inferred from gravity and that which we can see in the form of baryons exists on all galactic and extra-Galactic scales, and it always points toward roughly five times more dark matter than baryons.

Cosmological Constant's evidence started with observations of distant supernovae [28, 29], indicating an accelerated cosmological expansion, suggesting that beside the ordinary matter and radiation, there must be a DE component. There is a possibility that this new form of energy remains constant with time, i.e. being a CC, This possibility was first introduced (and later abandoned) by Einstein. Cosmologists have explored other forms though, many of which behave quite differently, implementing a model building for explaining a more general DE component. In chapter 4, we will see one dynamical form of this generalizations.

Inflation is the most likely mechanism for generating the initial perturbations that grew into the structure observed today. This theory considers that there is a very early epoch in the early Universe in which the scale factor grew exponentially rapidly with time. We will briefly explain this concept in Sec. 4 and we refer to

comprehensive works for reviews [38, 39, 40, 41, 42]. Indeed the epoch of inflation shares some features with DE, which is that the dominant form of energy remained roughly constant as the Universe expands.

2 The expanding Universe

The "general covariance" states that an observer in a uniform gravitational field makes exactly the same measurements as the one in an accelerated reference frame. In four space-time dimensions we have the following metric definition:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu \,. \tag{2.4}$$

In flat cosmology, the statial metric is Minkowski, except that spatial coordinates must be multiplied by the scale factor. Thus, the Euclidean Friedmann–Lemaître–Robertson–Walker (FLRW) metric is given by

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & a^2(t) & 0 & 0 \\ 0 & 0 & a^2(t) & 0 \\ 0 & 0 & 0 & a^2(t) \end{pmatrix} \,. \tag{2.5}$$

In a curved space, the concept of a straight line gets generalized to a geodesic, the shortest path between two points. GR states that this is precisely the path followed by a particle in the absence of any forces apart from gravity which can be illustrated by the following equation by introducing the parameter λ which monotonically increases along the particle's path.

$$\frac{d^2 x^\mu}{d\lambda^2} + \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} = 0 \,, \tag{2.6}$$

where the coefficient, known as Christoffel symbol, $\Gamma^\mu_{\alpha\beta}$, can be obtained directly from the metric:

$$\Gamma^\mu_{\alpha\beta} = \frac{g^{\mu\nu}}{2} \left[\frac{\partial g_{\alpha\nu}}{\partial x^\beta} + \frac{\partial g_{\beta\nu}}{\partial x^\alpha} - \frac{\partial g_{\alpha\beta}}{\partial x^\nu} \right] \,. \tag{2.7}$$

2.1 Cosmic distances

As we mentioned before in Sec. 1, by expanding the Universe, the comoving distance between two objects remains constant. Let us Consider a comoving distance between a distant light source and us. In a small time interval dt, light travels a comoving distance $dx = dt/a$, so the total comoving distance traveled by light that began its journey from an object at time t when the scale factor would be equal to a (or redshift $z = 1/a - 1$) is

$$\chi = \int_t^{t_0} \frac{dt'}{a(t')} = \int_{a(t)}^1 \frac{da'}{a'^2 H(a')} = \int_0^z \frac{dz'}{H(z')} \,. \tag{2.8}$$

In order to make the final expression in the equation above more clear, we can think that for small redshifts z we can write the comoving distance as $\chi \approx z/H_0$ (see the discussion of the Hubble diagram at small redshifts in Sec. 1). Now let us define the comoving distance η, that light could have traveled since $t = 0$,

$$\eta(t) \equiv \int_0^t \frac{dt'}{a(t')} \,. \tag{2.9}$$

No information can be propagated farther than η on the coordinate grid. Therefore, the regions which are separated by distances higher than η are not causally connected. Then we can think of η as the "comoving horizon". We can also think of η, which is monotonically increasing, as a time variable and call it the conformal time which is usually done.

A classical way to determine distances in astronomy is to measure the angle θ specified by an object of known physical size l ("standard ruler"). Since this angle is always very small in astronomy, the distance to
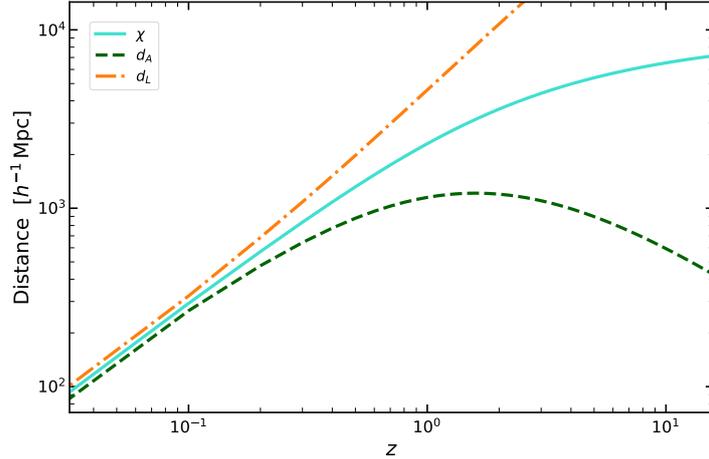
Figure 2.3: Three different so-called distance measurements in the Euclidean expanding Universe in the concordance cosmological model. The comoving distance, χ, (in orange/dot-dashed) Eq. 2.8, the angular diameter distance, d_A, (in green/dashed) Eq. 2.10, and the luminosity distance, d_L, (in cyan/solid) Eq. 2.12.

that object would be equal to $d_A = \dfrac{l}{\theta}$, which is known as the "angular diameter distance". In order to compute the angular diameter distance in an expanding Universe, by considering the fact that the comoving size of the object is l/a, using the Eq. 2.8, the angle specified in a Euclidean Universe is $\theta = (l/a)/\chi(a)$. Thus, the angular diameter distance reads

$$d_A^{Euc} = a\,\chi = \frac{\chi}{1+z}\,. \tag{2.10}$$

As can be seen also from Fig. 2.3, the angular diameter distance is equal to the comoving distance at low redshifts, but actually decreases at very large ones. Objects at large redshifts appear larger than they would at intermediate redshifts. This is a consequence of the fact that the entire Universe was smaller and hence the emitting galaxy and we as observers, were physically much closer. Another classical way of calculating distances in astronomy is to measure the flux from an object of known luminosity ("standard candle") with the discovery of gravitational-wave sources, we now also have "standard sirens", to which all of the following applies as well but here we are not going to discuss them). The flux, F, observed at a distance, d, from a nearby source of known luminosity L would be

$$F = \frac{L}{4\pi d^2}\,. \tag{2.11}$$

In an expanding Universe, knowing the fact that the source centered at the origin, the flux we observe is $F = \dfrac{L(\chi)}{4\pi\chi^2(a)}$, in which $L(\chi)$ represents the luminosity through a comoving spherical shell with radius $\chi(a)$. If we assume that photons are all emitted with the same energy, then $L(\chi)$ is this energy multiplied by the number of photons passing through a comoving spherical shell per unit time. Considering the fact that for a fixed time interval, the physical distance at early times is smaller (by a factor of a) than at late times, the number of photons crossing a shell in a fixed time interval will be smaller today than at emission. If we consider this difference, the energy per unit time passing through a comoving shell at a distance $\chi(a)$ from the source will be a factor of a^2 smaller than the luminosity at the source. Therefore the flux we observe will be $F = L\,a^2/4\pi\chi^2(a)$ in which L represents the luminosity at the source. Using Eq. 2.11 we can define the luminosity distance in a Euclidean expanding Universe as follows

$$d_L^{Euc} \equiv \frac{\chi}{a}\,. \tag{2.12}$$

2.2 Evolution of energy

Here we characterize the background properties of matter, radiation and other components. The energy density and pressure can be combined in to a relativistic tensor, called "energy-momentum tensor".

$$
T^\mu_\nu = \begin{pmatrix} -\rho & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix}, \tag{2.13}
$$

where P is the pressure and ρ is the energy density that can evolve according to the continuity and Euler equations, which can be written in a more general way as the local energy and momentum conservation.

$$
\nabla_\mu T^\mu_\nu \equiv \frac{\partial T^\mu_\nu}{\partial x^\mu} + \Gamma^\mu_{\alpha\mu} T^\alpha_\nu - \Gamma^\alpha_{\nu\mu} T^\mu_\alpha = 0 \,. \tag{2.14}
$$

Thus, by assuming isotropy the conservation law in an expanding Universe reads

$$
\frac{\partial \rho}{\partial t} + 3\frac{\dot{a}}{a}(\rho + P) = 0 \,. \tag{2.15}
$$

and by rearranging terms in the equation above we have: $a^{-3}\partial[\rho a^3]/\partial t = -3(\dot{a}/a)P$. Since non-relativistic matter has effectively zero pressure, $\partial[\rho_m a^3]/\partial t = 0$, implying that the energy density of matter follows $\rho_m \propto a^{-3}$. Also radiation has $P_r = \rho_r/3$, so we have $\partial\rho_r/\partial t + (\dot{a}/a)4\rho_r = a^{-4}\partial[\rho_r a^4]/\partial t = 0$, which means that $\rho_r \propto a^{-4}$ considering the decrease in energy per particle dy expanding the Universe. We already illustrated these behaviors in Fig. 2.1. It is convenient to summarize the cases of matter and radiation in one equation, and generalize the evolution results to other constituents, by defining the "equation of state parameter", $w_s \equiv P_s/\rho_s$, where s stands for any constituent. Matter corresponds to $w = 0$, radiation to $w = 1/3$, and a cosmological constant has $w = -1$. However, in general the equation of state does not have to be time-independent. Eq. 2.15 can be integrated to find the evolution of any components s with a time-dependent equation of state $w_s(a)$:

$$
\rho_s(a) \propto exp\left\{ -3 \int^a \frac{da'}{a'}[1 + w_s(a')] \right\} \overset{w_s=const}{\propto} a^{-3(1+w_s)} \,. \tag{2.16}
$$

It is common in cosmology to have all energy densities in the same units. The usual way to do this is to divide all energy densities by the critical density today, mentioned in Sec. 1, and define the density parameters as follows

$$
\Omega_s \equiv \frac{\rho_s(t_0)}{\rho_{cr}} \,, \tag{2.17}
$$

where again "s" in the equation above, indicates a specific component. It can be CDM (c), baryons (b), photons (γ), neutrinos (ν), CC (Λ) or DE. It is also common to use r for all radiation constituents (photons and ultra-relativistic neutrinos), and m for the total non-relativistic matter: $\Omega_m = \Omega_b + \Omega_c$. We can write the density of a given s as a function of scale factor as follows assuming that its equation of state, ω_s is time-independent:

$$
\rho_s(a) = \Omega_s \, \rho_{cr} \, a^{-3(1+\omega_s)} \,. \tag{2.18}
$$

The critical density is $\rho_{cr} = 3H_0^2/8\pi G$ and that H_0 is not perfectly known. This means that any precise constraint on the physical mean density, for instance the density of baryons, indeed constrains the parameter combination $\Omega_b h^2$. That's why in the literature constraints are often phrased in terms of this combinations, usually with this symbol: $\omega_s \equiv \Omega_s h^2$.

3 The fundamental equations of cosmology

There are essentially two classes of equations which are relevant in Cosmology, and take their name from the scientists who derived those, namely Einstein and Boltzmann. Einstein equations determine the behavior of gravitational interactions, in the background as well as perturbations. Boltzmann equations deal with all aspects of statistical mechanics, for matter and radiation. In the following, we give the general aspects of them, focusing on the ones which are relevant in the following.

3.1 Einstein equations

GR equations relating metric and stress energy tensors are written as

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} \,, \tag{2.19}$$

where $G_{\mu\nu}$ is the Einstein tensor: $G_{\mu\nu} \equiv R_{\mu\nu} - (1/2)g_{\mu\nu}R$, and $R_{\mu\nu}$ and R are the Ricci tensor and its contraction ($R \equiv g^{\mu\nu}R_{\mu\nu}$), depending only on the metric and its derivatives. Λ is the CC, G is the Newton's constant, and $T_{\mu\nu}$ is the energy-momentum tensor, whose expression in the background Universe has been already described in Sec. 2.2. Thus, the left-hand side of Eq. 2.19 is a function of the metric, the right, a function of the constituents of the Universe: the Einstein equations relate the two. The Ricci tensor can be demonstrated in terms of the Christoffel symbols:

$$R_{\mu\nu} = \Gamma^\alpha_{\mu\nu,\alpha} - \Gamma^\alpha_{\mu\alpha,\nu} + \Gamma^\alpha_{\beta\alpha}\Gamma^\beta_{\mu\nu} - \Gamma^\alpha_{\beta\nu}\Gamma^\beta_{\mu\alpha} \,. \tag{2.20}$$

In this expression, commas denote derivatives with respect to x. For instance: $\Gamma^\alpha_{\mu\nu,\alpha} \equiv \partial\Gamma^\alpha_{\mu\nu}/\partial x^\alpha$. The Christoffel symbols are proportional to the first derivative of the metric with respect to the coordinates.

3.2 Boltzmann equation

In cosmology, we are not interested in the fate of individual particles, but in their statistical behavior. Imagine we have a collection of particles in some region of the space. In classical physics, we can describe these particles by their positions and momenta $\{\mathbf{x}_i, \mathbf{p}_i\}$. Then this is common to define a distribution function by relating it to the number of particles in a small phase-space element around $(\mathbf{x}, \mathbf{p})$ (see Eq. 2.21), By writing

$$N(\mathbf{x}, \mathbf{p}, t) = f(\mathbf{x}, \mathbf{p}, t)(\Delta x)^3 \frac{(\Delta p)^3}{(2\pi)^3} \,, \,. \tag{2.21}$$

$N(\mathbf{x}, \mathbf{p}, t)$ represents the number of particles, while $f(\mathbf{x}, \mathbf{p}, t)$ is the distribution function. Δx and Δp are the space and momenta elements, respectively.

The total number of particles is conserved, which is provided by the vanishing of the total time derivative of f vanishes:

$$\frac{df(\mathbf{x}, \mathbf{p}, t)}{dt} = 0 \quad \text{where} \quad \frac{d}{dt} = \frac{\partial}{\partial t} + \dot{\mathbf{x}}.\nabla_x + \dot{\mathbf{p}}.\nabla_p \,. \tag{2.22}$$

If we want to take into account the particle-particle interactions as well, the equation must be modified in order to include a source term in the right hand side. This is known as "collision term", which describes how particles are moved from one phase-space element to another.

$$\frac{df}{dt} = C[f] \,. \tag{2.23}$$

3.2.1 Boltzmann equation in an expanding Universe

In order to generalize the Boltzmann equations in an expanding cosmology, we have to take into account the fact that the equation of motion of particles ($\dot{\mathbf{x}} = p/m, \quad \dot{\mathbf{p}} = m\mathbf{a}(\mathbf{x}, \mathbf{p}, t)$) is modified along the geodesics:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x^i}.\frac{dx^i}{dt} + \frac{\partial f}{\partial p}.\frac{dp}{dt} + \frac{\partial f}{\partial \hat{p}^i}.\overset{0}{\cancel{\frac{d\hat{p}^i}{dt}}} \,. \tag{2.24}$$

In the equation above, since the direction of the particle momentum does not change in comoving coordinates, we can drop the last term. Then we have to find the terms: dx^i/dt and dp/dt. Here, the three-momentum, $\mathbf{p}$, becomes a four-vector and can be written usually in the following form: $P^\mu \equiv dx^\mu/d\lambda$, where λ parametrizes the particle's path. Note that for the FLRW metric we can write it as: $E \equiv (P^0)^2 = p^2 + m^2$. Using the four-momentum for (0,0) and (i,i) components, and defining $P^i \equiv dx^i/d\lambda$, $P^0 \equiv dt/d\lambda$, we can write the following relation:

$$\frac{dx^i}{dt} = \frac{dx^i}{d\lambda}\frac{d\lambda}{dt} = \frac{P^i}{P^0} = \frac{p}{E}\frac{\hat{p}^i}{a} \,. \tag{2.25}$$

We can now use the time component of geodesic equation for momentum

$$\frac{dP^0}{d\lambda} = -\Gamma^0_{\alpha\beta}P^\alpha P^\beta \,. \tag{2.26}$$

Using the definition of Christoffel symbols for FLRW metric, leading to

$$p\frac{dp}{dt} = -Hp^2 \quad \rightarrow \quad \frac{dp}{dt} = -Hp\,. \tag{2.27}$$

This equation says that the physical momentum of any particle decays as $1/a$ in an unperturbed expanding Universe. Our final expression for the Boltzmann equation in the homogeneous expanding Universe reads

$$\frac{\partial f}{\partial t} = \frac{p}{E}\frac{\hat{p}^i}{a}\frac{\partial f}{\partial x^i} - Hp\frac{\partial f}{\partial p} = C[f]\,. \tag{2.28}$$

The perturbed cosmology is more complicated (and phenomenologically richer) with respect to the homogeneous Universe. In the following, we outline the main steps in deriving the system of coupled equations, the so-called Einstein-Boltzmann equations in a perturbed spacetime.

3.2.2 Relativistic Perturbation Theory

In the early Universe and on large scales, we can treat the inhomogeneities as perturbations of the smooth FLRW background, $\bar{g}_{\mu\nu}$, and write

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}, \tag{2.29}$$

where $|h_{\mu\nu}| \ll g_{\mu\nu}$, is a symmetric 4-tensor. Due to the gauge-freedom and under the infinitesimal coordinate transformation $x^\mu \rightarrow \tilde{x}^\mu = x^\mu + \xi^\mu(x)$, only 6 parameter from the 10 independent parameters of $h_{\mu\nu}$ are physical [78]. We have to choose the gauge by fixing the 4 degrees of freedom. $h_{\mu\nu}$ can be decomposed in the well-known *scalar-, vector-, and tensor-perturbations* as follows:

$$h_{00} = -E, \tag{2.30}$$

$$h_{i0} = a\left[\frac{\partial F}{\partial x^i} + G_i\right], \tag{2.31}$$

$$h_{ij} = a^2\left[A\delta_{ij} + \frac{\partial B}{\partial x^i \partial x^j} + \frac{\partial C_i}{\partial x^j} + \frac{\partial C_j}{\partial x^i} + D_{ij}\right], \tag{2.32}$$

where A, B, E and F are scalars, C_i and G_i are divergence-less vector fields and D_{ij} is a traceless, symmetric and divergence-less tensor field. These different modes correspond to different physical phenomena: the gravitational potential (scalar mode), gravito-magnetism (vector mode) and gravitational radiation (tensor mode) (see e.g. [78]) In the following, we will be considering scalar modes only. Due to gauge-freedom, we can choose a coordinate system in which $B = F = 0$. Choosing $A = 2\Phi$ and $E = 2\Psi$, we recover the common notation of the Newtonian gauge in which the full metric becomes

$$g_{\mu\nu} = \begin{pmatrix} -1 - 2\Psi & \\ & a^2\delta_{ij}(1 + 2\Phi) \end{pmatrix}. \tag{2.33}$$

The Newtonian gauge has the advantage that it can easily be related to the Newtonian limit of gravity.

Analogously, we also perturb the stress-energy tensor $T_{\mu\nu} = \bar{T}_{\mu\nu} + \delta T_{\mu\nu}$. For photons we parametrize the temperature inhomogeneities as $T = \bar{T}(t)(1 + \Theta(\mathbf{x}, \hat{\mathbf{p}}, t))$ and label n-th multipole of Θ as Θ_n. Analogously for neutrinos, where we name the perturbations $\mathcal{N}$. For CDM, we write density perturbations as $\rho = \bar{\rho}(t)(1 + \delta(\mathbf{x}, t))$ and peculiar velocities as $v(\mathbf{x}, t)$ and ignore all higher order moments due to the coldness of CDM. Similarly to CDM, the baryonic inhomogeneities can be parametrized by $\delta_b(\mathbf{x}, t)$ and $v_b(\mathbf{x}, t)$.

To study the evolution of the perturbed metric and energy density, we can split the problem into two parts: studying the effect of the inhomogeneous potential on the content of the Universe via the Boltzmann equation and the effect of the perturbed components on the metric via the Einstein field equations. Computing these reactions give rise to a system of coupled differential equations known as the Einstein-Boltzmann equations. Assuming the perturbations are small, the system can be simplified by only considering terms linear in perturbations and transforming the equations to Fourier space.

Using the conformal time η (represented in Eq. 2.9) and defining $\mu = \hat{\mathbf{p}} \cdot \hat{\mathbf{k}}$ as the cosine between the wave-vector, $\mathbf{k}$, and the photon momentum, $\mathbf{p}$, the linear Boltzmann equations can be written as [75, 5]

$$\dot{\Theta} + ik\mu\Theta = -\dot{\Phi} - ik\mu\Psi - \dot{\tau}\left[\Theta_0 - \Theta + \mu v_b - \frac{1}{2}\mathcal{P}_2(\mu)\Pi\right], \tag{2.34}$$

$$\Pi = \Theta_2 + \Theta_{P2} + \Theta_{P0}, \tag{2.35}$$

$$\Theta_p + ik\mu\Theta_p = -\dot{\tau}\left[-\Theta_p + \frac{1}{2}(1 - \mathcal{P}_2(\mu))\Pi\right], \tag{2.36}$$

$$\dot{\delta} + ikv = -3\dot{\Phi}, \tag{2.37}$$

$$\dot{v} + \frac{\dot{a}}{a}v = -ik\Psi, \tag{2.38}$$

$$\dot{\delta}_b + ikv_b = -3\dot{\Phi}, \tag{2.39}$$

$$\dot{v}_b + \frac{\dot{a}}{a}v_b = -ik\Psi + \frac{\dot{\tau}}{R}[v_b + 3i\Theta_1], \tag{2.40}$$

$$\dot{\mathcal{N}} + ik\mu\mathcal{N} = -\dot{\Phi} - ik\mu\Psi, \tag{2.41}$$

where $\mathcal{P}_l$ is the Legendre polynomial of order l, τ is the optical depth, Θ_p the perturbations in the photon polarization field and $R \equiv 3\rho_{b,0}/4\rho_{\gamma,0}$ is the baryon/photon ratio. Eq.(2.34), determines the evolution of the photon temperature including Compton scattering between photons and baryons, Eq.(2.36), describes the generation of photon polarization that can be observed in the CMB, Eq.(2.37) and Eq.(2.38) govern the evolution of CDM overdensities and peculiar velocities, and Eqs.(2.39) and (2.40) the ones of baryons which unlike the CDM are coupled to the photons by Compton scattering. Eq.(2.41) describes the evolution of the neutrino temperature inhomogeneities, assuming massless neutrinos.

The second part, the effects of the perturbations in the stress energy tensor on the potentials Φ and Ψ, are obtained from Einstein field equations. The linear approximations are as follows [75, 5]

$$k^2\Phi + 3\frac{\dot{a}}{a}\left(\dot{\Phi} - \frac{\dot{a}}{a}\Psi\right) = 4\pi G a^2[\rho_m\delta_m + 4\rho_r\Theta_{r,0}], \tag{2.42}$$

$$k^2(\Phi + \Psi) = -32\pi G a^2\rho_r\Theta_{r,2}, \tag{2.43}$$

in which subscripts m and r indicates all non-relativistic matter and relativistic radiation respectively. Note that if the radiation quadrupole $\Theta_{r,2}$, is small, then $\Phi \simeq \Psi$ is a good approximation.

4 Initial conditions

As we anticipated, the process of generation of initial conditions for cosmological perturbations can be activated by an initial phase of quasi-exponential expansion known as Inflation, which we briefly review here. Inflation ([38, 39, 40, 41, 42]) has been introduced in order to explain how regions that could not have been in causal contact with each other (Fig. 2.4) have the same temperature, which is another way to say, why the Universe is so homogeneous on large scales. This problem is known as "horizon problem" in cosmology.

By considering the comoving distance as an integral over the scale factor, rewriting Eq. 2.9 changing integration variable from t' to $\ln a'$, we have

$$\eta(a) = \int_0^a d\ln a' \frac{1}{a'H(a')}, \tag{2.44}$$

Which is saying that the comoving horizon is the logarithmic integral of the comoving Hubble radius, $1/aH$. The "comoving Hubble radius" approximately the distance over which light could travel in the course of one expansion time, which is indeed the time that the scale factor increases by a factor of e. This concept provides a benchmark to estimate whether particles can, at the given epoch, communicate within one $e-$fold of expansion. Therefore if there was an epoch of early acceleration during which the comoving Hubble radius decreased, then η_* may have received large contributions from very early times when the Hubble radius was much larger. This
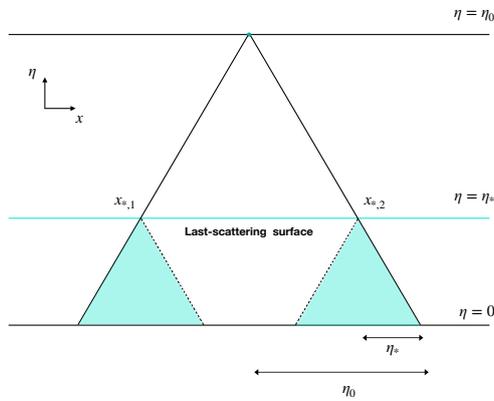
Figure 2.4: The diagram of the horizon problem, represented in η VS. x axis. The observer in the top center of the diagram can detect light signals only from his/her past light cone. When this cone intersects the last-scattering region (where radiation has decoupled from the rest of particles, and the CMB originates), $\eta = \eta_*$ (the cyan line here), perturbations at this time becomes accessible to observations. The CMB photons coming from points $x_{*,1}$ and $x_{*,2}$ from the last-scattering surface can not influence each other, receiving signals only signals from shaded regions beneath them. These two shaded regions do not overlap, and therefore no form of causal process could have allowed them to adjust to the same temperature if they started from different initial conditions. In other words, the comoving horizon at the last-scattering surface, η_*, is much smaller than the comoving horizon today, η_0.

15

epoch corresponds to the "inflation". The simplest possibility to generate such a transitory epoch of accelerated expansion in the literature is through the potential energy of a scalar field ϕ, with mass m where the potential would be $V(\phi) = m^2\phi^2/2$ (by the way, this is also what rules a class of models of dark energy named by quintessence). Using the stress-energy tensor, the time-time component would give us the the energy density as $\rho = \dot{\phi}^2/2 + V(\phi)$ and the pressure for a homogeneous field would be $P = \dot{\phi}^2/2 - V(\phi)$. Using the conservation equation, Eq. 2.15, we have: $\ddot{\phi} + 3H\dot{\phi} + dV/d\phi = 0$. By switching to the conformal time, η as the time variable rather than t, we have

$$\phi'' + 2aH\phi' + a^2 dV/d\phi = 0\,. \tag{2.45}$$

Once the field reaches the minimum of the potential, the inflationary epoch would end. The quasi-exponential expansion may be obtained through a quasi-flat potential, leading to an effective slow roll of the field towards the minimum. Here we introduce the following slowly-rolling variables. First, the fractional change in the Hubble rate during one $e-$fold of expansion, $\epsilon_{sr} \equiv \dfrac{d}{dt}\left(\dfrac{1}{H}\right) = -\dfrac{H'}{aH^2}$, which quantifies the departure of the spacetime from an exact de Sitter space, which is an empty Universe with a positive cosmological constant. Second, $\delta_{sr} \equiv \dfrac{1}{H}\dfrac{\ddot{\phi}}{\dot{\phi}} = -\dfrac{1}{aH\phi'}[aH\phi' - \phi'']$.

One of the greatest successes of Inflation is the capability of predicting that quantum-mechanical perturbations in the very early Universe are first produced when the relevant scales were causally connected. Then these scales are moved suddenly and quickly outside the horizon by inflation through the quasi-exponential growth of the scale factor, only to re-enter much later to serve as initial conditions for the growth of structure in the Universe. The perturbations can be describe in terms of their Fourier modes. The mean of a given Fourier mode, for example for the gravitational potential, is zero: $\langle\Phi(k)\rangle = 0$. Notice that any given Fourier mode is uncorrelated with a different one. However, a given mode has nonzero variance, so

$$\langle\Phi(\mathbf{k})\Phi^*(\mathbf{k}')\rangle = P_\Phi(k)(2\pi)^3\delta_D^{(3)}(\mathbf{k}-\mathbf{k}'), \tag{2.46}$$

where $\delta_D^{(3)}(\mathbf{k}-\mathbf{k}')$ represents the Dirac delta function which enforces the independency of the different modes. An important prediction of single-field inflation is that it generates adiabatic perturbations: different patches of the Universe have different overdensities, but the fractional density perturbations are the same for all species:

$$\frac{\delta\rho_s}{\dot{\rho}_s} = \frac{\delta\rho}{\dot{\rho}}\,, \tag{2.47}$$

with the same relations for the velocity perturbations. CMB has been confirmed the adiabatic nature of perturbations with a great precision. There exist another types of perturbations which allows for different primordial density perturbations in the different species (referred to as isocurvature perturbations which we will briefly note in the chapter 5 and chapter 6.) at most as a percent-level fraction of the adiabatic perturbations. The scalar perturbations generated during inflation are most commonly parametrized in terms of the power spectrum of the gauge-invariant curvature perturbation $\mathcal{R}$ which defines as follows:

$$P_\mathcal{R}(k) = \frac{2\pi}{k^3}\frac{H^2}{m_{pl}^2\epsilon_{sr}}\bigg|_{aH=k} \equiv 2\pi^2\mathcal{A}_s k^{-3}\left(\frac{k}{k_p}\right)^{n_s-1}, \tag{2.48}$$

where $\mathcal{A}_s$ is the variance of curvature perturbations in a logarithmic wavenumber interval centered around the pivot scale k_p and known as the amplitude of the primordial perturbations, and n_s is the scalar spectral index. The pivot scale is a matter of convention, and is often determined as the scale best constrained by a given set of observations (CMB anisotropies; the Planck team adopts $k_p = 0.05\,\mathrm{Mpc}^{-1}$). In our fiducial cosmology,

$$\mathcal{A}_s = \frac{k_p^3}{2\pi^2}P_\mathcal{R}(k_p) \simeq 2.1 \times 10^{-9}. \tag{2.49}$$

Therefore the usual amplitude of curvature perturbations on the scale k_p is $\sqrt{\mathcal{A}_s} \simeq 4.6 \times 10^{-5}$, which has the similar order of magnitude as the temperature fluctuations in the CMB.

The primordial tensor modes are conventionally parametrized via their total power spectrum $P_T(k)$ defined by

$$P_T(k) = \frac{32\pi}{k^3} \frac{H^2}{m_{pl}^2}\bigg|_{aH=k} \equiv 2\pi^2 \mathcal{A}_T k^{-3} \left(\frac{k}{k_p}\right)^{n_T},\tag{2.50}$$

which gives us the opportunity to define the conventional tensor amplitude, $\mathcal{A}_T$, and the tensor spectral index, n_T. We have to note that this convention which is common says that a scale-free scalar spectrum corresponds to $n_s = 1$, whereas for the tensor mode, corresponds to $n_T = 0$. In practice $\mathcal{A}_T$ is often replaced with the tensor-to-scalar ratio, r, which we will come back to it and open it as a cosmological parameter in our analysis in chapter 4, to see the effect of the dynamical early dark energy on it.

$$r(k) \equiv \frac{P_T(k)}{P_\mathcal{R}(k)} \overset{k=k_p}{=} \frac{\mathcal{A}_T}{\mathcal{A}_s}.\tag{2.51}$$

In the following section we introduce the cosmic microwave background.

5   The cosmic microwave background

The density perturbations in the primordial Universe could be investigated in the matter and radiation distribution. Understanding the evolution of the photon perturbations allows us to make predictions for the power spectrum of CMB anisotropies. Einstein-Boltzmann equations control the evolution of these photon perturbations. Photon perturbations evolve totally in a different manner before and after the epoch of recombination (roughly at redshift $z \simeq 1100$). Before the recombination epoch, $z \gtrsim 1100$, photons were tightly coupled to the baryons. They could thereby be described as a single "baryon-photon" fluid. After the recombination epoch, $z \lesssim 1100$, photons can free-stream from the "surface of last-scattering" to us. Note that after decoupling of the photons from baryons, the photon perturbations do not grow, because the gravitational potentials in the Universe are too weak to trap photons. So, immediately after decoupling, photons start to travel freely and they keep the perturbations they had at the level of decoupling with themselves. But the story is completely different for baryons and CDM, these components grow by orders of magnitude between decoupling and today.

The path of a single photon as scatters off a sea of electrons can be illustrated by a random walk form. Between each scattering event, photon travels a mean comoving distance λ_{MFP} which can be written by $\lambda_{\mathrm{MFP}} = (n_e \sigma_T a)^{-1} = -1/\tau$, where τ is the optical depth, n_e is the density of electrons and σ_T is the Thomson cross section ($\sigma_T = 0.665 \times 10^{-24}\,\mathrm{cm}^2$). During a Hubble time, H^{-1}, a photon scatters of order $n_e \sigma_T H^{-1}$ performing a random walk. We know that the total distance traveled in the course of a random walk is the mean free path times the square root of the total number of steps. Therefore, a cosmological photon moves a mean comoving distance in a Hubble time:

$$\lambda_D \sim \lambda_{\mathrm{MFP}} \sqrt{n_e \sigma_T H^{-1}} = \frac{1}{a\sqrt{n_e \sigma_T H}}.\tag{2.52}$$

Any perturbation on scales smaller than λ_D can be expected to be washed out (). In Fourier space, this effect corresponds to the damping of all high-k modes. There is indeed one caveat to this free-streaming picture. We are assuming that nothing happens in the journey of the photons from the last scattering surface to us. But indeed this is not a completely correct picture. While gravitational potentials are constant deep in the matter domination era, they evolve right after the recombination (because of the presence of radiation) and at late times (due to dark energy). Evolving potentials produced additional perturbations to the photons via the ISW effect. We will describe this in Chapter 4.

5.1   The angular power spectrum

The photon distribution depends not only on $\mathbf{x}$ and time but also on the direction of propagation of the photon, $\hat{\mathbf{p}}$. In Fourier space, therefore, the photon perturbations depend on k and η and also on $\mu = \hat{\mathbf{p}} \cdot \hat{\mathbf{k}}$. In general, the ℓ-th moment of the temperature field is defined as:

$$\Theta_\ell(k,\eta) \equiv \frac{1}{(-i)^\ell} \int_{-1}^{1} \frac{d\mu}{2} \mathcal{P}_\ell(\mu) \Theta(\mu, k, \eta),\tag{2.53}$$

where $\Theta(k, \mu, \eta)$ is the photon perturbation variable and $\mathcal{P}_\ell$ stands for the Legendre polynomial of order ℓ. The higher moments contain information about the small-scale anisotropies of the radiation field. Now we want to see how is the relation between these perturbations and the observed anisotropy pattern. We have to first describe how the temperature of the CMB radiation field is characterized today as:

$$T(\mathbf{x}, \hat{\mathbf{p}}, \eta) = T(\eta)[1 + \Theta(\mathbf{x}, \hat{\mathbf{p}}, \eta)]. \tag{2.54}$$

Even though this field is defined at every point in space and time, we can only observe it at $\mathbf{x}_0$ and t_0. So the only remained parameter is the direction, indeed the only wealth of information we get from observing comes from the changes in the temperature as the direction vector $\hat{\mathbf{p}}$ changes. Commonly the locations on the sky are $\hat{p}_x$, $\hat{p}_y$ and $\hat{p}_z$ components of $\hat{\mathbf{p}}$ which usually labeled by their polar coordinates, θ and ϕ.
By expanding the temperature perturbations in terms of spherical harmonics we have:

$$\Theta(\mathbf{x}, \hat{\mathbf{p}}, \eta) = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}(\mathbf{x}, \eta) Y_{\ell m}(\hat{\mathbf{p}}), \tag{2.55}$$

where ℓ and m subscripts stand for the conjugates to real-space unit vector $\hat{\mathbf{p}}$ and $Y_{\ell m}(\hat{\mathbf{p}})$ are the complete set of eigenfunctions for expansion on the surface of a sphere. All of the information contained in the temperature field T, is also contained in the $(\mathbf{x}, \eta)$-dependant amplitudes $a_{\ell m}$.

By multiplying both sides of the Eq. 2.55 by $Y_{\ell m}^*(\hat{\mathbf{p}})$ and using the orthogonality property of the spherical harmonics ($\int d\Omega Y_{\ell m}(\hat{\mathbf{p}}) Y_{\ell' m'}^*(\hat{\mathbf{p}}) = \delta_{\ell\ell'}\delta_{mm'}$) we have:

$$a_{\ell m}(\mathbf{x}, \eta) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{x}\cdot\mathbf{k}} \int d\Omega \, Y_{\ell m}^*(\hat{\mathbf{p}}) \, \Theta(\mathbf{k}, \hat{\mathbf{p}}, \eta), \tag{2.56}$$

in which we used $\Theta(\mathbf{k})$ in terms of the Fourier transform of $\Theta(\mathbf{x})$. Having only the density distribution, we can not make predictions for any particular $a_{\ell m}$. Their distribution is Gaussian (with a zero mean value and a non-zero variance which is called $\mathcal{C}(\ell)$) that traces its origin to the quantum fluctuations laid down during inflation. We can introduce them as follows:

$$\langle a_{\ell m} \rangle = 0; \qquad \langle a_{\ell m} a_{\ell' m'}^* \rangle = \delta_{\ell\ell'}\delta_{mm'}\mathcal{C}(\ell), \tag{2.57}$$

where $\langle \rangle$ shows an ensemble average. There is a underlying uncertainty in the knowledge we could get about the $\mathcal{C}(\ell)$. This uncertainty, which is more pronounced at low ℓ, is called cosmic variance. Quantitatively the uncertainty scales as the inverse of the square root of the number of samples. More precisely, it is the uncertainty on the estimation of $\mathcal{C}(\ell)$ after using the $2\ell + 1$ samples to infer it:

$$\left(\frac{\Delta\mathcal{C}(\ell)}{\mathcal{C}(\ell)}\right)_{\text{Cosmic Variance}} = \sqrt{\frac{2}{2\ell + 1}}. \tag{2.58}$$

In practice, it is not possible to achieve this limit, because even if an instrument observes the full sky (for instance the satellite experiments COBE, WMAP, and Planck), a large amount of foreground emission in the Milky Way plane force us to mask some parts of the sky. For a quantification based on a fraction f_{sky} of the full sky, the error bar would increase by approximately a factor $1/\sqrt{f_{\text{sky}}}$.
To obtain an expression for $\mathcal{C}(\ell)$ in terms of $\Theta_\ell(k)$, it is common to introduce the following ratio in which $\mathcal{R}$ is the primordial curvature perturbation:

$$\mathcal{T}(\mathbf{k}, \hat{\mathbf{p}}) \equiv \frac{\Theta(\mathbf{k}, \hat{\mathbf{p}}, \eta_0)}{\mathcal{R}(\mathbf{k})}. \tag{2.59}$$

Now computing $\langle \Theta(\mathbf{k}, \hat{\mathbf{p}}) \Theta^*(\mathbf{k}', \hat{\mathbf{p}'}) \rangle$ and taking into account the fact that $\mathcal{T}$ does not depend on the initial amplitude of each mode and it is not random, so it can be removed from the averaging over the distribution, by expanding τ in Legendre polynomials and doing the integrations we have:

$$\mathcal{C}(\ell) = \frac{2}{\pi} \int_0^\infty dk\, k^2 P_{\mathcal{R}}(k) |\mathcal{T}_\ell(k)|^2, \tag{2.60}$$

where $P_{\mathcal{R}}(k)$, is the definition of the power spectrum of curvature perturbations (see also Eq. 2.48). Indeed for a given ℓ, then the variance $\mathcal{C}(\ell)$ is an integral over all Fourier modes of the variance of $\Theta_\ell(k)$, which is given by $|\mathcal{T}_\ell(k)|^2$ times the variance of curvature perturbations.

5.2 The CMB power spectrum

Since in this thesis in chapter 4, we investigate the CMB spectra, ISW effect and CMB lensing as well as extensively using the Planck data set [2] (including lensing and polarization), it is important to summarize the physics of CMB fluctuations.

5.2.1 Large angular scales

To have a quantitative idea of the large-angle CMB anisotropies, we have to take into account the extremely large-scale modes (low ℓs). The computation of large-angle anisotropy has been done for the first time by (*Sachs & Wolf 1967*) [87]. (Here we only explain the main steps, for a comprehensive review see [5]) It has been argued that at these large-scales we can neglect the dipole moments of the photon perturbations and the only remained terms is determined by $\Theta_0 + \Psi$ evaluated at the recombination. The large-scale photon perturbation satisfies $\Theta_0(\mathbf{k}, \eta_*) = -\Phi(\mathbf{k}, \eta_*) + \mathcal{R}(k) = \frac{2}{5}\mathcal{R}(\mathbf{k})$ (regarding the fact that in matter dominated era $\Phi = 3/5\mathcal{R}$). The observed anisotropy $(\Theta_0 + \Psi)$ which is a good approximation to $\Theta_0 - \Phi$ (since $\Psi \simeq -\Phi$) becomes: $\Theta_0 + \Psi = -\frac{1}{5}\mathcal{R}(\mathbf{k})$. Then using Eq. 2.60 we have:

$$\mathcal{C}(\ell) \simeq \frac{2}{25\pi} \int_0^\infty dk\, k^2 P_{\mathcal{R}}(k) |j_\ell[k(\eta_0 - \eta_*)]|^2. \tag{2.61}$$

Using the definition of the curvature perturbation 2.48 we have:

$$\mathcal{C}(\ell) \simeq \frac{4\pi}{25} \mathcal{A}_s k_p^{1-n_s} \int_0^\infty dk\, k^{n_s-2} j_\ell^2[k(\eta_0 - \eta_*)] = \frac{4\pi}{25} \mathcal{A}_s (\eta_0 k_p)^{1-n_s} \int_0^\infty dx\, x^{n_s-2} j_\ell^2(x). \tag{2.62}$$

Considering the fact that $\eta_* \ll \eta_0$ and defining the integration variable $x \equiv k\eta_0$ we have the second equality of the Eq. 2.62. The integral over the spherical Bessel function can be analytically expressed in terms of the Gamma function as follows:

$$\mathcal{C}(\ell) \simeq 2^{n_s-2} \frac{\pi^2}{25} \mathcal{A}_s (\eta_0 k_p)^{1-n_s} \frac{\Gamma(\ell + \frac{n_s}{2} - \frac{1}{2})}{\Gamma(\ell + \frac{5}{2} - \frac{n_s}{2})} \frac{\Gamma(3 - n_s)}{\Gamma^2(2 - \frac{n_s}{2})}. \tag{2.63}$$

If the spectrum is scale-invariant, $n_s = 1$, then the first ratio of the gamma functions $\Gamma(\ell)/\Gamma(\ell + 2)$ would be equal to $[\ell(\ell+1)]^{-1}$ ans also $\Gamma(2)/\Gamma^2(3/2) = 4/\pi$, therefore

$$\ell(\ell+1)\mathcal{C}(\ell) = \frac{8}{25} \mathcal{A}_s, \tag{2.64}$$

which is a constant. $\ell(\ell+1)\mathcal{C}(\ell)$ is the variance of the temperature anisotropies per logarithmic interval in ℓ and it is common to plot $\ell(\ell+1)\mathcal{C}(\ell)$ vs ℓ which then becomes approximately constant at low ℓ.
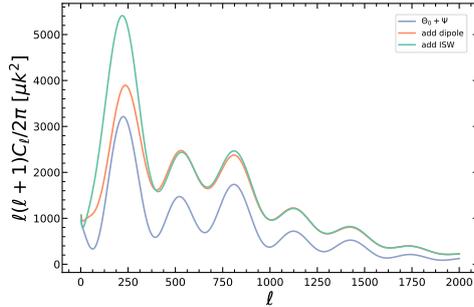
Figure 2.5: The blue line is obtained if only the monopole at recombination were present. It contains most of the structure of the final anisotropy spectrum. The orange line shows the CMB power spectrum including the dipole. Since the dipole is out of phase with the monopole, the troughs become less pronounced. Adding the integrated Sachs–Wolfe effect (green line) enhances the anisotropy mostly on scales comparable to or larger than the horizon at recombination. The first peak gets most of the additional power. We use the CAMB code [4] to obtain the CMB power spectra.

5.2.2 Acoustic peaks

On scales inside the horizon at recombination the anisotropy spectrum depends on the monopole, Θ_0, the dipole, Θ_1, and the integrated Sachs-Wolf effect: $\propto \int d\eta (\Psi - \Phi)'$. All these contributions are shown in Fig. 2.5.

At recombination the monopole $(\Theta_0 + \Psi)(k, \eta_*)$ free streams to us today and creates anisotropies on angular scales $\ell \sim k\eta_0$. The first interesting thing about the monopole (which is shown in Fig. 2.5 in blue) is the fact that the "zeros" in the monopole spectrum at $\ell \sim 70, 400, 650$ and 1000, etc are smoothed out because many Fourier modes contribute to anisotropy on a given angular scale. The other feature of free-streaming is the fact that the positions of the peaks are not precisely on angular scale $\ell = k\eta_0$. Rather, there is a noticeable shift, suggesting that a given k-mode contributes to a slightly smaller ℓ than we anticipated. This shift partially comes from the spherical Bessel function in which the peaks come not at $\ell = k\eta_0$ but at slightly smaller ℓ values.

The dipole at recombination is smaller than the monopole and is not in phase with it (see the orange line in Fig. 2.5). The effect of adding the dipole to the monopole is indeed raising the overall anisotropy level. In particular it is lowering the amplitude of the peaks in comparison with only monopole case. In other words at the positions where the monopole contributes least to the anisotropy, the dipole contributes the most.

The third contribution is from the integrated Sachs–Wolfe effect due to the time evolution of the potentials after recombination, which is mostly due to the fact that the energy density in radiation is not entirely negligible at recombination. In Fig. 2.5 in green we showed what happens when we include the ISW effect in to the anisotropy spectrum. If the Universe was purely matter dominated there would be no such an effect. But, the transition to pure matter domination is not immediate, and even for $a_{eq} \sim 10^{-4}$, an ISW effect occurs right after recombination. This effect is called early ISW effect. The late-time ISW effect occurs when potentials decay during the dark energy epoch at $z \lesssim 1$. Indeed this late-time effect is restricted to extremely large scales, $\ell \lesssim 30$, and it is hardly visible when we plot $\mathcal{C}(\ell)$ as a function of ℓ. We will talk about the ISW effect more in the chapter 4.

6 Large-Scale Structures in linear regime

In this section, we concentrate on the density and velocity perturbations of the dark matter. To study these perturbations apart from the attractive force of gravity, we have to take into account two counteracting effects: **first**, the effect of expanding the background Universe which drags all different species of particles apart. Considering this effect, if we have a faster expansion, the structures in the Universe would grow slower. Imagine, if we do not have any expansion in the Universe, a small overdensity could grow exponentially through gravity which is the only force affecting. Since we have the expansion, this exponential growth is slowed down to a power-low behavior or even logarithmic growth in time. That is why the structures grow more slowly during radiation domination rather than the matter domination epoch and the growth rate of structures slows down once dark energy begins to dominate. The **second** effect only affects baryons and photons (not the cold dark matter at least) which brings into play the concept of pressure. Since gas tends to move through the

direction of lower pressure, the overdensity in baryons does not accumulate as fast as the dark matter. For the moment we focus on dark matter evolution. Since in general, we have two different gravitational growth: (super-horizon, when $k\eta \gg 1$, and sub-horizon, when $k\eta \ll 1$), it would be pedagogical to see the evolution of dark matter in short, long, and medium wavenumbers. Fig. 2.6 - left panel, illustrates the evolution of the gravitational potential as a function of scale factor for different long and short wavelength modes (small and large values of wavenumbers, k, respectively). At high redshifts (smaller scale-factor), all of the modes are outside of the horizon ($k\eta \ll 1$) and the gravitational potential seems to be constant. At intermediate redshifts, the wavelengths enter the horizon and the Universe evolves from radiation-dominated are ($a \ll a_{eq}$) to the matter-dominated era ($a \gg a_{eq}$). The left panel of Fig. 2.6 represents that the gravitational potential affected strongly by crossing the horizon. The long-wave modes (large-scale ones, small k,) enter the horizon after the matter-radiation equality, a_{eq} while the small-scale modes much sooner (before equality). At late times all the modes evolve identically at least before dark energy dominates and cause decay in the potential.
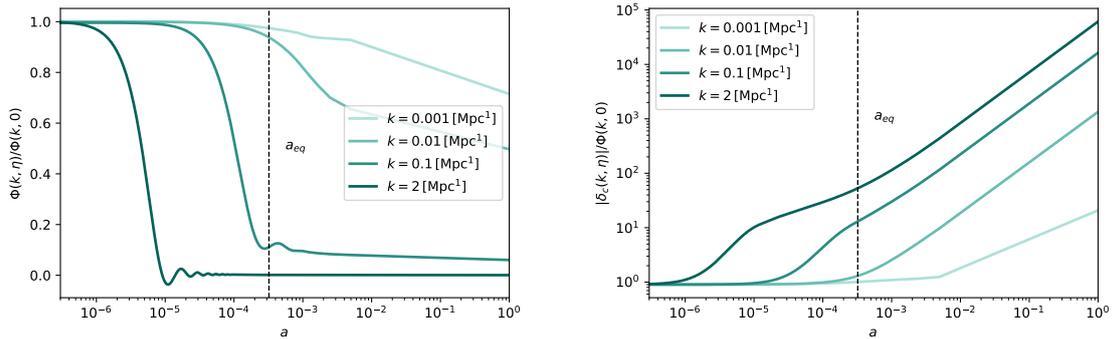
Figure 2.6: **Left panel:** Linear evolution of the gravitational potential, Φ, as a function of scale factor for different wavelength modes in the concordance ΛCDM model for different large-scale and small-scale modes. The vertical line represents the matter-radiation equality a_{eq} in both panels. **Right panel:** The evolution of dark matter density perturbation as a function of scale factor for the same wavelength modes shown in the left panel. We see that during matter-domination, when we are in sub-horizon scales, all different modes grow identically. We used the Boltzmann code CAMB [4] to compute the evolution of the potential and CDM density field.

At late epochs, when all modes are evolving identically, schematically if we want to relate the potential to the primordial curvature perturbation, $\mathcal{R}$, generated during inflation, which we introduced it in Eq. 2.48, we can write:

$$\Phi(\mathbf{k}, a) = \frac{3}{5}\mathcal{R}(\mathbf{k}) \times \left\{ \text{Transfer Function}(\mathbf{k}) \right\} \times \left\{ \text{Growth Factor}(a) \right\}. \tag{2.65}$$

The Transfer Function in the schematic equation above is a function of modes, that describes how the perturbations evolve through the horizon crossing and explains the transition from radiation-dominated era to the matter-dominated one. But the Growth Factor indeed is a function of scale factor and there is no wavenumber relation in it, indeed it is there to describe how perturbations grow at late time. Defining the transfer function as $T(k) \equiv \Phi(\mathbf{k}, a_{\text{late}})/\Phi_{\text{large-scale}}(\mathbf{k}, a_{\text{late}})$ and taking in to account the fact that conventionally the transfer function on large-scales is defined to be equal to 1. We have also the following convention for the growth factor:

$$\frac{\Phi(\mathbf{k}, a)}{\Phi(\mathbf{k}, a_{\text{late}})} \equiv \frac{D_+(a)}{a} \qquad (a > a_{late}). \tag{2.66}$$

If we want to write Eq. 2.65 in a more mathematical way, we have to consider since all sub-horizon modes evolve identically during matter-domination (as illustrated in the right panel of Fig. 2.6), the growth factor would be $D_+(a) = a$.

$$\Phi(\mathbf{k}, a) = \frac{3}{5}\mathcal{R}(\mathbf{k})T(k)\frac{D_+(a)}{a} \qquad (a > a_{late}). \tag{2.67}$$

In the right panel of the Fig. 2.6, we could see the evolution of the CDM overdensity of matter follows the evolution of the potential. Note that for all modes inside the horizon, when the potential is constant, the overdensity grows in time, $\delta_c(\mathbf{k}, a) \propto D_+(a)$. In the late Universe baryons closely follow the dark matter, so we typically describe them together in the form of matter overdensity δ_m. In chapter 5, we show that this convention can be broken. Using Poisson's equation we can relate the matter overdensity to the potential at late time for large k.

$$k^2 \Phi(\mathbf{k}, a) = 4\pi G \rho_m(a) a^2 \delta_m(\mathbf{k}, a) \qquad (a > a_{late}, k \gg aH). \tag{2.68}$$

Using $\rho_m = \Omega_m \rho_{cr}/a^3$ and $4\pi G \rho_{cr} = (3/2) H_0^2$ for the background density together with Eq. 2.67, we can relate the overdensity to the primordial potential

$$\delta_m(\mathbf{k}, a) = \frac{2}{5} \frac{k^2}{\Omega_m H_0^2} \mathcal{R}(\mathbf{k}) T(k) D_+(a) \qquad (a > a_{late}, k \gg aH), \tag{2.69}$$

In the case of adiabatic perturbations as we sow in the previous section, using $P_\mathcal{R}(k) = (2\pi^2/k^3)\mathcal{A}_s(k/k_p)^{n_s-1}$, finally we can write the "linear power spectrum" at late times as

$$P_L(k, a) = \frac{8\pi^2}{25} \frac{\mathcal{A}_s}{\Omega_m^2} D_+^2(a) T^2(k) \frac{k_s^n}{H_0^4 k_p^{n_s-1}}. \tag{2.70}$$
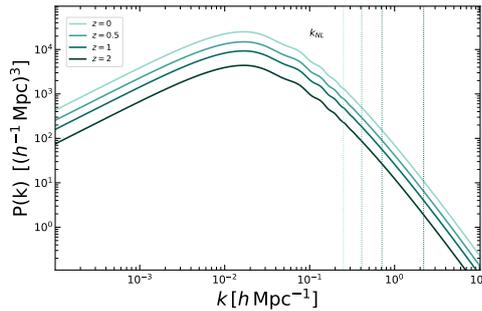
Figure 2.7: The linear matter power spectrum at different redshifts in fiducial ΛCDM cosmology. The vertical dotted lines indicate $k_{\mathrm{NL}}(z)$ at different redshifts. Note that these scales are evolving approximately in linear form. To obtain the linear matter power spectra here we used the CAMB code [4]. The $k_{\mathrm{NL}}(z)$ scales are computed using the condition in Eq. 2.71.

Fig. 2.7 represents the linear matter power spectrum in the fiducial ΛCDM cosmology at different redshifts. On large scales we see the expected power-low behavior $P_L(k) \propto k^{n_s}$, but in small-scales the power spectrum turns over. The vertical lines in this figure shows the scale k_{NL} in which we can not ignore non-linearities. To compute this scale we consider an infinitesimal logarithmic wavenumber bin $|\ln k' - \ln k| < \epsilon$, where ϵ is infinitesimally small, and then we use the variance of density perturbations $\Delta_L^2(k)$ generated by the modes inside this bin as we can see in Eq. 2.71.

$$\Delta_L^2(k, a) = \frac{1}{\epsilon} \int_{|\ln k' - \ln k| < \epsilon} \frac{d^3 k'}{(2\pi^3)} P_L(k', a) = \int_{|\ln k' - \ln k| < \epsilon} k'^3 \frac{dk'}{k'} \int \frac{d\Omega'}{(2\pi^3)} P_L(k', a) = \frac{k^3 P_L(k, a)}{2\pi^2}. \tag{2.71}$$

The regime where $\Delta_L \ll 1$ corresponds to small inhomogeneities, while $\Delta_L \gtrsim 1$ indicates nonlinear perturbations. For instance solving the condition $\Delta_L^2(k_{\mathrm{NL}}, a) \approx 1$ for k_{NL} yields $k_{\mathrm{NL}}(a = 1) \approx 0.25 \, h \, \mathrm{Mpc}^{-1}$.

This would be interesting to see how the shape of the linear matter power spectrum (Eq. 2.70), changes with different values of Ω_m in the Euclidean ΛCDM Universe. Fig. 2.8 shows that doing so will clearly change the shape of the power spectrum and in particular the turnover scale, $k_{eq} \equiv a_{eq} H(a_{eq}) = \sqrt{2\Omega_m} H_0 a_{eq}^{-1/2}$. As we decrease the value of Ω_m, the matter-radiation equality shifts to later times so that k_{eq} pushed to lower values. the opposite behavior would happens is we increase the amount of Ω_m.
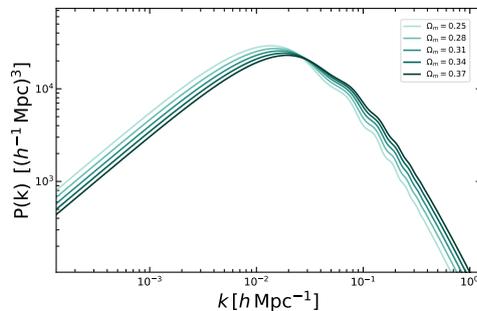
Figure 2.8: The linear matter power spectrum at $z = 0$ in the concordance ΛCDM cosmology. Here we kept Ω_b and h fixed and keep varying Ω_m, considering the fact that we want to remain in Euclidean ΛCDM Universe so we kept $\Omega_K = 0$ and then by varying Ω_m we had to change Ω_Λ as well, using the condition: $\Omega_m + \Omega_\Lambda = 1$. Changing Ω_m changes the epoch of equality (k_{eq}) and also the shape of the matter power spectrum. (By increasing Ω_m here we mean increasing Ω_{cdm} with a fixed value for Ω_b. since in this section we are mostly interested in CDM evolution.) We used the CAMB code [4] to compute the matter power spectrum.

<div style="text-align: right; font-size: 4em; color: gray;">**3**</div>

Cosmological observables: probes & methods

In the previous chapter, we summarized the concordance model of cosmology together with the fundamental Einstein-Boltzmann equations and the definition of physical quantities that will be of use later. We introduced the CMB and the contributions of the observed CMB anisotropies. Moreover, we have shown how small perturbations with $\delta \ll 1$ can be treated analytically in the linear approximation. However, as perturbations grow and $\delta \sim 1$, the predictive power of these solutions decreases. Higher-order approximations (such as n-loop order approximations [88] and renormalized perturbation theory [89]) only converge slowly, and in general, a full non-linear treatment using numerical simulations is needed.

In this chapter we begin with introducing cosmological parameters on the CMB spectrum (Sec. 1), then after introducing statistical analysis to explore parameter space which is a well-known method for both CMB and LSS parametrizations (Sec. 2.1) and the two-point statistics which is a practical method in LSS clustering (Sec. 2.2), we will turn to the probes of the matter distribution in the Universe, galaxy clustering as a tracer of large-scale matter distribution particularly we will concentrate on the BAO feature. We will also refer to the full non-linear exploration of the growth of structures using cosmological N-body simulations and then we will outline how structures and underdense regions as identified in N-body simulations can be analyzed (Sec. 3).

The ambitious goal of this chapter is thereby to introduce at a quantitative level all the observables and methods that will be used in the core of the Thesis.

1 Cosmological parameters

The CMB anisotropy power spectrum has a wealthy structure and its shape depends on cosmological parameters. We can constrain the various parameters by measuring the CMB power spectrum precisely. The price of this multi-dimensional parameter space is that there are partial degeneracies. It means that the effect of varying one parameter can be mimicked by varying, in general, several other parameters in specific ways. In the following we considered the six ΛCDM parameters:

- Curvature parameter, $\Omega_K \equiv 1 - \Omega_m - \Omega_\Lambda$, often set to zero in the concordance model, Or Cosmological constant, Ω_Λ (Or indeed the $100\theta_{\mathrm{MC}}$ parameter, which is the $100\times$(the ratio of the sound horizon to the angular diameter distance and commonly used in the COSMOMC code [90])

- Normalization of the primordial spectrum, $\mathcal{A}_s$

- Scalar spectral index, n_s

- Reionization, parametrized by the optical depth $\tau_{\mathbf{rei}}$ to a redshift after recombination is completed

- Baryon density, $\Omega_b h^2$

- CDM density, $\Omega_c h^2$

We note that obviously, this list does not include all possible cosmological parameters. Some favorites missing are neutrino masses (we will set the sum of neutrino masses to its minimum experimentally allowed value, $\sum m_\nu = 0.06\,\mathrm{eV}$, the equation of state for dark energy w (fixed here at -1 corresponding to a cosmological constant), and tensor to scalar ratio (amplitude r fixed to zero, explained in chapter 2, Sec. 4). The reason that these parameters were not in the main cosmological parameter list is: these parameters are not directly constrained by the CMB temperature power spectrum. Dark energy and neutrino masses are best constrained by combining the CMB with large-scale structure probes and tensor to scalar ratio most constrained by CMB polarization. In the following, we will see the effect of each parameter on the CMB anisotropy power spectrum.

- **Curvature and Λ or angular diameter distance:**

In a non-Euclidean Universe (open or closed) the photon geodesics which start parallel to each other, would converge or diverge. In the case of an open Universe, as a result of the convergence of the photon trajectories, a fixed physical scale in CMB power spectra, for instance, the position of the first peak gets projected on to a much smaller angular scale therefore, the peaks shift to higher ℓ. In a closed Universe, the opposite happens. Fig. 3.1 shows the effect of varying Ω_K on the CMB anisotropy spectrum.
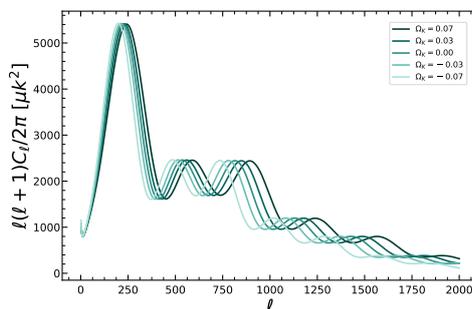
Figure 3.1: The CMB anisotropy spectrum in an open, closed and Euclidean Universe. The peaks are shifted to smaller scales for an open Universe ($\Omega_K > 0$) and to larger scales for a closed Universe ($\Omega_K < 0$). The only varying parameters here are Ω_Λ and Ω_K and we kept fixed the rest of the cosmological parameters. We used the Boltzman code, CAMB [4] to gain the CMB spectrum.

The magnitude of this effect is determined by the angular diameter distance to the last scattering surface (see chapter 2, Sec. 2.1, Eq. 2.10, for a brief explanation on angular diameter distance). In a Euclidean Universe, we can have it by $\eta_0 - \eta_*$ and in a Universe with curvature by

$$d_A = \frac{a}{H_0\sqrt{|\Omega_K|}} \begin{cases} \sinh[\sqrt{\Omega_K}H_0\chi], & \Omega_K > 0, \\ \sin[\sqrt{-\Omega_K}H_0\chi], & \Omega_K < 0. \end{cases} \tag{3.1}$$

Because of the large distance to the last scattering, the CMB peaks respond very sensitively to curvature, resulting in a correspondingly tight constraint. Current best constraints on Ω_K, obtained by combining CMB and large-scale structure probes, are at the level of $|\Omega_K| < 0.002$, (from Planck collaboration-2018 [2]). Now an exact Euclidean Universe is only one point in the parameter space in which the sum of the energy densities exactly equals the critical density (see chapter 2, Sec. 1 for the definition of the critical density), and no data will ever rule out all values except for this one point. Changing the cosmological constant has a similar effect to the curvature. It shifts the peak position due to the change in the angular diameter distance to the last scattering (we note that by varying Ω_Λ while keeping $\Omega_m h^2$ fixed, we are also modifying H_0). This fact also explains why the CMB constraint on Ω_K partially degenerates with that on Ω_Λ, if no large-scale structure probes are included to break the degeneracy. Moreover, changing Λ also affects the late-time ISW contribution at $\ell \lesssim 30$, although the constraining power of this effect is unfortunately limited due to the large cosmic variance errors.

- **Amplitude, spectral index, and optical depth**

In chapter 2, Sec. 4, we introduced the amplitude and spectral index of the curvature perturbations. Fig. 3.2 illustrates the changes in the temperature (TT) CMB angular power spectra as the amplitude of the scalar

perturbations or scalar spectral index varied. We see that by increasing $\mathcal{A}_s$, the CMB spectra would increase which is what we expected from amplitude but the effect of n_s is more complicated. By increasing n_s, we see a decrease in the first peak but from the third peak to the last we see an increase. Indeed it depends to the fact that if n_s is < 1 or > 1.
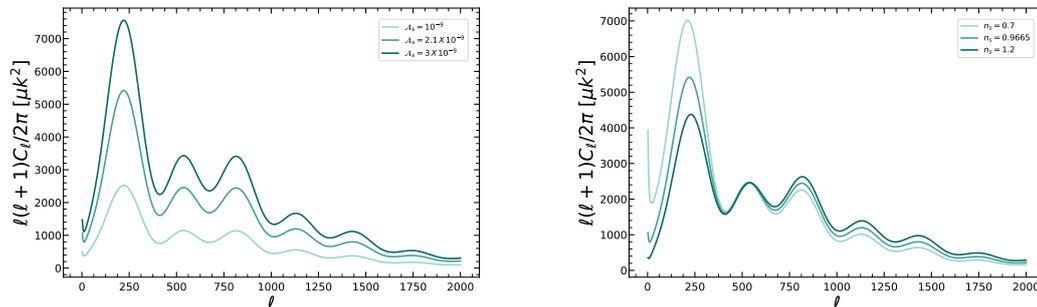
Figure 3.2: The effects of the amplitude of the curvature perturbations $\mathcal{A}_s$, and the spectral index, n_s, on the TT CMB angular power spectra. The left/right panel illustrates the changes in the anisotropy spectrum as the amplitude of the scalar perturbations, $\mathcal{A}_s$/spectral index, n_s is varied. Here we used the CAMB [4] code to obtain the CMB angular power spectrum, which is one of the widely-used Boltzmann solvers.

Indeed changing $\mathcal{A}_s$ by a factor means multiplying all $\mathcal{C}(\ell)$ by the same factor. Shifting $n_s \to n_s + \alpha$ changes the small-scale $\mathcal{C}(\ell)$ by a factor $(\ell/\ell_p)^\alpha$, where ℓ_p is the angular wavenumber corresponding to the pivot scale k_p. On large-scales due to the wide support of j_ℓ for low ℓ this reasoning is not quiet correct. We would like to consider the effect of optical depth due to reionization. After recombination, the gas in the Universe was neutral. On the other hand, most of the gas we observe in the late-time Universe is ionized. For instance, we see no evidence for neutral gas in the absorption spectra of high-redshift quasars until we go back as far as $z \sim 6$ [91]. Therefore at some point, the gas had to be reionized. We believe that this happened between redshifts 15 and 6. After reionization, the CMB photons could scatter off the other free electrons again.
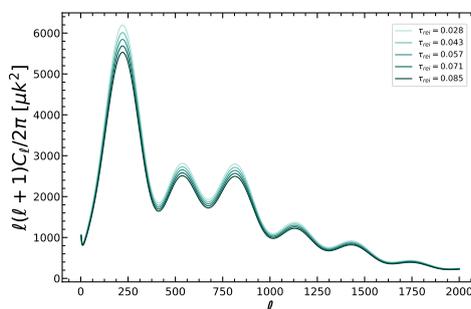
Figure 3.3: The effect of varying the optical depth to reionization on the CMB power spectrum. On scales $\ell \gtrsim 150$ the effect is essentially an overall multiplicative factor, while on very large-scales the CMB is insensitive to τ_{rei}. We obtained the CMB spectrum using the CAMB [4] code.

To understand this in a bit more detail, we can imagine a traveling of photons toward our direction with a temperature $T(1+\Theta)$, where T is the background temperature, and Θ is the perturbation. If these photons hit a region with optical depth τ_{rei}, only a fraction $e^{-\tau_{\mathrm{rei}}}$ will escape and continue on their way to us. Moreover, we will observe a fraction $1 - e^{-\tau_{\mathrm{rei}}}$ scattered into the beam while traveling through the ionized region (since scattering conserves the total number of photons). These scattered-in photons come from all different directions. So we can assume a mean temperature, T, for them. Therefore the temperature we see today is

$$T(1+\Theta)e^{-\tau_{\mathrm{rei}}} + T(1 - e^{-\tau_{\mathrm{rei}}}) = T(1 + \Theta e^{-\tau_{\mathrm{rei}}}). \tag{3.2}$$

Subtraction from the mean temperature T yield the fractional anisotropy as the primordial one set up at $z \simeq 1100$, multilied by $e^{-\tau_{\mathrm{rei}}}$. This scattering only affects the perturbations inside the horizon at the reionization epoch, which means only the multipoles ℓ larger than $\eta_0/\eta_{\mathrm{rei}}$ would be suppressed by $e^{-\tau_{\mathrm{rei}}}$. This effect can be

seen in Fig. 3.3 which shows the effect of changing $\tau_{\rm rei}$ on the CMB power spectrum. By increasing $\tau_{\rm rei}$ we see a suppression on small-scale anisotropies. But indeed for scales $\ell \lesssim 100$ we see no effect.

- **Baryon and CDM densities**

Now we turn to the effect of varying the baryon density, $\Omega_b h^2$ and the CDM density $\Omega_c h^2$. Fig. 3.4 illustrates these effects on the CMB anisotropy power spectra. These changes lead to richer variations in the anisotropy spectrum Other than a shift or a tilt. Indeed, they induce a small relative shift in the locations of the peaks and troughs in the spectrum, as well as changing their amplitudes.
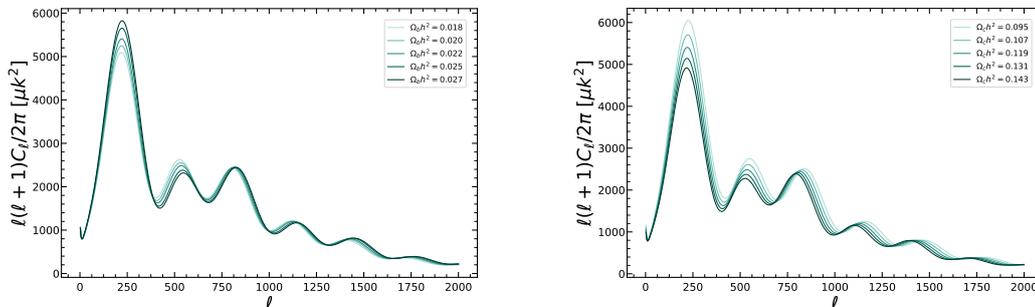
Figure 3.4: **Left panel:** The effect of changing the baryon density $\Omega_b h^2$ on the CMB anisotropy spectrum. **Right Panel:** Changes in the anisotropy spectrum as the cold dark matter density $\Omega_c h^2$ is varied. In each case we kept a Euclidean Universe and compensate the change in the density parameters through Ω_Λ. The CMB spectrum obtained using the CAMB [4] code.

To understand these effects better, we have to consider the fact that since in a Euclidean Universe, inhomogeneities on scales k show up at $\ell = k\eta_0$, the peaks will show up at $\ell_{pk} \simeq k_{pk}\eta_0 \simeq n\pi\eta_0/r_s(\eta_*)$ (in which r_s is the definition of the sound horizon $r_s(\eta) \equiv \int_0^\eta d\tilde{\eta} c_s(\tilde{\eta})$ where $c_s \equiv \sqrt{\dfrac{1}{3\left(1 + \dfrac{3\rho_b}{4\rho_\gamma}\right)}}$ stands for the sound speed of the baryon-photon fluid.)

The effect of changing the baryon density (Left panel of Fig. 3.4) is a shift in the peak locations due to the change in the sound horizon $r_s(\eta_*)$, as well as modifications in the heights of the peaks. From the left panel of Fig. 3.4 it is visible that when the baryon density is large, the ratio of the heights of the odd to even peaks is higher. Another effect is that increasing the baryon density reduces the diffusion length. So larger baryon density means damping moves to smaller angular scales, so the anisotropy spectrum on scales $\ell > 1000$ is larger when we increase the baryon density.

Since the gravitational potential is dominated by cold dark matter, varying the CDM density $\Omega_c h^2$ changes the driving term for the acoustic oscillations. In addition, CDM determines to a large extent the epoch of equality, affecting both the evolution of perturbations (more growth by increasing $\Omega_c h^2$) and the early ISW effect (less ISW for increased $\Omega_c h^2$, since the potentials decay less after recombination).

2   Probes of structures

In the previous section, we sow there is a wealth of information that can be extracted from the CMB power spectrum which is sensitive to cosmological parameters. Other than CMB, there are many observables that indirectly probe the matter distribution in the Universe such as galaxy clustering, which uses galaxies (or any other astronomical objects) as tracers of large-scale matter distribution. In this section, before explaining the LSS probes, we focus on the well-known methods to explore parameter space and constraining them which would be a useful tool for both CMB and LSS parametrizations, then before getting to the details of the galaxy clustering, we briefly explain the two-point statistics which would be a useful method for characterizing the clustering.

2.1   Statistical analysis

In the previous section, we introduced the cosmological parameters where can be constrained by measuring precisely the CMB angular power spectra. In this section, we focus on the analysis techniques to deal with the

complex current cosmological data. We begin by introducing the concepts of likelihood, prior, and posterior. Then we move to explain the Markov Chain Monte Carlo (MCMC) sampling method, which allows for an efficient determination of best-fit parameters and error bars even for very complex likelihoods.

2.1.1 The likelihood function

The likelihood function is defined as the probability that an experiment yields the observed data given a theory. Once we have the likelihood function, we can determine the parameters of the theory along with errors. Imagine, as an example, if we want to measure somebody's weight scientifically, we would set up for instance 100 different scales and record the weight of the person in each of these different scales. The likelihood function, $\mathcal{L}$, gives the probability of getting a hundred numbers given a theory. If we show the person weigh by w, with the noise drawn from a Gaussian distribution with mean zero and variance σ_w^2, then our theory has two free parameters: w and σ_w^2. If only one data point d was taken, the likelihood, $\mathcal{L}$, which is the probability of getting d given the theory would be defined as the following which in the limit of very small σ_w, would sharply peak at $d = w$.

$$\mathcal{L}(d|w, \sigma_w) \equiv P(d|w, \sigma_w) = \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left\{\frac{-(\mathrm{d} - \mathrm{w})^2}{2\sigma_\mathrm{w}^2}\right\}. \tag{3.3}$$

Since we assume to make $m = 100$ independent measurements, the likelihood function would be the product of all individual likelihood functions as follows:

$$\mathcal{L}(\{d\}_{i=1}^m|w, \sigma_w) = \frac{1}{(2\pi\sigma_w^2)^{m/2}} \exp\left\{\frac{-\Sigma_{\mathrm{i}=1}^\mathrm{m}(\mathrm{d} - \mathrm{w})^2}{2\sigma_\mathrm{w}^2}\right\}. \tag{3.4}$$

Since we are interested in the values of the theoretical parameters (w and σ_w), so we can use the probability theory: $P(A, B) = P(B|A)P(A) = P(A|B)P(B)$, which gives us:

$$P(w, \sigma_w|\{d_i\}) = \frac{P(\{d_i\}|w, \sigma_w)P(w, \sigma_w)}{P(\{d_i\})}. \tag{3.5}$$

This is the so-called Bayes' theorem. To get the probability of the theory, $P(w, \sigma_w|\{d_i\})$, which is what we are searching for, we need the likelihood function (the first term in the numerator) and also the prior probability, $P(w, \sigma_w)$. So then we can write: $P(w, \sigma_w|\{d_i\}) \propto \mathcal{L}(\{d_i\}|w, \sigma_w)P_{\mathrm{prior}}(\{d_i\})$.

Other than the proportionality constant, which is independent of the parameters, the resulting probability distribution is the so-called "posterior" for w, σ_w given the data. In cosmology, we are mostly interested in the posteriors for cosmological parameters (introduced in Sec. 1). The idea of using priors to get the posterior might seem a bit unsatisfying since it added some ambiguity to the posterior. Indeed if we want to be conservative, we usually use a prior with a uniform distribution. The primary purpose of putting a prior is to include additional information to incorporate the information consistently. To adopt a uniform prior for the parameters w and σ_w in our example, if we want to find a parameter space in which $P(w, \sigma_w|\{d_i\})$ would be the largest we need to differentiate the likelihood function, $\mathcal{L}$, with respect to each of the parameters,

$$\frac{\partial\mathcal{L}}{\partial w} = \frac{\Sigma_{j=1}^m(d_j - w)}{\sigma_w^2(2\pi\sigma_w^2)^{m/2}} \exp\left\{-\frac{\Sigma_{\mathrm{j}=1}^\mathrm{m}(\mathrm{d_j} - \mathrm{w})^2}{2\sigma_\mathrm{w}^2}\right\} = 0 \quad \rightarrow \quad \Sigma_{\mathrm{j}=1}^\mathrm{m}(\mathrm{d_j} - \mathrm{w}) = 0, \tag{3.6}$$

Thus the likelihood would be at its maximum when: $w = \hat{w} = (1/m)\sum_{i=1}^m d_i$, which gave us actually the sample mean. The other parameter of the model is the variance σ_w^2 and we can find what is the most probable value of σ_w^2 is given the data as follows

$$\frac{\partial\mathcal{L}}{\partial\sigma_w^2} = \mathcal{L} \times \left[-\frac{m}{2\sigma_w^2} + \frac{\Sigma_{i=1}^m(d_i - w)^2}{2\sigma_w^4}\right] \quad \rightarrow \quad \hat{\sigma_w^2} = \frac{1}{m}\sum_{i=1}^m(d_i - w)^2, \tag{3.7}$$

which gives us the estimator of the variance from the data for a known w. The error on these estimated parameters corresponds to the width of the likelihood function. For instance for parameter w, the posterior, $P(w|\{d_i\})$ which is the likelihood multiplied by prior, allows us to make confidence intervals as follows. For instance, the values of w on either side of the maximum (we call them as w_+ and w_-) have the equal probability and also satisfy the following relation:

$$\int_{w_-}^{w_+} dw P(w|\{d_i\}) = 0.68. \tag{3.8}$$

The relation above, defines the 1-σ (or 68% confidence-level) error bar. The value 0.68 comes from the Gaussian probability to find a value within 1-σ of the mean. In the case of Eq. 3.4, the definition translates into the condition that, at the boundaries of the 1-σ confidence interval, the log-likelihood is reduced from the maximum value by $\Delta \ln\mathcal{L} = -1/2$. Since in this example, the likelihood is Gaussian in w, the errors are symmetric. Using the so-called "Taylor" expansion for the log of the likelihood about the maximum value for w, we have

$$\ln\mathcal{L}(\mathrm{w}) = \ln\mathcal{L}(\hat{w}) + \left.\frac{\partial \ln\mathcal{L}}{\partial w}\right|_{\mathrm{w}=\hat{w}} (w - \hat{w}) + \frac{1}{2}\left.\frac{\partial^2 \ln\mathcal{L}}{\partial \mathrm{w}^2}\right|_{\mathrm{w}=\hat{w}} (\mathrm{w} - \hat{\mathrm{w}})^2 \tag{3.9}$$

$$= \ln\mathcal{L}(\hat{w}) - \frac{\mathrm{m}}{2\sigma_{\mathrm{w}}^2} \cdot (\mathrm{w} - \hat{\mathrm{w}})^2 \xrightarrow{=\ln\mathcal{L}(\hat{w})-1/2} \mathrm{Var}[\hat{\mathrm{w}}] = \frac{\sigma_{\mathrm{w}}^2}{\mathrm{m}}. \tag{3.10}$$

So, the square root of this variance, $\sigma_w/m^{1/2}$, would give us the 1-σ error on the estimation of w. Usually, we have not only a single unknown parameter, instead, we have several of them (a multi-dimensional likelihood function). If we are not interested in some parameters, we have to "marginalize" over them which means we have to work with the likelihood after integrating over these parameters, these parameters could be some nuisance parameters that describe observational or astrophysical effects. For the example of weight here, we can imagine a case in which we are interested in measuring w but we do not have any information about σ_w. Then since the full posterior is $P(w, \sigma_w|\{d_i\})$, the marginalized posterior would be

$$P(w|\{d_i\}) = \int_0^\infty d\sigma_w P(w, \sigma_w|\{d_i\}). \tag{3.11}$$

2.1.2 Sampling the likelihood function

Finding analytically the maximum of the likelihood is not generally possible because of the fact that the likelihood is in general not close to Gaussian in the cosmological parameters, moreover, we usually have to integrate over several nuisance parameters. Perhaps we can use a kind of brute-force approach and compute this likelihood function at many points in parameter space and find its maximum (included a set of best-fit parameters) and describe a contour region in which for instance 95% of the volume lies around this maximum which gives us the 95% confidence-region of the parameter. However, this brute-force approach is impractical for the multi-dimensional parameter sets needed for modern experiments. Imagine we have an algorithm that, given any posterior, returns us some points ("samples") in the parameter space $\{\lambda_\alpha^i\}_{i=1}^{m_{\mathrm{sample}}}$. Suppose these parameters are statistically independent from each other and their distribution follows the posterior distribution. The one dimensional example of this situation presented in Fig. 3.5. Then the best-fit parameter can be calculated using the mean values of these samples and the marginalized error can be estimated as the sample variance as follows

$$\bar{\lambda}_\alpha = \frac{1}{m_{\mathrm{sample}}} \sum_{i=1}^{m_{\mathrm{sample}}} \lambda_\alpha^i, \qquad \mathrm{Var}[\lambda_\alpha] = \frac{1}{m_{\mathrm{sample}-1}} \sum_{\mathrm{i}=1}^{m_{\mathrm{sample}}} (\lambda_\alpha^\mathrm{i} - \bar{\lambda}_\alpha^\mathrm{i})^2. \tag{3.12}$$

If the number of samples is very large, these mean and variance values would converge to the central values of the posterior. Indeed if the number of samples, m_{sample} is sufficiently large, a normalized histogram of λ_α^i gives us the marginalized probability distribution of λ_α and then since this becomes a one-dimensional function, we can find the maximum and the confidence intervals as well.
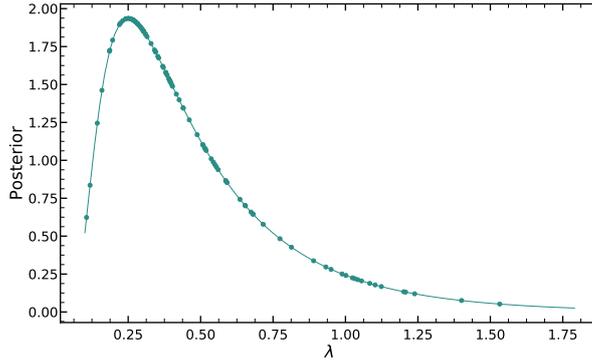
Figure 3.5: The posterior distribution of a parameter λ (the solid curve). The points represent 100 different sampled from the aforementioned distribution. The true mean value and the variance of the posterior distribution are 0.5 and 0.125. The mean value of 100 different samples (points) here is equal to 0.508 and the variance of them is 0.098 which are close to the true mean and variance of the distribution. If more points were sampled the mean and variance would continue to approach the true values of the distribution.

Such algorithms are popular and known as *Markov Chain Monte Carlo (MCMC)*. The "Monte Carlo" refers to the fact that in this process, we use a random generator (like throwing a dice), and the "Markov Chain" stands for the fact that to generate the sample $i + 1$, we only use the previous sample point (sample i), which means that the algorithm does not have a memory. To understand this better, if λ stands for the parameter vector $\{\lambda_\alpha\}$ and $P(\lambda)$ is the desired posterior, we introduce λ' as the output or actually the next sample after λ. Given the Markov nature, the algorithm is described completely by the conditional probability, $K(\lambda'|\lambda)$, that takes us from a sample λ to the next sample λ'. The fundamental requirements on K, for the MCMC sampler to sample from the right posterior, is *detailed balance* which would be written as follows:

$$P(\lambda)K(\lambda'|\lambda) = P(\lambda')K(\lambda|\lambda'). \tag{3.13}$$

Starting with a distribution of λ, that obeys $P(\lambda)$, using the above condition, the algorithm preserve the distribution. One choice for $K(\lambda'|\lambda)$ is the Metropolis-Hastings algorithm which has been invented in the 1950s by Metropolis and later generalized by Hastings. To simplify the situation, we focus on a single parameter and we select a possible next sample λ' by drawing from a distribution centered around λ and symmetric in its two arguments: $g(\lambda', \lambda)$. The simplest example is a Gaussian, $g(\lambda', \lambda) \propto \exp[-(\lambda - \lambda')/2\sigma^2]$. After that, this new sample is "accepted" with the probability as follows

$$p_{\mathrm{acc}}(\lambda', \lambda) = \min\left\{\frac{\mathrm{P}(\lambda')}{\mathrm{P}(\lambda)}, 1\right\}. \tag{3.14}$$

When we evaluate the posterior for λ', and we compute the ratio $\alpha \equiv P(\lambda')/P(\lambda)$, if the value of α is larger than 1, then we accept λ' as our next step in the chain, otherwise, we uniformly draw a random number between zero and one and accept λ' into the chain only if that random number is less than α. Therefore, if the proposed λ' is much less likely than the previous λ it would be unlikely (but not impossible) to be accepted. However if the random number is greater than α, then we would discard λ' and we insert the previous value λ to the chain. Then we repeat and draw a new proposal using the last element of the chain.

$$\frac{P(\lambda')}{P(\lambda)} = \frac{K(\lambda'|\lambda)}{K(\lambda|\lambda')} = \frac{p_{\mathrm{acc}}(\lambda', \lambda)}{p_{\mathrm{acc}}(\lambda, \lambda')}, \tag{3.15}$$

in which the second equality follows since $g(\lambda', \lambda)$ is symmetric. Using Eqs 3.14 and 3.15, the two following cases would happen:

$$\text{if } P(\lambda') < P(\lambda) \rightarrow P_{\text{acc}}(\lambda, \lambda') = 1 \rightarrow P_{\text{acc}}(\lambda', \lambda) = \frac{P(\lambda')}{P(\lambda)},$$

$$\text{if } P(\lambda) < P(\lambda') \rightarrow P_{\text{acc}}(\lambda', \lambda) = 1 \rightarrow P_{\text{acc}}(\lambda, \lambda') = \frac{P(\lambda)}{P(\lambda')}. \tag{3.16}$$

The algorithm could be also generalized to multiple parameters, by performing successive steps in $\lambda_1, \lambda_2, \dots,$ and also one of the advantages of the Metropolis-Hastings algorithm is the fact that we can tune the function g. For instance, for the choice of Gaussian above, σ is a free parameter, which is chosen to be very small, the sampler will take a long time to map out the likelihood function, or might even get stuck near a local maximum. If it is too large, the acceptance rate will be low, because most draws of λ' will end up in low probability regions of parameter space. Consequently, throughout the burn-in phase, the step-size parameter is usually adjusted dynamically. Nowadays there are several cosmological codes for exploring cosmological parameter space, that have implemented the Metropolis-Hastings MCMC algorithms and also several others. Among the ones that I have been using during my P.hD. research projects we can mention COSMOMC [90] which performs brute force (but accurate) theoretical matter power spectrum estimation and C_ℓ calculations with the Boltzmann code CAMB [4], MONTE PYTHON [92] which interfaces with the Boltzmann code CLASS [93], and also emcee [94] package which is an implementation of a MCMC sampler.

2.2   Two-point statistics

Before describing observables and the tracers of LSS, we briefly explain the two-point statistics, and then we present galaxy clustering, which is used to measure BAOs.

Imagine we have N points (astrophysical sources, galaxies, quasars, etc.) in a volume V, naively the simplest descriptor of these points is the average numerical density $\rho_0 = N/V$, which clearly is not sufficient to discriminate among these points. To have an idea about the fact that how much these N points are clustered close to a spot, we need a bit more precise descriptor. That is why we draw our attention to a small volume dV chosen randomly inside the volume V. Then $\rho_0\, dV$ yields the average number of particles in the infinitesimal volume. Using the fact that $dN_{ab} = \langle n_a\, n_b \rangle$ gives us the average number of pairs in volumes dV_a and dV_b (which is the product of the number of particles in one volume times the number of particles in another volume), separated by r_{ab}, then we can define the *2-point Correlation Function (2PCF)* as the next order descriptor of the points as follows

$$dN_{ab} = \langle n_a\, n_b \rangle = \rho_0^2\, dV_a\, dV_b [1 + \xi(r_{ab})], \tag{3.17}$$

which also can be written as a spatial average of the product of the density contrast $\delta(r_a) = n_a/(\rho_0\, dV_a) - 1$ at two different points

$$\xi(r_{ab}) = \frac{dN_{ab}}{\rho_0^2\, dV_a\, dV_b} - 1 = \langle \delta(r_a)\, \delta(r_b) \rangle. \tag{3.18}$$

The 2PCF is defined as a deviation of a probability from the probability expected from a random distribution of tracers [25] and can be written as a sample average over all possible positions as follows:

$$\xi(r) = \frac{1}{V} \int \delta(\mathbf{y})\, \delta(\mathbf{y} + \mathbf{r})\, dV_y. \tag{3.19}$$

To calculate the 2PCF from a data or simulation catalogue for any type of tracer of the LSS, we can use the Landy & Szalay estimator [95] which we have used to calculate the void-tracer correlation functions in chapter 6. Using the data catalogue, D, and an artificial random catalogue, R, for each tracer from the simulation box, where for instance for the void-void cross correlation, $DD(r)$ shows the number of void-void pairs in the data catalogue which are separated by distance r, $DR(r)$ and $RD(r)$ denote the number of void-random pairs (voids from the data catalogue and random positions from a random catalogue) separated by r and $RR(r)$ is the number of random-random pairs with distance r. The estimator reads:

$$\xi(r) = \frac{DD(r) - DR(r) - RD(r) + RR(r)}{RR(r)} \, . \tag{3.20}$$

To relate the definition of the power spectrum which we described in chapter 2, Eq. 2.46 in Fourier Space, to the definition of the 2PCF in the configuration space, we can use the Fourier transform of the density contrast of a density field $\delta(\mathbf{x})$ as

$$\delta_k = \frac{1}{V} \int \delta(\mathbf{x}) \, e^{-i\mathbf{k}\cdot\mathbf{x}} \, dV \, . \tag{3.21}$$

Using the definition of the power spectrum as a variance of the perturbations $P(\mathbf{k}) = V|\delta_k|^2 = V\delta_k\delta_k^*$, we have

$$P(k) = \frac{1}{V} \int \delta(\mathbf{x})\delta(\mathbf{y})e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} \, dV_x \, dV_y \, , \tag{3.22}$$

Using $\mathbf{r} = \mathbf{x} - \mathbf{y}$, and Eq. 3.19, we obtain:

$$P(k) = \int \xi(r)e^{-i\mathbf{k}\cdot\mathbf{r}} \, dV \, . \tag{3.23}$$

2.3 Galaxy clustering

The positions of galaxies in a redshift survey look like the one we have shown in chapter 1, in the left panel of Fig. 1.1. A reasonable statistics that we can compute to compare with theory, would be a three-dimensional power spectrum, $P_{g,\text{obs}}(\mathbf{k})$. If the position of a galaxy from us would be at comoving distance $\chi(z)$ (Eq. 2.8), then the three-dimensional position vector can be written as:

$$\mathbf{x}_{\text{obs}}(z, \theta, \phi) = \chi(z)\hat{\mathbf{n}}(\theta, \phi), \quad \text{where} \quad \hat{\mathbf{n}} = \frac{\mathbf{x}_{\text{obs}}}{|\mathbf{x}_{\text{obs}}|} \, . \tag{3.24}$$

In survey analysis, generally, it is assumed a fiducial cosmology with a distance-redshift relation, $\chi_{\text{fid}}(z)$ (which could be different from the correct one):

$$\chi_{\text{fid}}(z) = \chi(z) + \delta\chi(z) \, . \tag{3.25}$$

Moreover, the observed redshift of a galaxy is a composition of the cosmological redshift and the Doppler shift due to the peculiar velocity of the galaxy $\mathbf{u}_g$ along the line-of-sight:

$$1 + z = \frac{1}{a_{emit}}[1 + u_\parallel], \quad \text{where} \quad u_\parallel = \mathbf{u}_g \cdot \hat{\mathbf{n}} \, . \tag{3.26}$$

Assuming $\mathbf{u}_g = 0$, and the fiducial cosmology to be true, the error made in Eq. 3.24 is given by

$$\Delta\mathbf{x}_{\text{RSD}} = \left.\frac{\partial\mathbf{x}_{\text{obs}}}{\partial\mathbf{u}_\parallel}\right|_{u_\parallel=0} \mathbf{u}_\parallel = \frac{1}{aH}u_\parallel\hat{\mathbf{n}}, \tag{3.27}$$

where the subscript RSD indicates the effect of redshift-space distortion which we will describe in the next subsection. If $u_\parallel > 0$, which means the galaxy is moving away from us, we would have an increase in the distance. Combining Eqs. 3.25 and 3.27, including the effect of a non-fiducial cosmology we have:

$$\mathbf{x}_{\text{obs}} = \mathbf{x} + \left(\delta\chi(z) + \frac{1}{aH} u_{\parallel}(\mathbf{x}) \right) \hat{\mathbf{n}}. \tag{3.28}$$

Now we have the observed coordinates of the galaxies, we want to measure the power spectrum in terms of these positions. To see how this distorted power spectrum relates to the true spectrum of the galaxies, we will follow the derivation in [96] by Kaiser in 1987. Considering the fact that whether we use the observed position of the galaxies or the true positions, the number of the galaxies in a particular region would be the same, we divide the volume of the surveys to different elements (voxels) and then we count the number of galaxies in each volume element. Since the number of galaxies would be the same, we have

$$n_{g,\text{obs}}(\mathbf{x}_{\text{obs}})d^3 x_{\text{obs}} = n_{\text{g}}(\mathbf{x})d^3 x, \tag{3.29}$$

where n_g represent the galaxy density field at $\mathbf{x}$ in real space and $n_{g,\text{obs}}$ shows the galaxy density field in redshift space. Using the infinitesimal volume element around each point in each case: $d^3 x_{\text{obs}} = x^2_{\text{obs}} dx_{\text{obs}} d\Omega$ and $d^3 x = x^2 dx d\Omega$, where $d\Omega$ is the angular volume element and would be identical in both real and redshift space. Therefore,

$$n_{g,\text{obs}}(\mathbf{x}_{\text{obs}}) = n_{\text{g}}(\mathbf{x}) \underbrace{\left| \frac{d^3 x}{d^3 x_{\text{obs}}} \right|}_{= \left| \frac{dx}{dx_{\text{obs}}} \right| \times \frac{x^2}{x^2_{\text{obs}}} \equiv J}, \tag{3.30}$$

where J is the Jacobian which can be calculated using Eq. 3.28

$$J = \left(1 + \frac{\delta\chi}{x} + \frac{u_{\parallel}}{aHx} \right)^{-2} \left| 1 + \underbrace{\frac{d}{dx}\delta\chi}_{\frac{dz}{dx}\frac{d\delta\chi}{dz} = H\delta(H^{-1}) = -H^{-1}\delta H} + \frac{1}{aH}\frac{\partial}{\partial x}u_{\parallel} \right|^{-1}, \tag{3.31}$$

in which $\delta H = H(z) - H_{\text{fid}}(z)$ in the underbraced equation, we also used $dz/dx = dz/d\delta\chi = H$. Discarding some second-order terms, an approximation for a simplified Jacobian is given by

$$J \simeq \bar{J}\left(1 - \frac{1}{aH}\frac{\partial}{\partial x}u_{\parallel} \right); \quad \bar{J} = 1 - 2\frac{\delta\chi(\bar{z})}{\bar{\chi}} + H^{-1}(\bar{z})\delta H(\bar{z}), \tag{3.32}$$

where $\bar{z}$ is the mean value of a fairly narrow redshift slice. This is typically done to avoid an evolving galaxy population.

The number densities of the true and observed galaxies are defined as $n_g = \bar{n}_g(1 + \delta_g)$ and $n_{g,\text{obs}} = \bar{n}_{g,\text{obs}}(1 + \delta_{g,\text{obs}})$, respectively, where $\bar{n}_g$ stands for the average number density of galaxies (usually find by counting galaxies in a redshift slice and dividing by its volume). Using the Eq. 3.30, the observed galaxy overdensity would be written as

$$1 + \delta_{g,\text{obs}}(\mathbf{x}_{g,\text{obs}}) = \bar{J}\left[1 + \delta_g(\mathbf{x}[\mathbf{x}_{\text{obs}}]) - \frac{1}{aH}\frac{\partial}{\partial x}u_{\parallel}(\mathbf{x}[\mathbf{x}_{\text{obs}}]) \right]. \tag{3.33}$$

2.4 Redshift-space distortions

Qualitatively the effect of peculiar velocities on galaxy clustering can be demonstrated as follows: If we have a large-scale overdense region towards which surrounding galaxies are falling, appears squashed in redshift space which means the galaxies closest to us are moving toward the center of the overdense region and therefore they are moving away from us so they would appear farther from us (closer to the center of the overdense region). Since moving galaxies towards each other increases their number density, we expect the clustering in redshift space would be stronger than the real space.

To relate the velocity to the density field we use the continuity equation for the total matter as follows

$$\delta'_m + i\mathbf{k} \cdot \mathbf{u}_m = -3\Phi', \tag{3.34}$$

in which the prime shows derivatives with respect to η (introduced in Eq. 2.9). Working in the sub-horizon scales, we set the right-hand side of the above equation equal to zero. Due to the fact that the time dependance of the linear density field is given by the growth factor $D_+(\eta)$ (see Eq. 2.67), we can solve the equation above for the velocity in terms of the density.

$$\mathbf{u}_m(\mathbf{k}, \eta) = \frac{i\mathbf{k}}{k^2} \frac{D'_+}{D_+} \delta(\mathbf{k}, \eta) = aHf \frac{i\mathbf{k}}{k^2} \delta(\mathbf{k}, \eta), \tag{3.35}$$

where $f = d \ln \mathrm{D}_+ / d \ln \mathrm{a}$, is the linear growth rate (note that the growth rate is close to unity for a ΛCDM Universe and exactly is equal to 1, for a flat matter-dominated cosmology).

In order to relate the galaxy density field δ_g to the underlying matter density field δ_m, we assume a *linear bias* relation as follows (for a comprehensive review on bias formalism see [97]).

$$\delta_g(\mathbf{x}, \eta) = b_1(\eta)\delta_m(\mathbf{x}, \eta). \tag{3.36}$$

This is due to the fact that since galaxies are complicated and they are indeed highly non-linear tracers of large-scale structures, their density is not the same as the underlying matter. The linear bias parameter, b_1, depends sensitively on the galaxy sampled considered and in general, it is redshift dependent. Also since galaxies are discrete tracers, the galaxy density field contains a noise that we will include at the end because it is independent of RSD. Using these results and including Eq. 3.33, with $\bar{J} = 1$, we realize that the overdensity in redshift space is indeed a summation of the overdensity in real space and a correction because of the peculiar velocity.

$$\delta_{g,\mathrm{RSD}}(\mathbf{x}) = b_1\delta_m(\mathbf{x}) - \frac{\partial}{\partial \mathbf{x}} \left[\frac{\mathbf{u}_m(\mathbf{x}) \cdot \hat{\mathbf{x}}}{aH} \right]. \tag{3.37}$$

Now we are going to use *distant-observer* approximation which is actually a flat-sky approximation saying that we can treat the direction vector $\hat{\mathbf{n}} = \mathbf{x}/x$ as fixed, neglecting variations from galaxy to galaxy. This can be justified if the angular scales are small. Since the galaxies are relatively close to each other, we can approximate $\hat{\mathbf{x}} \cdot \mathbf{u}_m \to \hat{\mathbf{e}}_z \cdot \mathbf{u}_m$, where $\hat{\mathbf{e}}_z$ denotes a radial vector pointing to the center of the sky (we choose this to be the $z-$axis). Regarding the distant-observer approximation, we can directly write the Fourier transform of the redshift-space overdensity as follows

$$\delta_{g,\mathrm{RSD}}(\mathbf{k}) = \int d^3x e^{-i\mathbf{k} \cdot \mathbf{x}} \left[b_1\delta_m(\mathbf{x}) - \frac{\partial}{\partial x} \left[\frac{\mathbf{u}_m \cdot \hat{\mathbf{e}}_z}{aH} \right] \right]$$

$$= b_1\delta_m(\mathbf{k}) - if \int d^3x e^{-i\mathbf{k} \cdot \mathbf{x}} \frac{\partial}{\partial x} \left[\int \frac{d^3k'}{(2\pi)^3} e^{-i\mathbf{k}' \cdot \mathbf{x}} \delta_m(\mathbf{k}') \frac{\mathbf{k}'}{k'^2} \cdot \hat{\mathbf{e}}_z \right], \tag{3.38}$$

The second equality comes from Eq. 3.35. The derivative with respect to x would act on the exponential and would bring down the factor of $i\mathbf{k}' \cdot \mathbf{x}$ which we set it to $i\mathbf{k}' \cdot \hat{\mathbf{e}}_z$, then

$$\delta_{g,\mathrm{RSD}}(\mathbf{k}) = b_1 \delta_m(\mathbf{k}) + f \int \frac{d^3 k'}{(2\pi)^3} \delta_m(\mathbf{k}') \left(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}_z \right)^2 \int d^3 x \, e^{i(\mathbf{k}'-\mathbf{k}) \cdot \mathbf{x}}. \tag{3.39}$$

The integral over x yields the Dirac Delta function as follows $(2\pi)^3 \delta_D^{(3)}(\mathbf{k}' - \mathbf{k})$, Using the distant-observer approximation we have

$$\delta_{g,\mathrm{RSD}}(\mathbf{k}) = \left[b_1 + f \mu_k^2 \right] \delta_m(\mathbf{k}). \tag{3.40}$$

In the equation above μ_k is defined to be $\hat{\mathbf{e}}_z \cdot \hat{\mathbf{k}}$, which is the cosine of the angle between the line of sight and the wavevector $\hat{\mathbf{k}}$. The Eq. 3.40 quantifies what we have expected about the large-scale redshift-space distortions. Since $f \mu_k^2 \geq 0$, the apparent overdensity in redshift space is larger than the real space, where it would be simply $b_1 \delta_m(\mathbf{k})$. Another important feature is that the enhancement of the density contrast is for perturbations with wavevector parallel to the line of sight. A perturbation with $\mathbf{k}$ perpendicular to the line of sight, does not experience the redshift space distortions.

In the same manner, the galaxy power spectrum in redshift space depends not only on the magnitude of $\mathbf{k}$ but also on its direction which can be parametrized as follows

$$P_{g,\mathrm{RSD}}(k, \mu_k, \bar{z}) = P_L(k, \bar{z}) \left[b_1 + f \mu_k^2 \right]^2 + P_N, \tag{3.41}$$

in which $P_L(k, \bar{z})$, denotes the linear matter power spectrum that we introduced in chapter 2, Eq. 2.70. In addition b_1 and f have to be evaluated at the mean redshift $\bar{z}$. P_N demonstrate the noise term in the galaxy power spectrum, which is assumed to be "white", which means is a scale independent and constant indeed. If the galaxies are Poisson-sampled from an underlying continuous field, then we have

$$P_N = \frac{1}{\bar{n}_g}. \tag{3.42}$$

Since a simple Poisson sampling is not a realistic assumption for actual galaxies, usually a scale-independent noise P_N would be expected at low k.

By measuring $P_{g,\mathrm{RSD}}(k, \mu_k)$, we can vary both k and μ_k, this allows us to disentangle the contributions multiplied by b_1 and f. This is usually done by performing multipole decomposition of $P_{g,\mathrm{RSD}}(k, \mu_k)$ with respect to μ_k as follows

$$P_{g,\mathrm{obs}}(k, \mu_k) = \sum_l \mathcal{P}_l(\mu_k) P_{g,\mathrm{obs}}^{(l)}(k)$$
$$P_{g,\mathrm{obs}}^{(l)}(k) = \frac{2l+1}{2} \int_{-1}^1 d\mu_k \mathcal{P}_l(\mu_k) P_{g,\mathrm{obs}}(k, \mu_k), \tag{3.43}$$

where $\mathcal{P}_l$ stands for the Legendre polynomials and also get the multipoles of the 2D correlation function in the configuration space, we have

$$\xi_l(r) = i^l \int_0^\infty \frac{k^2 \, d\log k}{2\pi^2} P_{g,\mathrm{obs}}^{(l)}(k) \, j_l(kr), \tag{3.44}$$

where j_l is the spherical Bessel function of order l.

2.5 Alcock–Paczynski effect and BAO feature

In the previous subsection, we saw that galaxies are displaced by $\delta\chi(z)$. At the lowest order, this would be the same for all the galaxies, but due to the fact that the distance-redshift relation evolves differently in different cosmologies, therefore we would have different displacements for different cosmological redshifts. The observed galaxy position $\mathbf{x}_{\text{obs}}$, which is made up of (θ, ϕ) and redshift, z, can be written as a 2 dimensional vector θ, by choosing the origin as follows:

$$\mathbf{x}_{\text{obs}} = \mathbf{0} \Leftrightarrow \theta = \mathbf{0},\ z = \bar{z}, \tag{3.45}$$

where $\theta = \mathbf{0}$ shows a point on the sky near the center of the survey footprint and $\bar{z}$ denotes the central value of the redshift slice. So the transverse components x^1_{obs} and x^2_{obs} could be assigned as follows

$$(x^1_{\text{obs}}, x^2_{\text{obs}}) = \chi_{\text{fid}}(z) \times (\theta^1, \theta^2), \tag{3.46}$$

while the components in true cosmology should be assigned as

$$(x^1, x^2) = \chi(z) \times (\theta^1, \theta^2) = \left[1 - \frac{\delta\chi(z)}{\chi_{\text{fid}}(z)} \right] (x^1_{\text{obs}}, x^2_{\text{obs}}). \tag{3.47}$$

The last component of the position which is the line-of-sight component, x^3_{obs}, would be determined by the redshift as:

$$x^3_{\text{obs}}(z) = \chi_{\text{fid}}(z) - \chi_{\text{fid}}(\bar{z}) \simeq \frac{1}{H_{\text{fid}}(\bar{z})}(z - \bar{z}), \tag{3.48}$$

where we use the expansion to the linear order in $(z - \bar{z})$, using the assumption that we are considering a narrow redshift slice and then we can use $d\chi/dz = 1/H$. Then to assign the line-of-sight component to the galaxy we have

$$x^3(z) \simeq \frac{1}{H(\bar{z})}(z - \bar{z}) = \frac{H_{\text{fid}}(\bar{z})}{H(\bar{z})} x^3_{\text{obs}}, \tag{3.49}$$

using $\delta H(z) = H(z) - H_{\text{fid}}(z)$, and assuming a linear order in δH, we can obtain the following relation

$$x^3(z) = \left[1 - \frac{\delta H(\bar{z})}{H_{\text{fid}}(\bar{z})} \right] x^3_{\text{obs}}. \tag{3.50}$$

Due to the fact that the zero-point of the third axis is defined by a fixed redshift, $\bar{z}$, the displacements between true and assigned positions is only induced by a difference in how $\chi(z)$ and $\chi_{\text{fid}}(z)$ vary around $\bar{z}$. Note that the line-of-sight displacement of the position (Eq. 3.50) is different from the transverse displacement (Eq. 3.47). Then we can summarize the relation between the true and observed galaxy positions due to a wrong distance-redshift relation as

$$\mathbf{x}(\mathbf{x}_{\text{obs}}) = \left([1 - \alpha_\perp] x^1_{\text{obs}}, [1 - \alpha_\perp] x^2_{\text{obs}}, [1 - \alpha_\parallel] x^3_{\text{obs}} \right), \quad \text{where}$$
$$\alpha_\perp = \frac{\delta\chi}{\chi_{\text{fid}}} \bigg|_{\bar{z}}; \quad \alpha_\parallel = \frac{\delta H}{H_{\text{fid}}} \bigg|_{\bar{z}}. \tag{3.51}$$

Given a measurement of $\alpha_\perp$, and $\alpha_\parallel$, we can express $\chi(\bar{z})$ and $H(\bar{z})$ in terms of the known quantities as: $\chi(\bar{z}) = \chi_{\text{fid}}(\bar{z})[1 + \alpha_\perp]$ and $H(\bar{z}) = H_{\text{fid}}(\bar{z})[1 + \alpha_\parallel]$.

Implying rescaling by Eq. 3.51 in Fourier space and using Eq. 3.40, since the number of galaxies in a given volume is independent of the coordinate, we obtain

$$\delta_{g,\mathrm{obs}}(\mathbf{k}_{\mathrm{obs}}) = [b_1 + f\mu_k^2]\delta_m(\mathbf{k})\Big|_{\mathbf{k}=([1-\alpha_\perp]k_{\mathrm{obs}}^1,[1-\alpha_\perp]k_{\mathrm{obs}}^2,[1-\alpha_\parallel]k_{\mathrm{obs}}^3)}. \qquad (3.52)$$

Also using Eq. 3.41, we can write the observed galaxy power spectrum as

$$P_{g,\mathrm{obs}}(\mathbf{k}_{\mathrm{obs}},\bar{z}) = \left(P_L(k,\bar{z})\left[b_1 + f\mu_k^2\right]^2\right)\Big|_{\mathbf{k}=([1-\alpha_\perp]k_{\mathrm{obs}}^1,[1-\alpha_\perp]k_{\mathrm{obs}}^2,[1-\alpha_\parallel]k_{\mathrm{obs}}^3)} + P_N. \qquad (3.53)$$

Note that the equation above contains both redshift-space distortions due to peculiar velocities, and rescaled coordinates based on distance-redshift relation effects. The noise term P_N would be constant and not affected by both effects.

Even if we ignore galaxy velocities, we would have anisotropy in the galaxy power spectrum originated by an incorrect distance-redshift relation. So, if we set $f = 0$, $\alpha_\perp$ would be still different from $\alpha_\parallel$. This means that using the wrong assumed cosmology would produce an anisotropy in the galaxy power spectrum. This fact was found by Alcock-Paczynski in 1979 [98] and is called *Alcock-Paczynski (AP)* effect which depends on the shape of the power spectrum.

Commonly, Eq. 3.53, is a target of galaxy redshift surveys to use BAO feature as a standard ruler, since the matter power spectrum displays an oscillatory modulation, roughly of the form $\cos(k r_s)$ where $r_s \approx 105\,\mathrm{h}^{-1}\mathrm{Mpc}$ which is the sound horizon at recombination (Fig. 3.6). Note that in configuration-space, the BAO feature appears as a peak in a two-point correlation function. (Fig. 3.7). In chapter 6 we will measure the position of this peak using N-body simulations. In the early Universe, this feature was imprinted only in baryons, but since baryons and CDM are coupled by gravity, this oscillatory pattern is transferred to the late-time power spectrum of matter, albeit with smaller amplitude. In chapter 6, we will show and discuss this transformation.
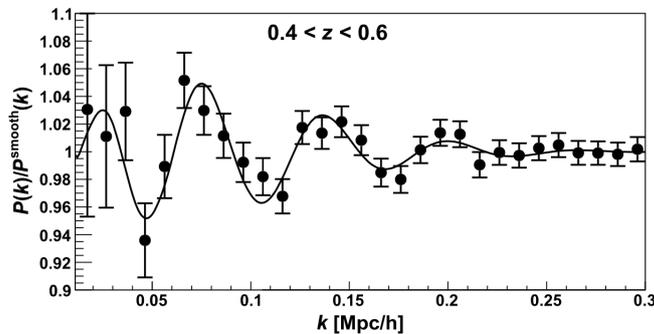
Figure 3.6: The BAO feature in the monopole of the 3D galaxy power spectrum for the CMASS sample (high redshift ($0.4 < z < 0.6$) galaxy sample using a set of color-magnitude cuts) observed by the Baryon Oscillation Spectroscopic Survey (BOSS). Both data (points) and the best-fit model (line) have been divided by a smooth model power spectrum without the BAO feature, in order to enhance the visibility of the feature. The fit parameters include $\alpha_\perp$ and $\alpha_\parallel$ would shift the model along the $k-$axis. Taken from [5] which has been adopted from [6].

Why the BAO feature is important for cosmology? Since we have a well-defined feature in the power spectrum at a *true* comoving scale $k \sim \pi/r_s$, using Eq. 3.53, we can observe the same feature in galaxy power spectrum but at an *apparent* scale, $k_{\mathrm{obs}[k]}$. Indeed r_s is extremely well defined by CMB so measuring this feature in the galaxy power spectrum at redshift $\bar{z}$, would allow us to measure $\alpha_\perp$ and $\alpha_\parallel$ precisely and this in turns allows us to measure $\chi(\bar{z})$ (using $\alpha_\perp$) and $H(\bar{z})$ (using $\alpha_\parallel$). In addition, we can apply all the equations outlined above and methods not only to galaxies but also to any LSS tracers like quasars, the Lyman-alpha forest, and unresolved line emitters and absorbers (known as intensity mapping).

Therefore, we have a two-fold approach that relies both on dynamical and geometrical probes: RSD allows us to measure the growth rate $f\sigma_8$, using the amplitude and anisotropy of the galaxy power spectrum or
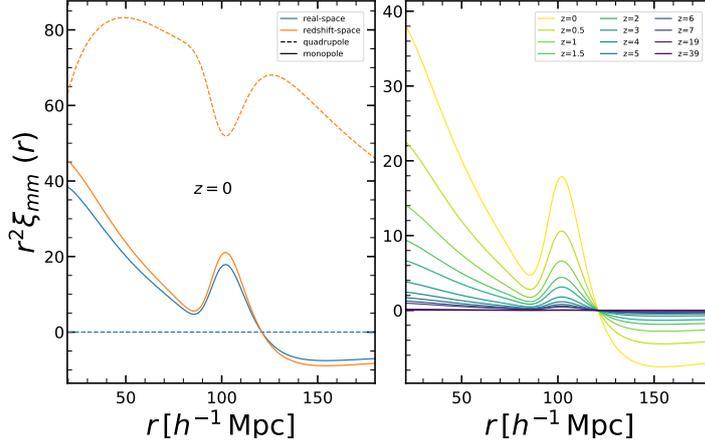
Figure 3.7: Linear theory prediction for matter-matter correlation function using Equation 3.44. **Left panel:** monopole, $l = 0$ (in solid line) and quadrupole, $l = 2$ (in dashed line) predictions in real space (in blue) and redshift space (in orange) at $z = 0$. In real space, the quadrupole term would be equal to zero and in redshift space, the quadrupole term would flip the BAO peak due to the negative sign appeared by implying i^2 in Equation 3.44. **Right panel:** redshift evolution of the monopole correlation function in realspace. All matter-matter two-point correlation functions are multiplied by r^2 in order to enhance the BAO peak.

correlation functions; BAO and AP effect allows us to measure $d_A(\bar{z})$ and $H(\bar{z})$ using the difference between fiducial and true distance-redshift relations.

3 Growth of structure: beyond linear theory

In chapter 2, Sec. 3.2.2, we reviewed the main steps of recovering the relativistic perturbed equations for different components of the Universe. The dominant clustering component of the late Universe is matter, which consists of dark matter (about 80%) as well as baryons (in the form of neutral and ionized gas as well as stars). In this section, we will lump together dark matter and baryons and refer to them as "matter". Indeed the reality is that the baryons behave differently from dark matter since they feel electromagnetic forces. After the decoupling of the baryons from the photons, the baryons cool rapidly (the temperature of the baryons scales as their kinetic energy, so $\propto a^{-2}$), therefore the pressure induced by electromagnetic interactions is actually only relevant on very small-scales ($k \gtrsim 1h\mathrm{Mpc}^{-1}$). For this reason, a good and practical approximation is to consider all of the matter (baryons and CDM) as a single component and neglect all non-gravitational forces. Here in this section, we neglect the different initial conditions for baryons and cold dark matter which is something that we will discuss extensively later. In chapter 5, we introduce a technique to perform gravity-only simulations considering the fact that CDM and baryons are not comove in their initial conditions. But in this section, we regard baryons and CDM as a single matter fluid. Considering continuity, Euler, and Poisson equations in the Fourier space for the matter which govern the linear evolution:

$$\delta_m' + iku_m + 3\Phi' = 0,$$
$$u_m' + \frac{a'}{a}u_m + ik\Psi = 0,$$
$$k^2\Phi = 4\pi Ga^2\rho_m\delta_m. \qquad (3.54)$$

In each of the above equations $\prime$ stands for a derivative with respect to the conformal time η, and δ_m denotes the matter density perturbation which defines as $\delta_m \equiv \delta\rho_m/\bar{\rho}_m$, u_m represents the velocity of the matter and Ψ and Φ are the metric perturbations as appeared in Eq. 2.33 before. Note that on the right-hand side of the Poisson equation we only include matter. This is because of the fact that the redshifts in which the structures begin to become non-linear are about $z \lesssim 10$ and in these redshifts the contribution of radiation is negligible. Concerning the fact that the late Universe has negligible anisotropic stress, we can set $\Phi = -\Psi$.

38

One basic fact is that large-scale fluctuations are small while small-scale fluctuations are very large. Mostly this is because of the fact that small-scale fluctuations had more time to grow, and what happens is that in structure formation, small-scale structures with larger perturbations grow faster and ensemble to gravitationally bound structures. Then the process of forming bound structures gradually extends to larger and larger scales. Indeed when we go to very large scales there are no bound structures at all and we can just see small density perturbations on the background. This fact allows us to do simulations of large volumes because then we can assume the entire simulation volume is basically at the background level. To understand this better, let's characterize the large-scale matter density field. We can smooth the matter density field on different scales and compute the variance of that smooth density field. To do so, we define a filtered density field as follows:

$$\delta_W(\mathbf{x}) = \int d^3y\, W(|\mathbf{x} - \mathbf{y}|)\delta_m(\mathbf{y})\,. \tag{3.55}$$

This sort of filtering corresponds to: $\delta_W(\mathbf{k}) = W(k)\delta_m(\mathbf{k})$ in Fourier space. By computing the variance of this filtered density field, we find:

$$\sigma_W^2 \equiv \langle (\delta_W)^2(\mathbf{x}) \rangle = \int \frac{d^3k}{(2\pi)^3} \int \frac{d^3k'}{(2\pi)^3} \langle \delta_W(\mathbf{k})\delta_W^*(\mathbf{k}') \rangle e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} \tag{3.56}$$

$$= \int \frac{d^3k}{(2\pi)^3} P_L(k)|W(k)|^2 = \frac{1}{2\pi^2} \int d\ln k\, k^3 P_L(k)|W(k)|^2\,.$$

The result is shown in Fig. 3.8 which is saying that the standard deviation is small on large scales and it is large on small scales. This effect is on the basis of the hierarchical structure formation paradigm.
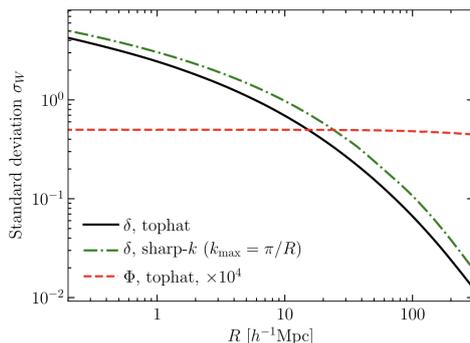
Figure 3.8: The standard deviation of the linear matter density field as a function of smoothing scale R at $z = 0$ using the real-space top-hat and sharp-k filters in black and green. The red dashed line shows the standard deviation of the gravitational potential multiply by 10^4. Credit: [5].

An important scale on Fig. 3.8, is when the standard deviation becomes in the order of ~ 1 which is roughly at $R \sim 10\,h^{-1}\mathrm{Mpc}$. This means that on length scales of $\sim 10\,h^{-1}\mathrm{Mpc}$, the density field has the fluctuations with the order of 1. So if we do any perturbation theory in small fluctuation, would not work on scales like $\lesssim 10\,h^{-1}\mathrm{Mpc}$ which means that our linear treatment based on Eq. 3.54 predicts a wrong result for the density field on small-scales. Another important point here is the fact that the gravitational potential, Φ, is always small (it does not matter in which scales we are looking) which means that the metric perturbations are small which means that the small perturbations to the background FLRW metric is a good approximation on all scales in structure formation so we do not have to run numerical relativity simulations. We can run quasi Newtonian simulations with small perturbations.

Now we would like to extend Eqs. 3.54 to the small scale in linear order. In order to do that we consider the Cartesian form of the Boltzmann Equations (Eq. 2.22).

$$\frac{df_m}{dt} = \frac{\partial f_m}{\partial t} + \frac{\partial f_m}{\partial x^i}\frac{dx^i}{dt} + \frac{\partial f_m}{\partial p^i}\frac{dp^i}{dt} = 0. \tag{3.57}$$

Thus, we need expressions for $\frac{dx^i}{dt}$ and $\frac{dp^i}{dt}$, and these will be derived from the relativistic context given by the geodesic equations (which tell us how particles position and momenta evolve). Here, where we are dealing

with small scales and with the non-relativistic matter, these equations come down back to the Newtonian equations except for the fact that we are dealing with comoving coordinates (we have the factors of a): $\dfrac{dx^i}{dt} = \dfrac{p^i}{am}$ and $\dfrac{dp^i}{dt} = -Hp^i - \dfrac{m}{a}\partial_i\Psi$. Then we need an equation for the gravitational potential Ψ, which comes from a Poisson equation: $\nabla^2\Psi = 3/2\Omega_m(\eta)(aH)^2\delta_m$, (zero-zero component of the Einstein equation in the subhorizon limit. here we replaced $4\pi G\rho_m$ with $3/2\Omega_m(\eta)H^2(\eta)$). Then we can have the fundamental set of equations governing the nonlinear growth of matter which is known as the collisionless, non-relativistic Boltzmann (Vlasov) equation coupled to gravity via the Poisson equation:

$$\frac{df_m}{dt} = \frac{\partial f_m}{\partial t} + \frac{\partial f_m}{\partial x^j}\frac{p^j}{ma} - \frac{\partial f_m}{\partial p^j}\left[Hp^j + \frac{m}{a}\frac{\partial\Psi}{\partial x^j}\right] = 0,$$

$$\nabla^2\Psi = \frac{3}{2}\Omega_m(\eta)(aH)^2\delta_m. \tag{3.58}$$

This coupled set of equations forms the starting point for the nonlinear evolution of matter (considered as a fluid with zero pressure). It is known as the Vlasov–Poisson system. A nonlinear system (through the coupling between Ψ and f_m) in $6+1$ dimensions, is difficult to solve. The following subsections will deal with perturbative as well as numerical techniques to make quantitative progress at a non-linear level.

3.1 Perturbation theory

The perturbative approach to the nonlinear growth of structure is to simplify the equations by taking moments of the Vlasov equation. To go from a $6+1$ dimension to the $3+1$ we can follow this approach. We can multiply the equation by p or p^2 and then integrate over d^3p to go from a $6+1$ phase-space dimension to the $3+1$ space+time. If we take the moments of the Vlasov equation 3.58 we obtain the following equations to solve:

$$\delta'_m + \theta_m = -\delta_m\theta_m - u_m^j\frac{\partial}{\partial x^j}\delta_m,$$

$$\theta'_m + aH\theta_m + \nabla^2\Psi = -u_m^j\frac{\partial}{\partial x^j}\theta_m - (\partial_i u_m^j)(\partial_j u_m^i),$$

$$\nabla^2\Psi = \frac{3}{2}\Omega_m(\eta)(aH)^2\delta_m. \tag{3.59}$$

The first equation above, that we can get from the zeroth moment of the Boltzmann equation, is nothing but the continuity equation where $\theta_m \equiv \partial_i u_m^i$ represents the velocity divergence. The second equation that we get by taking the first moment of the Boltzmann equation is the Euler equation. One thing that we can clearly see is that now these equations only contain the fields define only in space and time (no momentum). In the first and second equations above indeed we moved those terms that are nonlinear in the variables we wish to solve for, to the right-hand side.

Given the structure of this and given the fact that at least on sufficiently large-scales perturbations are small, we will take a perturbative approach. As the first step we will plug in our linear solution into these source terms (represented by (1) in the following equations) And these source terms drive the evolution of a second-order correction to the linear solution (denoted by (2))

$$\delta^{(2)\prime} + \theta^{(2)} = -\delta^{(1)}\theta^{(1)} - (u^{(1)})^j\frac{\partial}{\partial x^j}\delta^{(1)},$$

$$\theta^{(2)\prime} + aH\theta^{(2)} + \frac{3}{2}\Omega_m(\eta)(aH)^2\delta^{(2)} = -(u^{(1)})^j\frac{\partial}{\partial x^j}\theta^{(1)} - [\partial_i(u^{(1)})^j][\partial_j(u^{(1)})^i]. \tag{3.60}$$

This approach is justified as long as we know that for instance $\delta^{(2)}$ is smaller compare to $\delta^{(1)}$. The basic idea of the perturbation theory is to expand all the fields (fractional matter density perturbation and the velocity divergence of matter) in orders of perturbations as the following:

$$\delta_m(\mathbf{x}, \eta) = \delta^{(1)}(\mathbf{x}, \eta) + \delta^{(2)}(\mathbf{x}, \eta) + \ldots + \delta^{(n)}(\mathbf{x}, \eta),$$

$$\theta_m(\mathbf{x}, \eta) = \theta^{(1)}(\mathbf{x}, \eta) + \theta^{(2)}(\mathbf{x}, \eta) + \ldots + \theta^{(n)}(\mathbf{x}, \eta), \tag{3.61}$$

where for instance $\delta^{(1)}$ is the linear solution, $\delta^{(2)}$ would be the second order solution, and so on. We can keep doing that iteratively and obtain a solution at any given order. This approach is expected to work as long as each successive term in the series is smaller than the previous one. And at each order the solution can be written as the so called following form:

$$\delta^{(n)}(\mathbf{k}, \eta) = D_+^{(n)}(\eta) \left[\prod_{i=1}^{n} \int \frac{d^3 k_i}{(2\pi)^3} \right] (2\pi)^3 \delta_D^{(3)}\left(\mathbf{k} - \sum_{i=1}^{n} \mathbf{k}_i \right) \times F_n(\mathbf{k}_1, \ldots, \mathbf{k}_n) \delta_0(\mathbf{k}_1) \ldots \delta_0(\mathbf{k}_n),$$

$$\theta^{(n)}(\mathbf{k}, \eta) = ahf\hat{\theta}^{(n)} = -aHfD_+^{(n)}(\eta) \left[\prod_{i=1}^{n} \int \frac{d^3 k_i}{(2\pi)^3} \right] (2\pi)^3 \delta_D^{(3)}\left(\mathbf{k} - \sum_{i=1}^{n} \mathbf{k}_i \right) \times G_n(\mathbf{k}_1, \ldots, \mathbf{k}_n) \delta_0(\mathbf{k}_1) \ldots \delta_0(\mathbf{k}_n).$$

$$(3.62)$$

Where F_n and G_n are the kernels describing the coupling between different modes and the Dirac delta function ensures the momentum conservation. We only scratched the surface of the perturbation theory here, we refer to [99] for a comprehensive review.

So far, we have been working with the Eulerian fields in which the densities and velocities are defined at fixed spatial positions and fixed time. There exists an alternative approach to the perturbation theory which is called the "Lagrangian approach", where we follow the mass elements along their trajectories, labeling each trajectory with a coordinate q. In this picture, we are interested in the displacement field $S(\mathbf{q})$ that maps the initial positions, $\mathbf{q}$, to the final ones $\mathbf{x}$ by the following equation in which the displacement vanishes at the initial condition.

$$\mathbf{x}(\mathbf{q}, \eta) = \mathbf{q} + S(\mathbf{q}, \eta). \tag{3.63}$$

To follow this we use the geodesic equations: ($\frac{dx^i}{dt} = \frac{p^i}{am}$ and $\frac{dp^i}{dt} = -Hp^i - \frac{m}{a}\partial_i \Psi$) and we plug Eq. 3.63 in the geodesic equations. Then we have: $S''(\mathbf{q}, \eta) = aHs'(\mathbf{q}, \eta) + \nabla_x \Psi(\mathbf{x}, \eta) = 0$ for the displacement. The only complication is that we need a gravitational potential along the trajectory. To determine this potential we use the Poisson equation which relates the potential and the density. Using the mass conservation in Lagrangian and Eulerian pictures we can find the density as follows:

$$\bar{\rho}_m(\eta) d^3\mathbf{q} = \rho_m(\mathbf{x}, \eta) d^3\mathbf{x},$$

$$\frac{\rho_m}{\bar{\rho}_m} = 1 + \delta_m = \left| \frac{\partial \mathbf{q}}{\partial \mathbf{x}} \right| = \left| \frac{\partial \mathbf{x}}{\partial \mathbf{q}} \right|^{-1} = |\delta_{ij} + \partial_{q,i} S_j(\mathbf{q}, \eta)|. \tag{3.64}$$

Then we can insert δ_m into the Poisson equation to obtain Ψ. Perturbation theory then proceeds by writing

$$\mathbf{S} = \mathbf{S}^{(1)} + \mathbf{S}^{(2)} + \ldots \tag{3.65}$$

and solving the displacement field order by order.

3.2 Simulations

The idea behind the development of N-body simulations to numerically solve the Vlasov-Poisson system, (Eq. 3.58) is to discretize the phase-space volume occupied by matter (this volume takes the form of a thin sheet which is shown in Fig. 3.9 due to the cold nature of dark matter) in other words dividing this thin sheet into some patches and then following the evolution of the phase-space elements by integrating the geodesic equation. We call these patches of the phase-space sheet, or phase-space elements, "particles", but it is important to keep in mind that they are not actually particles. They stand for fluid elements of this sheet in the phase-space. Indeed each element could consist of a huge number ($\sim 10^{30}$ or 10^{50}) of the "real" dark matter particles. It actually does not matter how many particles make up this sheet, we are only interested in that there is such a sheet. The mass, m, of the particles (which we assume here is the same for all particles) is only a numerical
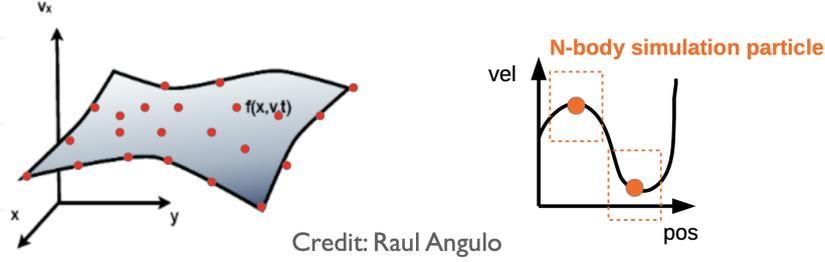
Figure 3.9: The initial conditions states that matter occupies a thin sheet in phase space, with a unique single-valued velocity $\mathbf{u}(\mathbf{x}, t)$ at each point in space. Credit: Raul Angulo.

parameter: it is determined by the total amount of matter in the simulation volume divided by the number of particles, so a higher-resolution simulation has more particles with correspondingly smaller m.

A small element of the sheet has a well-defined position x and momentum p. Since the motion of dark matter particles in this small region of phase space is described by the geodesic equation, so is that of the element of the phase-space sheet itself:

$$
\begin{aligned}
\frac{dx^i}{dt} &= \frac{p^i}{ma}, \\
\frac{dp^i}{dt} &= -Hp^i - \frac{m}{a}\frac{\partial \Psi}{\partial x^i}.
\end{aligned}
\tag{3.66}
$$

We can transform the geodesic equations above using the "superconformal" momentum $\mathbf{P}_c \equiv \mathbf{p}$ and then we have

$$
\begin{aligned}
\frac{dx^i}{dt} &= \frac{p_c^i}{ma^2}, \\
\frac{dp_c^i}{dt} &= -m\frac{\partial \Psi}{\partial x^i}.
\end{aligned}
\tag{3.67}
$$

The advantage of p_c is that it is conserved in the absence of perturbations, i.e. when the gradient of Ψ vanishes. And it is always good for code evolution purposes if you have a conserved quantity in the absence of perturbations. Of course, we need the potential again so we need to solve the Poisson equation as well.

Therefore we need to calculate the potential once and compute p_c^i and then compute the position x up to today. And that is nothing other than the first-order treatment of the Lagrangian perturbation theory (known as Zel'dovich approximation [100], we described the Lagrangian perturbation theory approach in Sec. 3.1).As illustrated in Fig. 3.10, we start with a uniform grid of particles (the left panel) and we displace each of them according to the potential gradient and move the particles to $z = 0$ with this potential and obtain the right-hand side panel of Fig. 3.10 in which we can see some structures. This technique is useful to generate initial conditions for the simulations.

There are two tasks we need to do for a N-body simulation. The first one is the integration of the particles trajectories and the other is the calculation of the potential of the forces. We start with

- **time integration:**

Typically the simulations evolve some version of what is called the leapfrog scheme where density and velocity are given at staggered epochs. So, we start with particle positions and velocities:

$$
\mathbf{x^{(i)}}(t) \quad \text{and} \quad \mathbf{p}^{(i)}(t - \Delta t/2).
\tag{3.68}
$$

Where Δt is the time step and superscript, i, denotes the index of the particle. The idea is that we have the position and the momentum ahead and next we can advance the position using that momentum and then

42
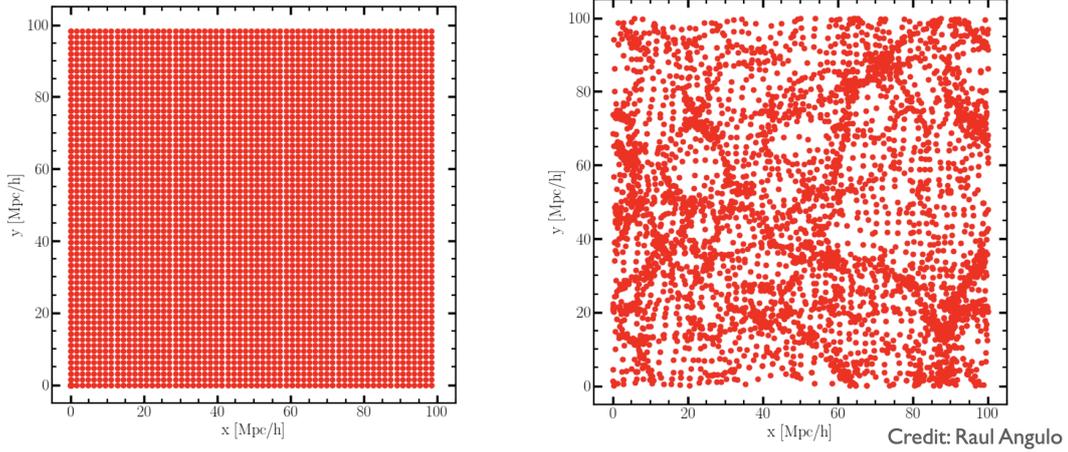
Figure 3.10: **Left panel:** A uniform grid of particles shown in $x - y$ plan to start the simulation with. **Right panel:** The particles displaced according to the potential gradient. This panel somehow looks like structures (we can see some particles accumulates and we have some empty regions connected by filaments). Credit: Raul Angulo.
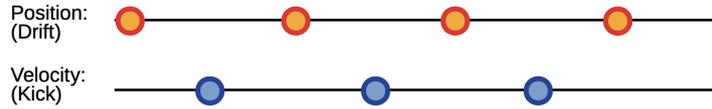
Figure 3.11: Positions and velocities (momentum) at staggered times. Credit: Raul Angulo

we use the position to calculate the force (the potential), afterwards we can move the momenta according to the force. This is a trick that is typically used in the N-body simulations to preserve energy to the cubic order in time.

The sequence is as follows:

1. Compute the gravitational potential generated by the collection of particles, and take its gradient to obtain $\nabla\Psi(\mathbf{x}, t)$

2. Change each particle's momentum ("kick") by

$$\mathbf{p}_c^{(i)}(t + \Delta t/2) = \mathbf{p}_c^{(i)}(t - \Delta t/2) - m\nabla\Psi(\mathbf{x}^{(i)}, t)\Delta t. \tag{3.69}$$

3. Move each particle position ("drift") by

$$\mathbf{x}^{(i)}(t + \Delta t) = \mathbf{x}^{(i)}(t) + \frac{\mathbf{p}_c^{(i)}(t + \Delta t/2)}{ma^2(t + \Delta t/2)}\Delta t. \tag{3.70}$$

4. Repeat

Note that particle positions and momenta are offset by half a time step. This scheme ensures that the energy of each particle is conserved to high accuracy (the numerical error in the energy is only of order $(\Delta t)^3$. The size of the time step Δt used for each particle is often adjusted to the local acceleration $|\nabla\Psi|$ to guarantee accuracy in high-density regions. N-body simulations are typically performed in a cubic volume with periodic boundary conditions so that particles exiting the volume on one side re-enter on the other side. This is appropriate if one aims to simulate a representative volume of a Universe that is statistically homogeneous.

For the initial condition of the N-body simulation, we need to start this chain of steps from somewhere. Indeed, that initial conditions consist of a set of particle's positions and momenta (because dynamically that is what we need to know to solve the future evolution) and we can compute those using the Zel'dovich approximation [100]. So there is no randomness involved, as long as we have a linear density field given. Linear density

field allows us to compute the linear potential, which allows us to compute the positions and momenta of the particles at linear order. So in this sense, there is no randomness involved but of course to generate the linear density field we typically sample from the linear power spectrum.

Now we turn to the second task:

- **solving for the force:**

A crucial step in N-body simulations is the calculation of the potential gradient, i.e. the gravitational forces. Here again, there exist numerical techniques and there is always a trade between accuracy and computation cost.

Particle-Particle: The simplest approach to compute the gravitational force acting on particles is called direct summation or Particle-Particle (PP)[101] in which we actually sum up the forces of all particles of any given particle:

$$\nabla\Psi(\mathbf{x}_i) = Gm_i \sum_{j\neq i} m_j \frac{\mathbf{x}_j - \mathbf{x}_i}{|\mathbf{x}_j - \mathbf{x}_i|^3}, \tag{3.71}$$

this is performed for each particle, j, summing over all the other $N-1$ particles, i. This method has a computational cost of the order of N^2, making it unaffordable for cosmological simulations, where the number of particles ranges from 10^6 to 10^{15}. Moreover, this method diverges when close encounters occur between two particles, so a "smoothing" of the force is required [102]. This is done by substituting the Newtonian force with a softened force $\nabla\Psi(\mathbf{x}_i)_{\text{soft}}$ that tends to $\nabla\Psi(\mathbf{x}_i)_{\text{Newt}}$, when $|\mathbf{x}_j - \mathbf{x}_i| \to \infty$ and tends to 0 when $|\mathbf{x}_j - \mathbf{x}_i| \to 0$. A simple example of such a smoothed force is:

$$\nabla\Psi(\mathbf{x}_i) = Gm_i \sum_{j\neq i} m_j \frac{\mathbf{x}_j - \mathbf{x}_i}{(|\mathbf{x}_j - \mathbf{x}_i|^2 - \epsilon^2)^{3/2}}, \tag{3.72}$$

where ϵ is called softening length and can be interpreted as the size of the particles. Unphysical two-body interactions can arise if the softening length is smaller than the mean inter-particle separation [103], leading to departures from the collisionless nature of the Vlasov-Poisson system.

Hierarchical Tree: The so-called Hierarchical Tree methods [104], alleviate the computational cost of the Poisson solver by using direct summation only for particles that are below a certain distance, while the force due to long-range interactions is approximated by expanding the gravitational field in multipoles and truncating the expansion to a given order. The computational cost of these algorithms is of the order of $N\log N$, In these methods, particles are assigned to a tree of computational cells, which is constructed by iterative subdivision of the volume of each cell in 8 parts (octree) halting the procedure when all the cells contain either 0 or 1 particles. Fig. 3.12.

Fig. 3.13, sketches the configuration for the multipole expansion. To compute the force exerted by a collection of particles on a particle at a given distance one can approximate the gravitational potential of the group

$$\phi(r) = -G \sum_i \frac{m_i}{|\mathbf{r} - \mathbf{x}_i|} = -G \sum_i \frac{m_i}{|\mathbf{r} - \mathbf{s} + \mathbf{s} - \mathbf{x}_i|}, \tag{3.73}$$

by Taylor expanding the denominator for $\mathbf{x}_i - \mathbf{s} \ll \mathbf{r} - \mathbf{s}$. The expansion gives [7]

$$\frac{1}{|\mathbf{y} + \mathbf{s} - \mathbf{x}_i|} = \underbrace{\frac{1}{|\mathbf{y}|}}_{\text{monopole}} - \underbrace{\frac{\mathbf{y}(\mathbf{s} - \mathbf{x}_i)}{|\mathbf{y}|^3}}_{\text{dipole}} + \frac{1}{2}\underbrace{\frac{\mathbf{y}^T[3(\mathbf{s} - \mathbf{x}_i)(\mathbf{s} - \mathbf{x}_i)^T - (\mathbf{s} - \mathbf{x}_i)^2]\mathbf{y}}{|\mathbf{y}|^5}}_{\text{quadrupole}} + \ldots \tag{3.74}$$

Where $\mathbf{y} = \mathbf{r} - \mathbf{s}$. These terms define the monopole, dipole and quadrupole of the group of particles:
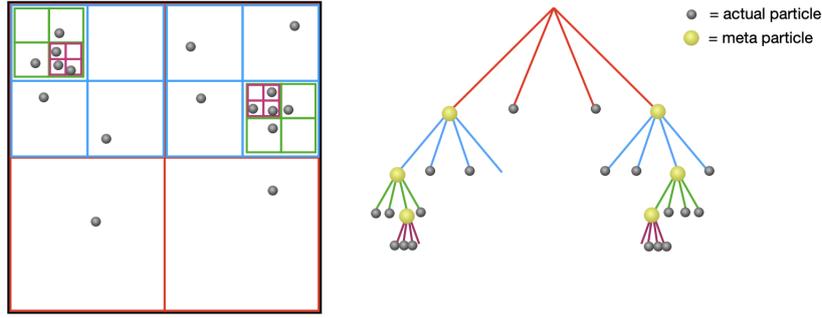
Figure 3.12: Instead of a direct summation, one organizes particles in a hierarchical tree with 'meta' particles located at the center of mass of the respective sub-tree at whose root they are located and carrying the total mass of the branch. Depending on the distance of the particle which wishes to compute the interaction, one can replace the entire tree branch with the meta-particle to good accuracy. For even better accuracy, one can carry out a multi-pole expansion of the branch so that quadrupoles, etc. can also be carried along. The effective algorithm ends up being $N \log N$. Credit:Oliver Hahn

$$
M = \sum_i m_i,
$$
$$
D_i = \sum_i m_i(\mathbf{s} - \mathbf{x}_i) = 0,
$$
$$
Q_{ij} = \sum_j m_j[3(\mathbf{s} - \mathbf{x}_i)(\mathbf{s} - \mathbf{x}_i)^T - (\mathbf{s} - \mathbf{x}_i)^2]. \tag{3.75}
$$

The dipole vanishes because we carried out the expansion around the center of mass:

$$
\mathbf{s} = \frac{1}{M} \sum_i m_i \mathbf{x}_i . \tag{3.76}
$$
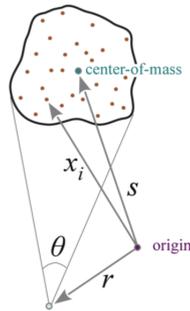
Figure 3.13: Multipole expansion for a group of distant particles. Provided the reference point $\mathbf{r}$ is sufficiently far away, the particles are seen under a small opening angle θ and the field created by the particle group can be approximated by the monopole term at its center of mass, augmented with higher-order multipole corrections if desired. Credit:[7]

If we stop the expansion at quadrupole order the potential is then given by:

$$\phi(\mathbf{r}) = -G\left(\frac{M}{|\mathbf{y}|} + \frac{1}{2}\frac{\mathbf{y}^T Q \mathbf{y}}{|\mathbf{y}|^5}\right). \tag{3.77}$$

from which the force can be obtained by differentiation. The opening angle θ is adjusted so that smaller distances correspond to larger opening angles. This procedure assures high resolution of the force in high-density regions but is inefficient for low contrast mass distribution.

Particle-Mesh: Particle-mesh (PM) methods [105], take a different approach to solve the Poisson equation. In these methods, a grid is superimposed to the particle distribution and physical quantities like density and velocity are interpolated from particle positions to grid nodes. The Particle Mesh algorithm assumes that particles have a certain size, mass, shape, and internal density. This determines the interpolation scheme used to assign densities to grid cells. Suppose we have N particles with mass m_i and coordinates $\mathbf{r_i}(i = 1, 2, \ldots, N)$ onto a mesh with uniform spacing $h = L/N_g$ and $\{\mathbf{r_p}\}$ denote the set of discrete cell centers with $\mathbf{p} = (p_x, p_y, p_z)$, being an integer index $0 \leq p_{x,y,z} < N_g$. We can define a one dimensional particle shape by $S(x)$ normalised by $\int S(x)\,dx = 1$. Then we assign a function $W_p(\mathbf{x_i})$ of particle i's mass that falls in to the cell which has the index $\mathbf{p}$.

$$\mathbf{W_p(x_i)} = \int_{\mathbf{x_p}-h/2}^{\mathbf{x_p}+h/2} \mathbf{S(x_i - x_p)dx}. \tag{3.78}$$

By introducing a top-hat function as follows we have: $\Pi(x) = \dfrac{1}{\Delta x}\begin{cases} 1, & \text{if } |\mathbf{x}| \leq 1/2, \\ 0, & \text{otherwise}, \end{cases}$

Then the integration boundaries would extend to all space and we would have

$$\mathbf{W_p(x_i)} = \int \Pi\left(\frac{\mathbf{x - x_p}}{h}\right)\mathbf{S(x_i - x_p)dx}. \tag{3.79}$$

Then the full density in grid cell $\mathbf{p}$ given by

$$\rho_{\mathbf{p}} = \frac{1}{h^3}\sum_{i=1}^{N} m_i\,\mathbf{W_p(x_i)}. \tag{3.80}$$

The common choices for the shape function in the literature are:
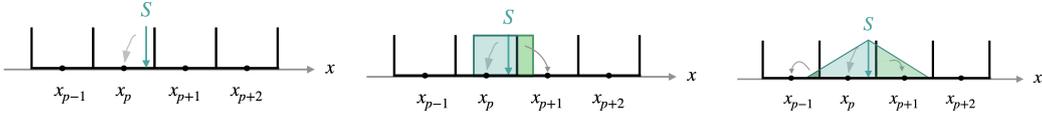
Figure 3.14: A 1D sketch of the geometric interpretation of NGP (**left panel**), CIC (**middle panel**) and TSC (**right panel**) assignments.

Nearest Grid Point (NGP): Particles are point-like and all the particle's mass is assigned to the single grid that contains it and the function $\mathbf{W_p(x_i)}$, would be (see Fig. 3.14, left panel)

$$\mathbf{W_p(x_i)} = \int \Pi\left(\frac{\mathbf{x - x_p}}{h}\right)\delta(\mathbf{x_i - x_p})\mathbf{dx} = \Pi\left(\frac{\mathbf{x - x_p}}{h}\right). \tag{3.81}$$

Cloud In Cell (CIC): Particles have a cubical could shape in 3 dimensions of uniform density and one grid cell size. And the assignment function defines as (see Fig. 3.14, middle panel)

$$\mathbf{W_p(x_i)} = \int \Pi\left(\frac{\mathbf{x} - \mathbf{x_p}}{h}\right) \frac{1}{h^3}\left(\frac{\mathbf{x} - \mathbf{x_p}}{h}\right) dx. \tag{3.82}$$

Triangular Shaped Cloud (TSC): which is ia systematic sequence of ever higher-order shape functions by adding more convolutions with the top-hat kernel. (see Fig. 3.14, right panel)

$$\mathbf{W_p(x_i)} = \int \Pi\left(\frac{\mathbf{x} - \mathbf{x_p}}{h}\right) \frac{1}{h^3}\left(\frac{\mathbf{x_i} - \mathbf{x} - \mathbf{x}'}{h}\right) \frac{1}{h^3}\Pi\left(\frac{x'}{h}\right) \mathbf{dxdx}' \tag{3.83}$$

$$= \frac{1}{h^6}\int \Pi\left(\frac{\mathbf{x} - \mathbf{x_p}}{h}\right)\Pi\left(\frac{\mathbf{x_i} - \mathbf{x}}{h}\right)\Pi\left(\frac{\mathbf{x}' - \mathbf{x}}{h}\right)\mathbf{dxdx}'. \tag{3.84}$$

Imposing periodic boundary conditions allows one to use Fast Fourier Transform (FFT) algorithms to compute the potential using the Poisson equation.

Hybrid approaches and refinement: Modern cosmological simulation codes use hybrid approaches that combine the methods described above for different regimes or introduce refinement of the grid to have a better resolution inside high density peaks. For instance TreePM hybrid methods use the Tree algorithm for small-range interactions and exploit the speed of FFT to compute long-range gravitational forces with the PM algorithm. Adaptive Mesh Refinement (AMR) methods allow to reach higher accuracy in high density regions with respect to PM methods by introducing sub-grids with finer spacing where the density in the grid cells exceeds a given threshold.

3.3 Dark matter halos

To identify bound structures in N-body simulations one searches for the very densest points in the matter distribution (matter never gets anymore dense from cosmology point of view than halos.), and then evaluates which particles nearby are gravitationally bound objects which are a well defined objects as follows. Now the question is that how can we evaluate the "gravitationally bound" particles in the simulation? Since we can compute the gravitational potential of the halo Ψ_h at a given position and compare it with the particle's kinetic energy. Indeed a particle with velocity v is bound if $v^2/2 < |\Psi_h|$. The other question is that how particles can be define as "nearby" particles? Particles are defined as "nearby" either they are within a spherical region whose interior density is above some threshold ("spherical overdensity" algorithm [106]), or if their nearest-neighbor distance to other halo particles is below a threshold value. ("friends-of-friends" algorithm [107]).

We believe all galaxies to be hosted by dark matter halo of some mass. For this reason we have actually some strong observational evidences, If we look at the dynamics within galaxies or of the galaxies in galaxy groups, or indeed if we look at gravitational lensing around galaxies, then we find that there is much more matter than the visible baryons and this additional matter is much more extended than the baryons. And that is where the term halo comes from. Because we can think of small galaxies sitting inside these much bigger structures of dark matter. So as we go through the most massive halos that exist in the Universe, those are actually associated with a whole assembly of galaxies (galaxy clusters) and that is why halos are highly relevant to observations but it is not easy to link them because we do not see the dark matter and still halos are mostly studied as objects in simulations. One of the issues we have is to define the halo mass. Maybe we think of that why dont we just count all the number of bound particles? The problem is that this is effected by numerical noise, it means if we change the simulation, this amount will change. The other problem is the fact that it is not practically relevant in the sense that we do not actually observe dark matter.

One of the most important tasks in the post-processing of cosmological simulations is the identification of gravitationally bound structures: the halos[1]. The existence of a local minimum of the gravitational potential is a necessary condition for the existence of a bound structure. A local minimum of the potential however corresponds also to local maximum of the matter density field. Thus, most halo finders focus on determining the region of gravitational boundness around a density maximum. Many methods exist, below the most commonly used for isolated halos are explained, and we discuss a few existing software packages after. Calculations using the model spherical collapse of a single top hat perturbation in Einstein-de Sitter (EdS) cosmology indicate that regions that are in virial equilibrium should have an overdensity ρ_{vir} of

---

[1]Credit: Oliver Hahn's lecture notes in 2018.

$$\rho_{\mathrm{vir}} = \Delta_{\mathrm{vir}}\rho_c \quad \text{with} \quad \Delta_{\mathrm{vir}} = 18\pi^2 \simeq 178, \tag{3.85}$$

where ρ_c is the critical density of the Universe (introduced in chapter 2, Sec. 1).

The Friends-of-Friends (FoF) Algorithm

Refs [108] and [109], introduced the so-called friends-of-friends algorithm that is particularly suited for N-body simulations. The algorithm builds equivalence classes among all particle pairs in the N-body simulation whose distance falls below a given threshold. The resulting equivalence classes thus comprise regions bounded by an isodensity surface. It is common practice to relate the distance threshold to the mean inter-particle distance b and adopt a value of $b = 0.2$. This choice can be roughly motivated by considering the density of an isothermal sphere at the virial radius. Expressed in units of the mean inter-particle separation, $b \sim (\Delta_{\mathrm{vir}}/3)^{-1/3} \approx 0.25$ in this case. One main problem of the friends-of-friends algorithm is that it tends to spuriously connect nearby halos that are linked by a bridge which may purely be a result of the stochastic nature of N-body simulations. This problem is most severe in regions of high environmental density. However, friends-of-friends haloes are not constrained in their shape by the algorithm and mass functions obtained with this algorithm show a nearly universal behaviour across redshift.

The Spherical Overdensity (SO) Algorithm

The second most commonly employed halo finding method is the spherical overdensity algorithm introduced by [110]. Here, particles are ranked according to their local density. Then a sphere is grown around the first particle until the mean density within that sphere falls below the virial overdensity, i.e. the virial radius R_{vir} is defined as $M_{\mathrm{vir}} = M(< R_{\mathrm{vir}})$ (where $M(< r)$ is the mass enclosed in a sphere of radius r) so that

$$\frac{(M(< R_{\mathrm{vir}}))}{\frac{4\pi}{3}R_{\mathrm{vir}}^3} = \Delta_{\mathrm{vir}}\rho_c. \tag{3.86}$$

All particles within the sphere make up the first halo and are removed from the list of density ranked particles. The procedure is then repeated with the next highest density particle left in the list and stopped, when no more groups are found that contain a given minimum number of particles. The resulting halos do not suffer from spurious linking of distinct halos. They are however spherical by definition which might not reflect the true shape of the gravitationally bound volume. A further advantage of the spherical overdensity method is that it can be readily used also in hydrodynamic simulations where particles and mesh cells come with basically arbitrary masses so that the simple friends-of-friends algorithm can- not be applied with a fixed distance threshold. Instead of ranking particles by their density, any method to identify peaks of the density field is sufficient to serve as the starting point before growing spheres around the peaks. For simplicity, often 200 is used instead of 178 and halos that are defined using this overdensity criterion are often denoted with an index $200c$, so that the usual mass definition for a halo is M_{200c}. In the sub-field of cosmology that uses galaxy clusters as cosmological probes, usually a higher density threshold of $500c$ is used that better correlates with the hot X-ray emitting region of the cluster halo. Sometimes also the mean density is used instead of the critical density so that $200m$ refers to $\rho_{\mathrm{vir}} = 200\rho_m = 200\,\Omega_m\rho_c$. In the presence of a cosmological constant, in principle the overdensity criterion becomes time dependent and a convenient parameterisation has been given by [111].

$$\Delta_{\mathrm{vir}}(a) = 18\pi^2 + 82x - 39x^2, \quad x := \frac{\Omega_m a^{-3}}{\Omega_m a^{-3} + \Omega_\Lambda} - 1. \tag{3.87}$$

There are several codes that are specifically designed to find halos in the simulations "halo finders". (e.g.[106, 107, 112, 110, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135]) Here in these works (chapters 5 and 6) we used Amiga Halo Finder (AHF) [119, 129], which identifies halos as spherical overdensities (SO) in the spatial distribution of particles in the simulations.

3.4 Cosmic Voids

Other than bound structures in the N-body simulations, cosmic voids are a novel probe to shed light on our Universe. Cosmic voids are defined as large underdense regions of the cosmic web, they are the largest structures in the Universe and make up most of its volume [136, 137]. Historically, their existence was one of the earliest predictions of the concordance cosmological model [138], and their observational detection goes back to roughly 40 years ago [139, 140]. Voids are in particular extremely underdense near their centers, and their spherically averaged density profile shows a characteristic shape [141, 142, 143, 144, 145]. Recently, cosmic voids are becoming one promising cosmological probes: in one side they could represent a population of statistically ideal spheres with a homogeneous distribution at different redshifts which size evolution could be used to probe the expansion of the Universe using Alcock & Paczynski test [146, 147, 148, 149, 150, 151, 152, 153]. Moreover, due to their low density, voids are naturally sensitive to dark energy and thus the interest to use them as probe of alternative Dark Energy models and modified gravity scenarios is increasing [154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166], as well as the possibility of using them to put constraints on neutrinos masses [167, 168, 169]. Finally, their imprint on the observed Cosmic Microwave Background (CMB) is also becoming an encouraging new cosmological probe, either through their ISW imprint as well as acoustic oscillations ([170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180]) or their lensing imprint ([181, 182, 183]). Furthermore, the observed cold spot of the CMB could be explained as the imprint of the ISW sourced by very large voids along the line of sight [184, 185, 186, 187]; some studies try to link high redshift intergalactic voids in the transmitted Lyman-α flux to the gas density [188]; Because of the fact that they are almost empty regions, their evolution during cosmic history is at most weakly nonlinear and their properties could possibly be impacted by the primordial density field from which they formed. This fact motivates us to investigate the effects of baryon-CDM relative perturbations on these regions and their statistics which we bring our results in chapter 6.

One of the 3D and widely used algorithms to identify voids in data catalogues and also N-body simulations is the well-known ZOBOV (ZOnes Bordering On Voidness) algorithm [9] which we used it to identify voids in our simulations in chapter 6. The ZOBOV algorithm performs a Voronoi tessellation of a set of points, identifies depressions in the density distribution of these points, and merges them in to the group of the Voronoi cells using a watershed transform [189] without pre-determined assumptions about void shapes, sizes or mean underdensities, which is the most appealing aspect of the watershed method. Here we briefly outline the basic steps of the void-finding technique in ZOBOV and we refer the interested readers to the main ZOBOV paper [9] for a detailed description. One can describe the ZOBOV mechanism with the four following main steps, that are summarized in Fig. 3.15 as well.

1. *Voronoi Tessellation Field Estimator ([190])*: The algorithm divides the space into cells around each particle (tracer) i in which the region inside the cell is closer to particle i than to any other one. It then estimates the density of each Voronoi region using the volume of each cell $1/V(i)$.

2. *Definition of the minimum density*: After estimating the density in each cell in the first step, the algorithm finds the minimum density cells, defined as Voronoi cells with a density lower than all their neighboring ones.

3. *Formation of basins*: The algorithm then joins adjacent higher-density cells to the minimum-density cell until no neighbor cell with a higher density can be found. It means that the void finder links all the particles to their minimum density neighbour. This procedure defines basins as the zones of these cells. At this point these basins themselves could be considered as voids because they are depression regions in the density field, but one single basin may also arise from spurious Poisson fluctuations due to the discreteness of the particles.

4. *Watershed transform*: The last step is when these basins are joined together using a watershed algorithm ([189]). For each basin b, the "water" level is set to the minimum density of b. It is then slowly elevated so that the it can flow to the neighbor basins, joining all of them to basin b. The process stops when the "water" flows into a basin with a lower minimum, which defines the final void distribution.

In chapter 6, we use the publicly available REVOLVER (REal-space VOid Locations from surVEy Reconstruction)[2] void finder to build our void catalogues with the ZOBOV algorithm [9], explained above, which is a 3D void finder and has been widely used in simulated and observed catalogues [191, 192, 169, 192]. In chapter 6, we will explain more the void statistics aw well as the main purposes of using them in our studies.

---

[2]https://github.com/seshnadathur/Revolver
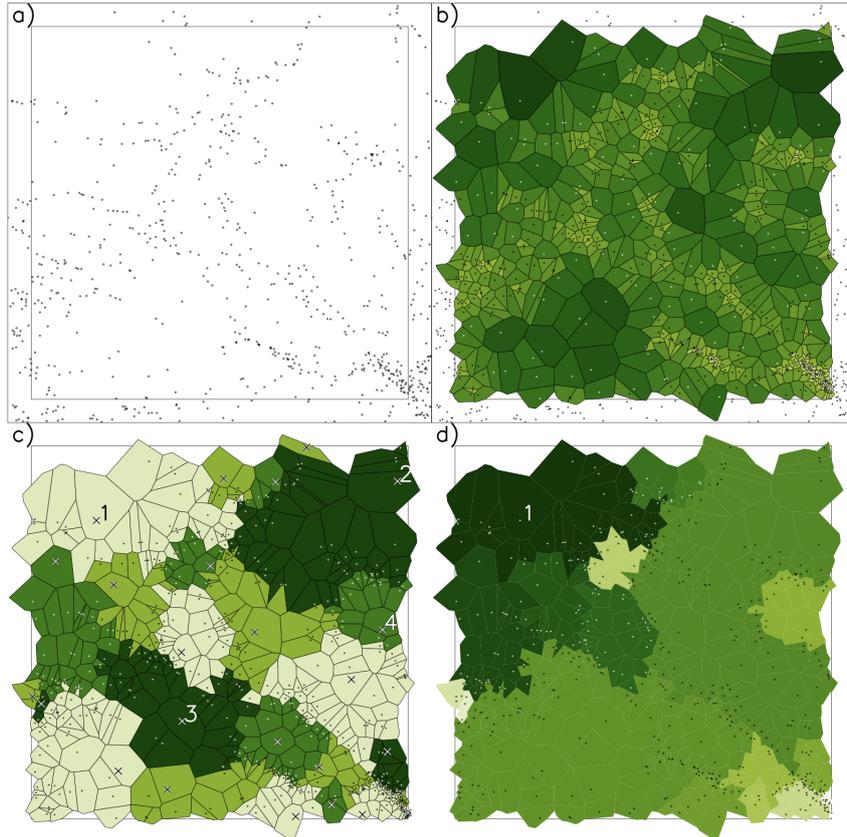
Figure 3.15: **Panel a:** Galaxies by [8], from a $40 \times 40 \times 5\,(\mathrm{h}^{-1}\,\mathrm{Mpc})^3$ slice. **Panel b:** The 2D Voronoi tessellation of galaxies in this slice, with each particle's Voronoi cell shaded according to its area. The galaxies outside the inner $(40\,\mathrm{h}^{-1}\,\mathrm{Mpc})$ boundary are shown because they contribute to the tessellation. **Panel c:** Basins. The cores (density minima) of each basin are shown with crosses, the different colors merely demarcate different basins. **Panel d:** Watershed transform, the growth of void 1, the deepest void in the sample. With analogy to a water tank, the water level (density) is increased and basins the water runs into are added to the void. Colours from dark to light indicate the stage at which the basin is added to the void. Credit: [9].

<div align="right">**4**</div>

Constraints on the Spacetime Dynamics of an Early Dark Energy Component

In the analysis we are about to present in this chapter, we study an Early Dark Energy cosmological model, and perform an analysis which takes into account both background and perturbation effects via the parameters c_{eff}^2 and c_{vis}^2, representing effective sound speed and viscosity, respectively. By using the latest available data at the time, we derive constraints on the amount of dark energy at early times and the present value of the equation of state. Our focus is on the effect that early dark energy has on the CMB data, including polarization and lensing, in a generalized parameter space including a varying total neutrino mass, and tensor to scalar ratio, besides the 6 standard parameters of the minimal cosmological model. (cosmological parameters are represented in details in chapter 3, Sec. 1) We find that the inclusion of Baryonic Acoustic Oscillations data and CMB lensing significantly improves the constraints on the EDE parameters, while other high redshift data like the Quasar Hubble diagram and the Lyman-α forest BAO have instead a negligible impact.

We find $\Omega_{\text{eDE}} < 0.0039$ and $w_0 < -0.95$ at the 95% C.L. for EDE accounting for its clustering through the inclusion of perturbation dynamics. This limit becomes stronger $\Omega_{\text{eDE}} < 0.0034$ if perturbations are neglected. The constraints on the EDE parameters are remarkably stable even when Σm_ν, and r parameters are varied, with weak degeneracies between Ω_{eDE} and r or Σm_ν. In general we expect smaller values for the upper limits on the total amount of EDE with an increasing neutrino mass, while with a decreasing value of the tensor to scalar ratio we expect the 2σ upper limits on EDE to increase. We compare this EDE model with a simple wCDM with zero dark energy at early times and we find $\sim 1-2\%$ different upper limits on total neutrino mass and $\sim 0.1-0.2\%$ difference on the equation of state at the present time.

Perturbation parameters are not constrained with current data sets, and tensions between the CMB derived H_0 and σ_8 values and those measured with local probes are not eased. This work demonstrates the capability of CMB probes to constrain the total amount of EDE well below the percent level.

We will proceed as follows. In Sec. 1 we describe the EDE model and the behaviour of the perturbations. In Sec. 2, we explain the data sets used in this analysis. Sec. 3 is devoted to a discussion on the effects of the perturbations on the CMB data with and without EDE. Finally in Sec. 4, we present the method and our results as well.

1 Early Dark Energy(EDE) models

In this Section we define the EDE framework in terms of background evolution and perturbation behavior. In subsection 1.1, we bring the background evolution picture of the EDE model and in subsection 1.2, we explain the perturbation equations.

1.1 Background evolution

An approach to construct Dark Energy (DE) models consists in modifying the energy momentum tensor $T_{\mu\nu}$ in the right hand side of the Einstein equations (described in chapter 2, Sec. 3.1) accounting for a generalized component with negative equation of state. In this way, a Quintessence scalar field ϕ with a potential $V(\phi)$ may describe a late time cosmic acceleration [193, 194] (we brought a brief explanation of this scalar field in the context of inflation in chapter 2, Sec. 4).

Unlike the CC scenario, (see chapter 2, Sec. 1 for a brief explanation), the equation of state of Quintessence models dynamically varies with time.

EDE represents the class of models in which the DE contribution to the energy density is relevant already in the early Universe, and it can have an impact both on background evolution of geometrical quantities and on structure formation. The notion of EDE has been introduced by Wetterich (2004) [55] and subsequently studied in several works by considering different possible effective parametrizations of physical properties of the DE. Here, we concentrate on the general parametrization by Doran and Robbers (2006) [56] However, notice that the number of parameters can in principle be reduced as shown in [195].

In the latter approach, instead of parametrizing $w(a)$, the fractional DE energy density, $\Omega_{\rm DE}(a)$ is written as

$$\Omega_{\rm DE}(a) = \frac{\Omega_{\rm DE}^0 - \Omega_{\rm eDE}(1 - a^{-3w_0})}{\Omega_{\rm DE}^0 + \Omega_{\rm m}^0 a^{3w_0}} + \Omega_{\rm eDE}(1 - a^{-3w_0}). \tag{4.1}$$

Here $\Omega_{\rm DE}^0$ and $\Omega_{\rm m}^0$ are the fractional energy densities of dark energy and matter today, i.e. when the scale factor is normalized to $a = 1$; w_0 is the equation of state of the dark energy fluid today and we also assume a flat Universe (explained in chapter 2, Sec. 1). Notice that $\Omega_{\rm eDE}$ becomes constant at high redshifts. The evolution of $\Omega_{\rm eDE}$ is connected to the equation of state w by the following relation:

$$\left\{ 3w - \frac{a_{\rm eq}}{a + a_{\rm eq}} \right\} \Omega_{\rm DE}(1 - \Omega_{\rm DE}) = -d\Omega_{\rm DE}/d\ln a. \tag{4.2}$$

Therefore the evolution of $w(a)$ reads

$$w(a) = -\frac{1}{3[1 - \Omega_{\rm DE}(a)]} \frac{dln\Omega_{\rm DE}(a)}{d\ln a} + \frac{a_{\rm eq}}{3(a + a_{\rm eq})}, \tag{4.3}$$

where $a_{\rm eq}$ is the scale factor at matter-radiation equality, ($a_{\rm eq}$, shown as a vertical line in chapter 2, Fig. 2.1) and also today's equation of state would be written as $w(a = 1) = w_0$. In order to track the dominant cosmological component, $w(a)$ behaves differently in three different epochs (see chapter 2, Fig. 2.1 for the behaviour of energy density in different epochs): during the radiation dominated era, one has $w \sim 1/3$, while during the matter domination epoch, $w \sim 0$; finally at present, $w \sim w_0$ (in chapter 2, Sec. 2.2, we argued these different behaviour of the equation of state).
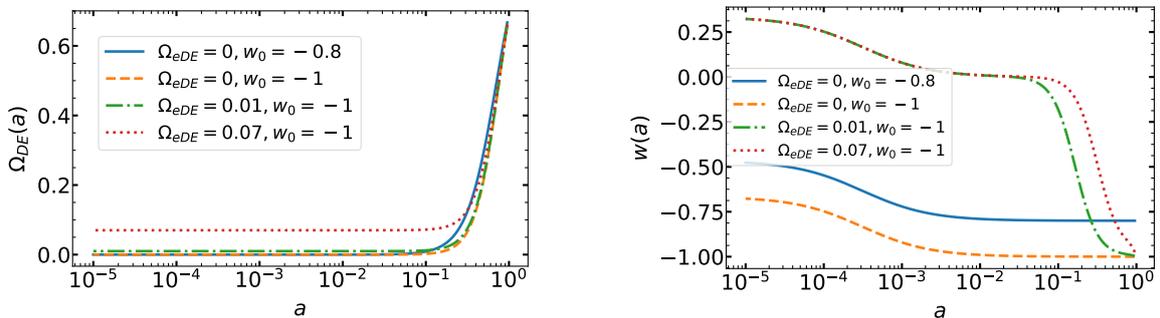
Figure 4.1: The behaviour of the EDE model as a function of scale factor. The left panel shows the evolution of the fractional DE density, while the right panel represents the evolution of the equation of state as a function of the scale factor for different values of $\Omega_{\rm eDE}$ and w_0.

In Figure 4.1 we plot $\Omega_{\rm DE}(a)$ and $w(a)$, for different values of $\Omega_{\rm eDE}$ and w_0.

1.2 EDE perturbations

Besides the background, the additional features that we define and discuss now, make EDE able to influence the behavior of cosmological perturbations. Therefore, DE density perturbations might leave an imprint in cosmological observables.

In the following we review the parametrization of the clustering feature of the EDE model we used in this work. Considering the fact that the gravitational influence of the DE component is controlled by the stress-energy tensor, $T_{\mu\nu}(\mathbf{x}, \eta)$, (defined in chapter 2, Sec. 2.2), which in general is a symmetric 4-tensor with 10 components that can be classified in 4 categories: the energy density ρ (1 component), the isotropic stress or pressure, p (1 component), the momentum density, $(\rho + p)\theta^i$ (3 components), and the anisotropic stress, $p\sigma^{ij}$ (5 components). The 5 components of the anisotropic stress can be further separated by their transformation properties under rotation into 1 scalar component, 2 vector and 2 tensor components. Here we neglect the tensor and vector components and we deal with the scalar component [196, 197]. One can choose these to be pressure, p and scalar anisotropic stress amplitude without loss of generality σ [198]. The non-adiabatic stress or entropy contribution is define as

$$p\Gamma = \delta p - c_s^2 \delta\rho, \tag{4.4}$$

where the adiabatic sound speed is defined as

$$c_s^2 = \frac{\dot{p}}{\dot{\rho}} = w - \frac{1}{3}\frac{\dot{w}}{1+w}\left(\frac{\dot{a}}{a}\right)^{-1}. \tag{4.5}$$

So, $p = w\rho$, which we have defined in chapter 2, Sec. 2.2, does not imply $\delta p = w\delta\rho$ due to temporal or spatial variations in w. The energy-momentum conservation yields the continuity equation for the density fluctuations, $\delta \equiv \delta\rho/\rho$ as follows

$$\left(\frac{\delta}{1+w}\right)^{\cdot} = -(k\theta + 3\dot{h}_\delta) - 3\frac{\dot{a}}{a}\frac{w}{1+w}\Gamma, \tag{4.6}$$

and the Euler equation

$$\dot{\theta} = -\frac{\dot{a}}{a}(1 - 3c^2)\theta + \frac{c^2}{1+w}k\delta + \frac{w}{1+w}k\Gamma - \frac{2}{3}\frac{w}{1+w}(1 - 3K/k^2)k\sigma + kh_\nu, \tag{4.7}$$

where $K = -H_0(1 - \Omega_{total})$ is the background curvature and the metric sources h_δ and h_ν depend on the choice of the gauge are defined as

$$h_\delta = \begin{cases} h_L, & \text{Synchronous,} \\ \Phi, & \text{Newtonian,} \end{cases} \tag{4.8}$$

$$h_\nu = \begin{cases} 0, & \text{Synchronous,} \\ \Psi, & \text{Newtonian,} \end{cases} \tag{4.9}$$

Notice that $h_L = h/6$ in the notation of [199]. Following the stress model introduced in [196], which considered the non-adiabatic pressure action to stabilize the perturbation in $w < 0$ regime by defining the following in the rest frame of DE where $T_i^0 = 0$.

$$w\Gamma = (c_{\text{eff}}^2 - c_s^2)\delta^{(\text{rest})}. \tag{4.10}$$

Here the assumption is that $c_{\rm eff}^2$ is only a function of time and if $c_{\rm eff}^2 > 0$, pressure support is obtained. The gauge transformation into an arbitrary frame gives

$$\delta^{(\rm rest)} = \delta + 3\frac{\dot{a}}{a}(1+w)(\theta - B)/k, \tag{4.11}$$

showing the gauge-invariant for the non-adiabatic stress [198, 200] and B illustrates the time-space component of metric fluctuations and vanishes in both the Synchronous and Newtonian gauges. Then the Euler equation could be written like

$$\dot{\theta} = -\frac{\dot{a}}{a}\theta - \frac{c_{\rm eff}^2}{1+w}k\delta^{(\rm rest)} - \frac{2}{3}\frac{w}{1+w}k(1-3K/k^2)\sigma + kh_\nu. \tag{4.12}$$

Consequently $c_{\rm eff}^2$ may be thought of as a rest-frame sound speed. The anisotropic stress could also affect the density perturbations which in the context of fluids it represents as viscosity and damps density perturbations. Generally the anisotropic stress component is the amplitude of the 3-tensor that is linear in the perturbation. The natural choice for the source of this is $k\theta$, the amplitude of the velocity shear tensor $\partial^i\theta^i$. However it should be gauge-invariant and generated by the corresponding shear term in the metric fluctuation H_T. The relation between velocity/metric shear and anisotropic stress would be parametrized with "viscosity parameter", $c_{\rm vis}$.

$$w\left(\dot{\sigma} + 3\frac{\dot{a}}{a}\sigma\right) = 4c_{\rm vis}^2(k\theta - \dot{H}_T), \tag{4.13}$$

where in the Newtonian gauge $H_T = 0$ and in the synchronous gauge $H_T = h_T = -h/2 - 3\eta$ [199].

As we mentioned above, the clustering features of different types of DE models are typically parametrized by an effective sound speed, that can be defined as the ratio of pressure perturbations to the density perturbations in the rest frame of the DE fluid, $c_{\rm eff}^2 \equiv \delta p/\delta\rho$ [196, 201]. In addition, another effective component in the density perturbation of an inhomogeneous DE model is the anisotropic stress which would be considerable for example if the DE behaves like a relativistic fluid with relevant viscosity effects. In order to parametrize the viscosity, we used $c_{\rm vis}^2$ as the viscous sound speed [196]. As mentioned in Refs. [199, 202, 203], by adopting the synchronous gauge in which the perturbation in the metric tensor is confined to the spatial sub-space, and by using the conservation of energy-momentum tensor $T_{\mu\nu}$ in Fourier space (indeed by using Eqs. 4.6, 4.11, 4.12 and 4.13), we can have the following relations for density perturbation, velocity perturbation and anisotropic stress:

$$\frac{\dot{\delta}}{1+w} = -\left[k^2 + 9\left(\frac{\dot{a}}{a}\right)\left(c_{\rm eff}^2 - w + \frac{\dot{w}}{3(1+w)(\dot{a}/a)}\right)\right]\frac{\theta}{k^2} - \frac{\dot{h}}{2} - 3\frac{\dot{a}}{a}(c_{\rm eff}^2 - w)\frac{\delta}{1+w}, \tag{4.14}$$

$$\dot{\theta} = -\frac{\dot{a}}{a}(1 - 3c_{\rm eff}^2)\theta + \frac{\delta}{1+w}c_{\rm eff}^2 k^2 - k^2\sigma, \tag{4.15}$$

$$\dot{\sigma} = -3\frac{\dot{a}}{a}\left[1 - \frac{\dot{w}}{3w(1+w)(\dot{a}/a)}\right]\sigma + \frac{8c_{\rm vis}^2}{3(1+w)}\left[\theta + \frac{\dot{h}}{2} + 3\dot{\eta}\right]; \tag{4.16}$$

δ, θ and σ represent the DE density perturbation, velocity perturbation and anisotropic stress, respectively, while h and η are the scalar perturbations of the space-space part in the metric, in the synchronous gauge, and "." denotes the derivative with respect to conformal time as we mentioned them all above. As we will see in the following, the equations above effectively control the DE clustering properties.

2 Data sets

In this section we present the CMB (in Sec. 2.1) and LSS (in Sec. 2.2) data sets we exploited in this work.

2.1 CMB

The analysis of Planck data follows dedicated pipelines for the so called low-ℓ and high-ℓ, corresponding to angular scales larger or smaller than a few degrees, respectively. In this work we utilise both temperature and polarization data sets from Planck. The details of the analyses and data sets are contained in the original publications by Planck [204]. We describe here their main features and properties, which are relevant in our context here.

2.1.1 Planck2018 low-ℓ data

For what concerns the low-ℓs, following Planck 2018 [204], the baseline low-ℓ likelihood adopted in the 2018 legacy release exerts the combination of the following three functions. The first one is a Gibbs-sampling approach in total intensity (TT-low-ℓ likelihood) and is based on the Bayesian posterior sampling framework which has been implemented by the COMMANDER code [205, 206] that has been used extensively in the Planck releases. The second one relies on the estimation of cross-spectra based on the High Frequency Instrument (HFI) channels, 100 and 143 GHz, extended to polarization and including the subtraction of the main diffuse Galactic foreground contamination [207]. The third function is an updated version of a pixel based likelihood using both total intensity and polarization for $l \leq 29$ [206]; it is based on the 70 GHz Planck channel of the Low Frequency Instrument (LFI), where the diffuse Galactic foregrounds in polarization have been subtracted using the 30 GHz and 353 GHz maps.

2.1.2 Planck2018 high-ℓ data

At high-ℓs, the 2019 Planck likelihood corresponds to those used in previous releases [206], and exploits a power spectrum estimation at multipoles ($30 < \ell < 2500$), using HFI data. It includes nuisance parameters introduced to control residual systematics, and point source contamination. Planck assumes a Gaussian distribution for the data, written as

$$-log\mathcal{L}(\hat{C}|C(\theta)) = \frac{1}{2}(\hat{C} - C(\theta))^T \mathbb{C}^{-1}(\hat{C} - C(\theta)) + const.\,, \tag{4.17}$$

where $\hat{C}$ is the data vector and $C(\theta)$ is the model with (cosmological and nuisance) parameters θ and $\mathbb{C}$ the covariance matrix. We utilise Planck data in from the 100, 143 and 217 GHz channels. Concerning the foreground residual contamination, and associated nuisance parameters, the details of the models and uncertainties are given in Planck 2015 [206], while the covariance matrix $\mathbb{C}$ is described in Planck 2013 [205]. In comparison with the 2015 releases, the Planck analysis improves the treatment of several systematics and foreground effects. For a comprehensive explanation of cut selections, masks, optical beams and binning of data and also the Galactic and extra-Galactic foregrounds, noise models and calibration, see Planck 2018 [204]. Overall, we used both temperature and polarization high-ℓ data (`TTTEEE`) from Planck.

2.1.3 Planck2018 CMB Lensing data

Gravitational lensing of the CMB can considerably improve the constraints on cosmological parameters. The effect of lensing actually is to remap the CMB fluctuations with an almost Gaussian field representing the lensing angle, with a standard deviation of about 2 arcminutes: the anisotropy in a given direction $\hat{n}$, is redirected onto the new path represented by $\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}})$ where $\phi(\hat{\mathbf{n}})$ is the CMB lensing potential and $\nabla\phi(\hat{\mathbf{n}})$ denotes lensing deflection angle. The Planck-2018 lensing likelihood [208] corresponds to the one used in the previous release [209], extended to cover the $8 \leq \ell \leq 400$ multipole interval, which might be important to improve the capabilities of CMB lensing to break geometrical degeneracies in the primary CMB anisotropies. The likelihood is approximated as Gaussian with a fixed covariance estimated from simulations, corresponding to

$$-2log\mathcal{L}_\phi = \mathcal{B}_i^L(\hat{C}_L^{\phi\phi} - C_L^{\phi\phi,th})\left[\Sigma^{-1}\right]^{ij} \mathcal{B}_i'^L(\hat{C}_{\prime L}^{\phi\phi} - C_{\prime L}^{\phi\phi,th})\,, \tag{4.18}$$

where Σ is the covariance matrix and $\mathcal{B}_i^L$ are the binning functions, see Planck 2018 [208] for details.

2.2 LSS

We consider LSS tracers, relevant for the dynamics of perturbation as well as for background: BAO [210, 211, 212] including Lyman-α quasar cross/auto correlations [213, 214, 215, 216], Type Ia supernova [217], the recent Hubble diagram for Quasars [218, 219], and prior on the present value of the Hubble constant H_0 which we discuss below.

2.3 Baryon Acoustic Oscillations

As we described in chapter 3, Sec. 2.5, BAO represent the imprints of the oscillations of the photon-baryon plasma in the early Universe and they can be used as a standard ruler in the distribution of the structures today corresponding to the size of the sound horizon at baryon drag

$$r_s(z_{\rm drag}) = \int_0^{\eta_{\rm drag}} c_s d\eta = \int_{z_{\rm drag}}^\infty \frac{c_s}{H(z)} dz \simeq 150\,{\rm Mpc}\,,$$

where η is the conformal time, and c_s the sound speed. A well known feature of the BAO is represented by a bump in the correlation function of the distribution of the same kind of galaxies and as wiggles in the matter power spectrum which is actually the Fourier transform of the correlation function [220, 221, 222, 223, 224, 225, 226, 227, 228]. Measurements of BAO from a galaxy sample constrain the angular diameter distance $D_A(z)$ and the expansion rate of the Universe $H(z)$, either separately or in combination through the Alcock-Paczynski test [146]. Indeed, the characteristic scale along the line-of-sight, $s_\parallel(z)$, provides a measurement of the Hubble parameter through $H(z) = \dfrac{c\,\Delta z}{s_\parallel(z)}$ while the tangential mode, $s_\perp$, provides a measurement of angular diameter distance $D_A(z) = s_\perp(1+z)\Delta\theta$ [228]. On the other hand the Alcock-Paczynski test [146] constrains the product of $D_A(z) \times H(z)$, more precisely the volume distance, D_v, defined as:

$$D_v(z) = \left[(1+z)^2 D_A^2(z) \frac{cz}{H(z)} \right]^{1/3}. \tag{4.19}$$

In this work we use the following BAO data sets: the six-degree-Fields Galaxy survey (6dFGS) at effective z ($z_{\rm eff}$) equal to 0.106 [210]; the Sloan Digital Sky Survey (SDSS) Main Galaxy Sample (SDSS-MGS) at $z_{\rm eff} = 0.15$ [211]; the complete SDSS-III Baryon Oscillation Spectroscopic Survey cosmological analysis of the Data Release (DR)12 galaxy sample which has been divided into three partially overlapping redshift slices centered at $z_{\rm eff} = 0.38$, 0.51, and 0.61 [212]; the measurement of BAO correlations at $z = 2.33$ with Baryon Oscillation Spectroscopic Survey (BOSS) SDSS DR12 $Ly\alpha$-Forest [213]. Moreover, we also include BAO from the complete SDSS-III $Ly\alpha$-quasar cross-correlation combined with $Ly\alpha$ auto-correlation function at $z = 2.40$ [214] and the BAO measurement at $z = 2.34$ from the recent analyses of correlations (auto-correlation and cross-correlation) of Lyα absorption performed by eBOSS DR14 [215, 216]. The first two data sets measure D_v/r_s, while the others measure $D_A(z_{\rm eff})$, $D_M(z_{\rm eff})$ (i.e., the comoving angular diameter distance $D_M = (1+z)D_A$) and $H(z_{\rm eff})$. In all cases, the BAO measurements are modelled as distance ratios, and therefore they provide no direct measurement of H_0. However, they provide a link between the expansion rate at low redshifts and the constraints that is placed by CMB data at $z \approx 1100$. Therefore, it is essential to combine CMB with BAO, because the latter can break the degeneracies from CMB measurements and can offer tighter constraints on the background evolution of different dark energy or modified gravity models [229, 230]. Finally, notice that BAO measurements are largely unaffected by the non-linear evolution of structures because the acoustic scale is considerably large.

2.3.1 Supernovae

As we explained in chapter 2, Sec. 1, SNe are known to be the most important probe of the accelerated expansion of the Universe and DE behavior [28, 29]. They provide accurate measurements of the luminosity distance as a function of z. However, the absolute luminosity measurements of SNe is considered to be uncertain and it is marginalized out, removing any constraints on H_0. Here we used the analysis of the SNe type-Ia by the Joint Light Curve Analysis (JLA) [217], which is actually constructed from the Supernova Legacy Survey (SNLS) and SDSS supernova data together with the low redshift supernova data sets.

The motivation for using JLA supernovae dataset rather than the more updated Pantheon dataset [231] is twofold. First, the JLA dataset was used in the Planck 2015 paper [229] on dark energy and modified gravity.

Using JLA allows us to directly compare our constraints on early dark energy parameters (third column of Table 4.5 ("EDE fixed c_{eff}^2 & c_{vis}^2") with those from the 2015 Planck paper (last column of Table 3), with the only difference of the updated BAO dataset of our analysis. The second reason is that the quasar dataset that we used in some of our runs is calibrated against the JLA supernova dataset, and, thus, using JLA we are fully consistent. Anyway, we also checked that if we replace the JLA with the Pantheon supernovae dataset the changes in our results are not statistically significant.

2.3.2   Quasars

Our analysis includes the Hubble diagram for Quasars (QSOs) as described in Risaliti et al. [218], where the constraining power is based on the non-linear relation between the ultraviolet (UV) and X-ray luminosity (L_X) of QSOs. Where the $L_X - L_{UV}$ relation is parametrized as a linear dependence between the logarithm of the monochromatic luminosity at 2500 Å(L_{UV}) and the α_{OX} parameter defined as the slope of a power law connecting the monochromatic luminosity at 2 keV (L_X), and L_{UV}: $\alpha_{OX} = 0.384 \times \log(L_X/L_{UV})$. Luminosities are derived from fluxes through a luminosity distance calculated adopting a standard ΛCDM model with the best estimates of the cosmological parameters Ω_m and Ω_Λ (described in chapter 3, Sec. 1). When expressed as a relation X-ray and UV luminosities the $\alpha_{OX} - L_{UV}$ relation becomes:

$$\log(L_X) = \beta + \gamma \log(L_{UV}) \, . \tag{4.20}$$

By using the definition of flux as $F = L/(4\pi D_L^2)$, the theoretical relation for the X-ray is

$$\log(F_X) = \Phi(F_{UV}, D_L) = \gamma \log(F_{UV}) + \left[\left[\beta + (\gamma - 1) \log(4\pi) \right] + 2(\gamma - 1) \log(D_L) \right], \tag{4.21}$$

where D_L is the luminosity distance (explained in chapter 2, Sec. 2.1) which, for a ΛCDM model with a fixed cosmological constant Λ, is given by

$$D_L(z, \Omega_m, \Omega_\Lambda) = \frac{(1+z)}{\sqrt{\Omega_K}} \sinh \sqrt{\Omega_K} \times \int_0^z \frac{dz}{H_0 \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda + \Omega_K(1+z)^2}} \, , \tag{4.22}$$

With $\Omega_K = 1 - \Omega_m - \Omega_\Lambda$ (see chapter 3, Sec. 1). By minimizing the likelihood function (LF) (see chapter 3, Sec. 2.1.1 for a description of a likelihood function), one can actually fit the equation 4.21 as follows:

$$\ln(LF) = \Sigma_{i=1}^N \left\{ \frac{[\log(F_X)_i - \Phi(F_{UV}, D_L)_i]}{s_i^2} + \ln(s_i^2) \right\} \, , \tag{4.23}$$

where s_i is the error, $s_i^2 = \sigma_2^2 + \delta^2$, with σ_i and δ indicating the measurement errors over F_X and the global intrinsic dispersion, respectively. We note that the dispersion δ is much higher than typical values of σ_i. And N is the number of QSOs, here $N = 1598$.

In this analysis we used the quasars data points at high redshifts from the recent 1598 quasar data set in the redshift range corresponding to $0.03 < z < 5.1$ [219]. In Figure 4.2 we plot the Hubble diagram of supernovae from the JLA survey and QSO data, showing the markedly different ranges probed and therefore the different constraining power of these two observables.

2.3.3   The Hubble constant

As discussed by Planck 2015 [229], dark energy and modified gravity are poorly affected by the physics of recombination, the main influence coming from the integrated Sachs-Wolfe effect and CMB lensing. Following the same reasoning, we use here a re-analysis of the Riess et al. 2011 [232] Cepheid data made by Efstathiou et al. 2014 [233]. By using a revised geometric maser distance to NGC 258 from Humphreys et al. 2013 [234], Efstathiou et al. 2014 [233] obtain the following value for the Hubble constant which we adopt (unless specified otherwise) as a conservative H_0 prior throughout this analysis:

$$H_0 = (70.6 \pm 3.3) \, \text{km/s/Mpc} \, . \tag{4.24}$$

Let us point out that, being very broad, this prior is consistent within 1σ both with the recent direct measurements by Riess et al. 2019 [235] and with Planck 2018. The motivation for choosing this particular H_0 prior is that it is broad enough to be compatible both with Planck and with supernovae data. Moreover, it has been used in Planck 2015 dark energy and modified gravity paper [229], and, thus, it is suitable for comparison.
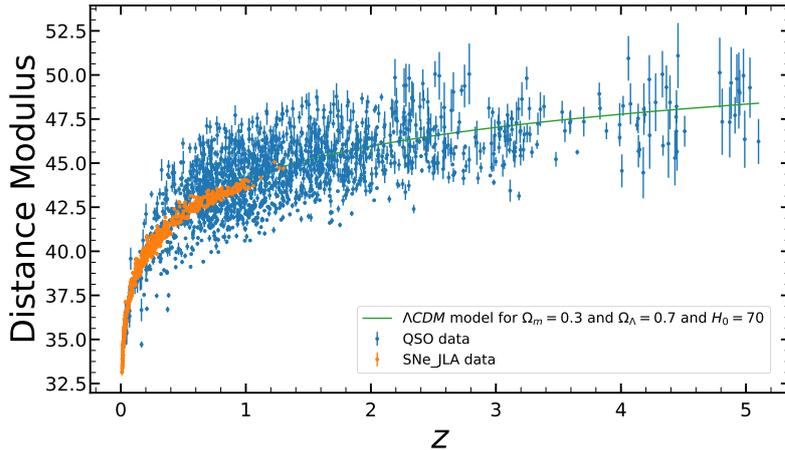
Figure 4.2: The Hubble diagram of supernovae from JLA survey (orange points) and quasars (blue points). The green curve shows the theoretical flat ΛCDM model with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70$ km/s/Mpc.

3 Impact of EDE on cosmological observables

The goal of this Section is to present the effect that a given amount of EDE has on the main observables considered here: the CMB. We will phenomenologically describe the effect of varying c_{eff}^2 and c_{vis}^2 from 0 to 1, separately for the cases in which the EDE component is present or not, i.e. with $\Omega_{\text{eDE}} \neq 0$ or $= 0$. We will then single out the Integrated Sachs Wolfe contribution (described in chapter 2, Sec. 5.2.2) and the effect on CMB lensing. Finally, we will also show the impact on the linear matter power spectrum, a quantity which is however not used in the present analysis, in order to see the implications that this model could have in terms of the rms value of the amplitude of density fluctuations at 8 Mpc/h, corresponding to the σ_8 density parameter.

The CMB angular power spectrum (see also chapter 2, Sec. 5) can be written as the covariance of the total intensity fluctuations in harmonic space:

$$C_l = 4\pi \int \frac{dk}{k} \mathcal{P}_\chi \, |\Delta_l(k, \eta_0)|^2 \,, \tag{4.25}$$

where $\mathcal{P}_\chi$ is the initial power spectrum and η_0 is today's conformal time. Here $\Delta_l(k, \eta_0)$ is the transfer function for photons, which has the following form on large scales:

$$\Delta_l(k, \eta_0) = \Delta_l^{LSS}(k) + \Delta_l^{ISW}(k) \,, \tag{4.26}$$

where $\Delta_l^{LSS}(k)$ is the contribution of the last scattering surface given by the ordinary Sachs-Wolfe (SW) effect and the total intensity anisotropy and $\Delta_l^{ISW}(k)$ is the contribution of the ISW effect. The latter is due to the time change of the potential ϕ along the line of sight as follows (We brought a brief explanation of the ISW effect on the CMB spectra in chapter 2, Sec. 5.2.2):

$$\Delta_l^{ISW}(k) = 2 \int d\eta \, e^{-\tau(\eta)} \, \dot{\phi} \, j_l[k(\eta - \eta_0)] \,, \tag{4.27}$$

where $\tau(\eta)$ is the optical depth coming from the scattering of photons along the line of sight, $j_l(x)$ is the spherical Bessel function and $\dot{\phi}$ is the derivative of the potential with respect to the conformal time. As already discussed in the literature (see e.g. in Bean & Dore 2004 [236] and Weller & Lewis 2003 [237]), perturbations in the DE density component with a constant equation of state have a large effect on the largest scales probed by the CMB.

3.1 Perturbation effects on the CMB angular power spectrum

Following the perturbation equations 4.14, 4.15 and 4.16, we plot the CMB angular power spectrum for different values of c_{eff}^2 and c_{vis}^2 in Fig. 4.3. Note that here we are switching off the presence of EDE, i.e. $\Omega_{\text{eDE}} = 0$, in order to investigate phenomenologically the pure effect of perturbations.
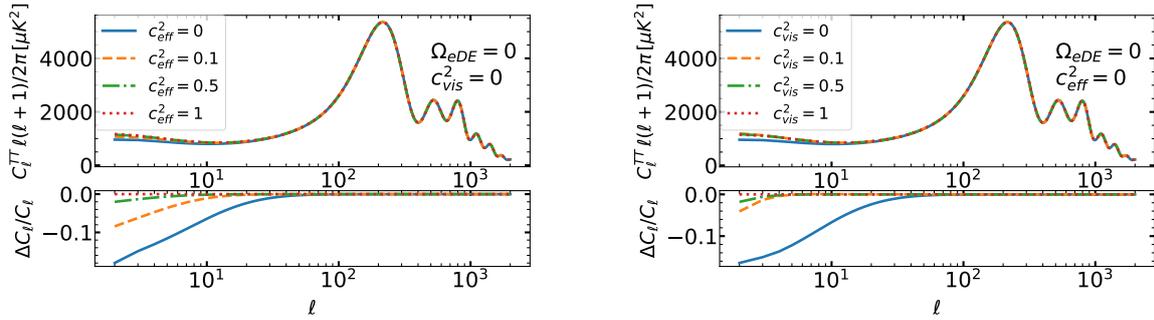
Figure 4.3: The effect of perturbations on the TT CMB angular spectrum for fixed values of $\Omega_{\mathrm{eDE}} = 0$ and $w_0 = -0.8$. Note that we use $w_0 = -0.8$ to make the difference visible, although $w_0 = -0.8$ is already excluded by data. In the left panel the value of the viscous sound speed has been fixed, $c_{\mathrm{vis}}^2 = 0$, in order to see the effect of varying the effective sound speed c_{eff}^2 on the CMB angular power spectra. In the right panel the value of the effective sound speed has been fixed, $c_{\mathrm{eff}}^2 = 0$, to see the effect of varying the viscous sound speed c_{vis}^2. The relative effects of each case has been shown in the lower panels.

As it can be seen in Fig. 4.3, the effect on the CMB spectrum are only confined at relatively large scales due to the ISW effect. The reason would be the fact that since in the scenario with the constant equation of state and a negligible energy component in the early Universe which means: $(w(a) = w_0$ and $\Omega_{\mathrm{eDE}} = 0)$, the dark energy has a contribution in the energy density only at late times. Therefore the CMB power spectrum can be only influenced by the late ISW effect. The effect that would be achieved by increasing c_{eff}^2 or c_{vis}^2 is the higher ISW power. And this fact replies the increased potential caused by the dark energy. Although the dark energy perturbation would help to keep the potential constant, increasing c_{eff}^2 or c_{vis}^2 can reduces the dark energy perturbation and this fact leads to diminish the decay of the potential. And by the decay of potential the ISW power increases. In the left panel of the Fig. 4.3, by fixing $c_{\mathrm{vis}}^2 = 0$ and increasing the value of c_{eff}^2 gradually from 0 to 1, as discussed above, the dark energy perturbation contribution decreases and leads to the decay in potential, Therefore the ISW effect increases and due to the equation 4.26 the transfer function of photons increases as well, so as can be seen in the equation 4.25, the CMB angular power spectrum increases. In comparison, the same effect is happening in the right panel, by fixing $c_{\mathrm{eff}}^2 = 0$ and increasing the amount of c_{vis}^2 gradually from 0 to 1. In this way the perturbations are suppressed and the ISW effect increases as a consequence of the dynamics deriving from the suppression itself. Therefore, in both cases, by fixing one parameter and increasing the value of the other, we have an increase in the amount of ISW component and the CMB angular power spectrum as well. As already discussed in the literature, the feasibility of accurately measuring one of these parameters is strongly undermined by the presence of cosmic variance on the angular scales in which the ISW is effective.

We now fix one of the parameters to 1 and increase the other parameter gradually from 0 to 1. As expected, and shown in Figs. 4.3 and 4.4, by fixing one perturbation parameter and changing the other we observe an effect which is similar to the one of perturbations on the CMB angular power spectra: the effect of increasing c_{eff}^2 or c_{vis}^2 reduces the DE perturbations and this can leads to the decaying of the potential and therefore to a larger ISW effect. For all the cases, the impact is mostly seen at large scales, at multipoles $l < 80$ and bound to be below the 3-4% level. In the following Section we are going to check the effect of the early dark energy model on the CMB.

3.2 Effects of the early dark energy on the CMB angular power spectrum

It is important now to focus on the effect of a given amount of early dark energy on the CMB. Therefore, in the following figures the combined effect of perturbations and a non-zero energy density in the dark energy fluid can be investigated.

Fig. 4.5 shows the effects of perturbations when $\Omega_{\mathrm{eDE}} \neq 0$. They are visible also on smaller scales with respect to a pure ISW, due to the contribution of the early ISW, associated with a non-zero EDE. The difference is particularly visible for the first acoustic peak. The reason is that, for $\Omega_{\mathrm{eDE}} \neq 0$, the EDE influences directly the recombination process so the EDE can affect on the evolution of the acoustic oscillations before recombination. Although the differences are small, but more significant with respect to the ISW, due to the reduced cosmic variance. Finally, we notice that, similarly to the previous case, by increasing the sound speed, perturbations in the EDE get more and more suppressed, leading to a stronger decay of the metric perturbations.
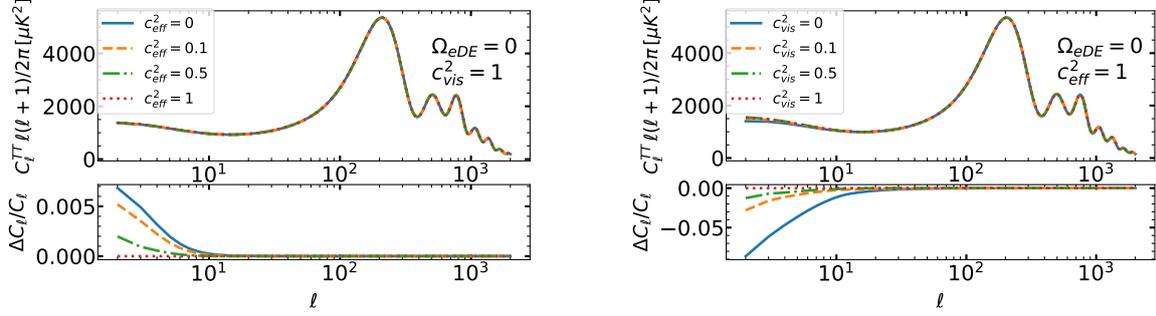
Figure 4.4: The effect of perturbations on the TT CMB angular power spectrum for the value of $\Omega_{\mathrm{eDE}} = 0$. The value of the viscous sound speed has been fixed, $c_{\mathrm{vis}}^2 = 1$ to see the effect of the effective sound speed c_{eff}^2 on the CMB angular power spectra (left panel). The value of the effective sound speed has been fixed, $c_{\mathrm{eff}}^2 = 1$ to see the effect of the viscous sound speed c_{vis}^2 in the right panel.
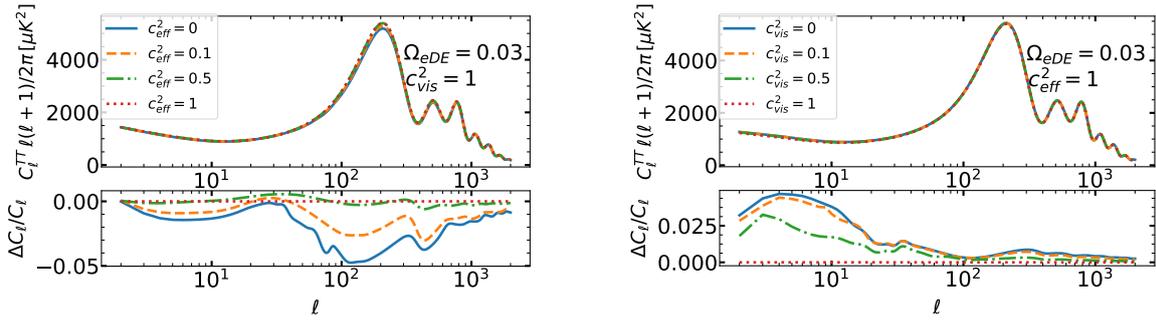
Figure 4.5: EDE effects on the TT CMB angular power spectrum for the indicated value of early dark energy $\Omega_{\mathrm{eDE}} = 0.03$. Note that, although the value of Ω_{eDE} chosen here is already excluded by data, we use it to make the difference more visible. The value of the viscous sound speed has been fixed, $c_{\mathrm{vis}}^2 = 1$ in the left panel and $c_{\mathrm{eff}}^2 = 1$ in the right panel, respectively.

The behaviour of the ISW effect is shown in detail in Fig. 4.6 which displays only the ISW component of the CMB spectra, highlighting the late $((l < 30))$ and early $(l \sim 120)$ contributions.
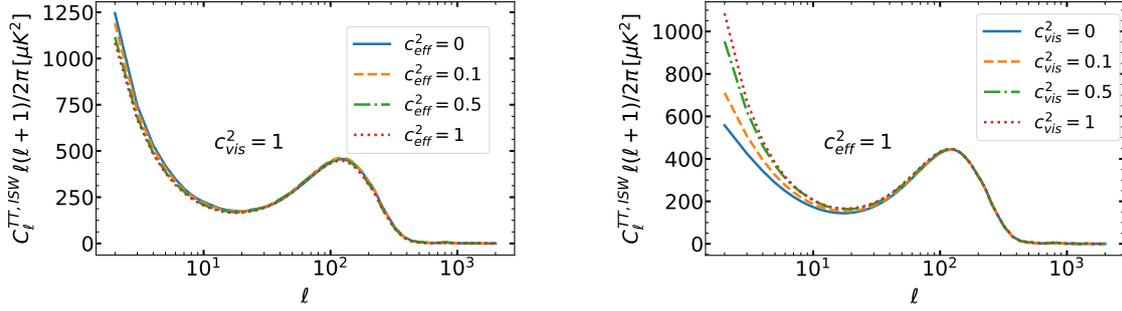
Figure 4.6: The behaviour of the ISW component of the TT CMB angular power spectrum when the EDE effect is on. The values of $\Omega_{eDE} = 0.03$ and $w_0 = -0.8$ are fixed in all the curves. As in previous figures, notice that the values of Ω_{eDE} and w_0 are chosen to make the differences more visible, although, as shown in Table 4.3, these values are already excluded by data. In the left(right) panel the value of the viscose(effective) sound speed has been fixed to see the effect of the effective(viscose) sound speed.

3.3   CMB Lensing

Similarly to the previous Section, Fig. 4.7 shows the lensing potential angular power spectra with and without EDE, for different values of the perturbation parameters. In the case $\Omega_{eDE} = 0$, one can clearly see that if by taking $c^2_{eff} = 1$ or $c^2_{vis} = 1$, perturbations are suppressed, making the lensing potential nearly equivalent in the two cases. In the case with $c^2_{eff} = 0$ and $c^2_{vis} = 0$, where no friction is caused to perturbation growth, the lensing potential is significantly enhanced (see the red dotted curve in Fig. 4.7). Similarly, when $\Omega_{eDE} = 0.03$, c^2_{eff} or c^2_{vis} is equal to 1 causes a suppression onto perturbations, while for c^2_{eff} and c^2_{vis} equal to 0, the lensing potential would be significantly enhanced. Thus, $\Omega_{eDE} \neq 0$ causes a stronger enhancement in comparison with non-EDE scenarios because of the fact that the presence of the EDE leads to a larger DE clustering, causing a more pronounced lensing power.

3.4   Effects on the matter power spectrum

Fig. 4.8 shows the impact of EDE for different values of the perturbation parameters. The left panel is without the EDE effect and the right panel we have EDE by $\Omega_{eDE} = 0.03$. The blue solid curve in the right panel, shows the matter power spectrum for $c^2_{eff} = 0$ and $c^2_{vis} = 0$, i.e. when perturbations are not affected by friction at all. We can clearly see that there would be a significant enhancement at small scales in comparison with the other combination of the perturbation parameters (c^2_{eff} and c^2_{vis}). Basically this enhancement can be decreases if we take $c^2_{eff} = 1$ or $c^2_{vis} = 1$ because of the suppression of the perturbations. In each panel the relative differences form the reference case i.e. ($c^2_{eff} = 1$ and $c^2_{vis} = 0$) is plotted as well.

Overall, the impact of the amount of EDE and varying c^2_{eff} and c^2_{vis} on cosmological observables, mainly CMB here, can be summarize as follows. In the absence of the EDE component ($\Omega_{eDE} = 0$), we can see two different effects on the CMB, mainly on the ISW power (see Figs. 4.3 and 4.4). If we switch off one of the c^2_{eff} or c^2_{vis} parameters, by increasing the other parameter we could achieve a higher ISW power about $\sim 10\%$. (Figs. 4.3). Contrary, if we fix one of the parameters to 1 and increase the other parameter, the impact depends on which parameter we are fixing: if we fix $c^2_{vis} = 1$, by increasing the other parameter we would have a decrease in the ISW power by about $\sim 0.5\%$; instead, if we fix $c^2_{eff} = 1$ and increase the other parameter, we would have $\sim 5\%$ decrease in ISW power (Fig. 4.4). Therefore it seems that even in the case where there is no EDE component ($\Omega_{eDE} = 0$), each parameter has its own effect on the CMB power spectra. So we need both parameters c^2_{eff} and c^2_{vis} to describe the ISW effect. By switching on the EDE component, this differences can be visible also on smaller scales (see Fig. 4.5).
Besides primary CMB, we also investigated the impact of switching on and off c^2_{eff} or c^2_{vis} parameters on the CMB lensing spectra in the presence or in the absence of a non-zero EDE component separately (Fig. 4.7). When $\Omega_{eDE} = 0$, fixing either $c^2_{eff} = 1$ or $c^2_{vis} = 1$, and setting to 0 the other perturbation parameter does not lead to any significant variation of the lensing spectra. Only if both $c^2_{eff} = 0$ and $c^2_{vis} = 0$, then there is
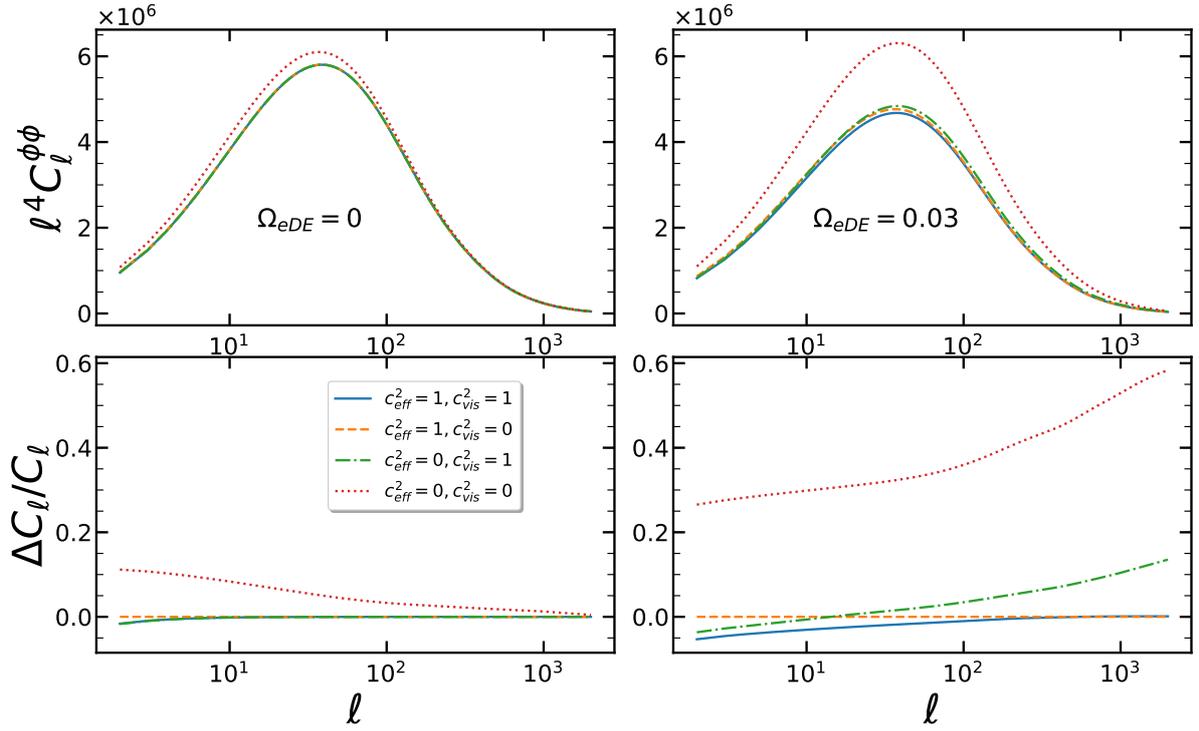
Figure 4.7: Lensing potential power spectrum in scenarios with vanishing or finite Ω_{eDE} for different values of perturbation parameters (c_{eff}^2 and c_{vis}^2). Left panels show the scenarios without early dark energy $\Omega_{\mathrm{eDE}} = 0$ and the right panels show the lensing potential with the early dark energy parameter $\Omega_{\mathrm{eDE}} = 0.03$. The ratios with respect to a model with $c_{\mathrm{vis}}^2 = 0$ and $c_{\mathrm{eff}}^2 = 1$ has been shown in the lower panels; $w_0 = -0.8$ is fixed and $\Omega_{\mathrm{eDE}} = 0.03$ when different from zero. We have to specify that the values chosen for w_0 and Ω_{eDE} are only meant to make the difference more visible. Table 4.3 will show that these values are actually excluded by data.
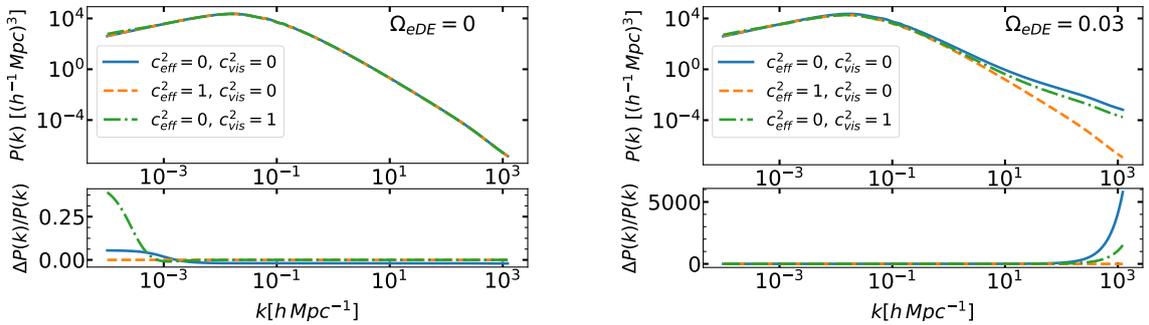
Figure 4.8: Linear matter power spectrum for $\Omega_{\mathrm{eDE}} = 0.03$ (right panel) and $\Omega_{\mathrm{eDE}} = 0$ (left panel) models for $w_0 = -0.8$ and different values of the perturbation parameters, as indicated. As already mentioned in the captions of previous figures, for the sake of visibility we chose $w_0 = -0.8$ and $\Omega_{\mathrm{eDE}} = 0.03$, although these values are already excluded by data.

Table 4.1: The impact of EDE and perturbation parameters on CMB spectrum.

Observables	Ω_{eDE}	c_{vis}^2	c_{eff}^2	Effect
ISW	$= 0$	0	Increasing	$\sim 10\%$ Increase at large scales
		Increasing	0	$\sim 10\%$ Increase at large scales
		1	Increasing	$\sim 0.5\%$ Decrease at large scales
		Increasing	1	$\sim 5\%$ Increase at large scales
	$\neq 0$	1	Increasing	$\sim$ up to 4% Increase at all scales
		Increasing	1	$\sim$ up to 4% Decrease at all scales
CMB lensing	$= 0$	0	0	$\sim 10\%$ Increase
	$\neq 0$	0	0	$\sim 45\%$ Increase

no friction in the growth of perturbations, and the lensing potential is significantly enhanced (Fig. 4.7-right panel). On the other hand, in the presence of a non-zero EDE component, even setting to 0 only one of the perturbation parameters leads to a noticeable increase of the lensing power. Still we can see the significant enhancement by switching off both parameters (Fig. 4.7-right panels).

We summarized all the effects on the CMB spectrum in Table 4.1.

Finally, we study the impact on the linear matter power spectrum. As can be seen in Fig. 4.8, each components can have different effects on the linear matter power spectrum both in the absence and in the presence of a non-zero EDE component.

4 Constraints on spacetime dynamics for EDE

In chapter 3, Sec. 2.1.2, we explained the MCMC sampling methodology which allows for an efficient determination of parameter estimation. Utilizing this method, in this Section we derive the constraints on the EDE scenarios, considering the latest data sets, and the phenomenology outlined above, in a wider parameter space that includes variation in the total neutrino mass Σm_ν, and the tensor-to-scalar-ratio r.

4.1 Methodology and parametrization

We analyze here the EDE models by using a modified version of the Boltzmann equation solver CAMB [4] in order to account for equations (4.1) and (4.14)-(4.16) through varying the following set of parameters:

$$\{\Omega_{\mathrm{b}}h^2, \Omega_{\mathrm{c}}h^2, 100\theta_{\mathrm{MC}}, \ln[10^{10}A_s], n_s, \tau, \Sigma m_\nu, r, \Omega_{\mathrm{eDE}}, w_0, c_{\mathrm{eff}}^2, c_{\mathrm{vis}}^2\}. \qquad (4.28)$$

We consider the standard six parameters of the concordance ΛCDM model [208] which we described them in chapter 3, Sec. 1, i.e. the baryon and CDM fractional densities today $\Omega_{\mathrm{b}}h^2$, $\Omega_{\mathrm{c}}h^2$, 100 times of the ratio between the sound horizon and the angular diameter distance at decoupling ($100 \times r_*/D_A$) which is usually denoted by $100\,\theta_{MC}$, the primordial scalar perturbations amplitude $\ln[10^{10}A_s]$, the scalar spectrum power-law index n_{s}, and the reionization optical depth τ. In addition, we include the neutrino masses Σm_ν, and the ratio of the tensor primordial power to the scalar curvature one at $k_0 = 0.05$ Mpc^{-1} which is called r. The last four parameters are related to the EDE scenario, in which the possibility of clustering also has been included. As already discussed in Sec. 1, the parameters could be described as follows: Ω_{eDE}, the non-negligible fractional DE density in the early Universe, w_0, the equation of state parameter today, c_{eff}^2, the effective sound speed, and c_{vis}^2, the viscose sound speed. As already discussed in Sec. 1.2, the two last parameters characterize perturbation. In order to derive constraints on the parameters, we used the last version of the MCMC package CosmoMC [90], that has a convergence diagnostic based on the Gelman and Rubin statistic and includes the support for the Planck data release 2018 Likelihood code [204]. We assume flat priors on the parameters as listed below in Table 4.2. For the ΛCDM parameters, they're significantly wider with respect to the present constraints. For the EDE parameters, we allow for full freedom in the interesting range.

4.2 Constraints on EDE

Table 4.3 and Fig. 4.9 show the constraints on cosmological parameters on the EDE scenario. All parameters, including EDE ones, are allowed to vary within the priors. Notice that although perturbations are included,

Table 4.2: Flat priors on the cosmological parameters assumed in this paper.

Parameter	Prior
$\Omega_\mathrm{b} h^2$	$[0.005, 0.1]$
$\Omega_c h^2$	$[0.001, 0.99]$
$100\theta_{MC}$	$[0.5, 10]$
$\ln[10^{10} A_\mathrm{s}]$	$[1.61, 3.91]$
n_s	$[0.8, 1.2]$
τ	$[0.01, 0.8]$
$\Sigma m_\nu \, [\mathrm{eV}]$	$[0.056, 1]$
r	$[0, 0.3]$
Ω_eDE	$[0, 0.1]$
w_0	$[-1, 0]$
c_eff^2	$[0, 1]$
c_vis^2	$[0, 1]$
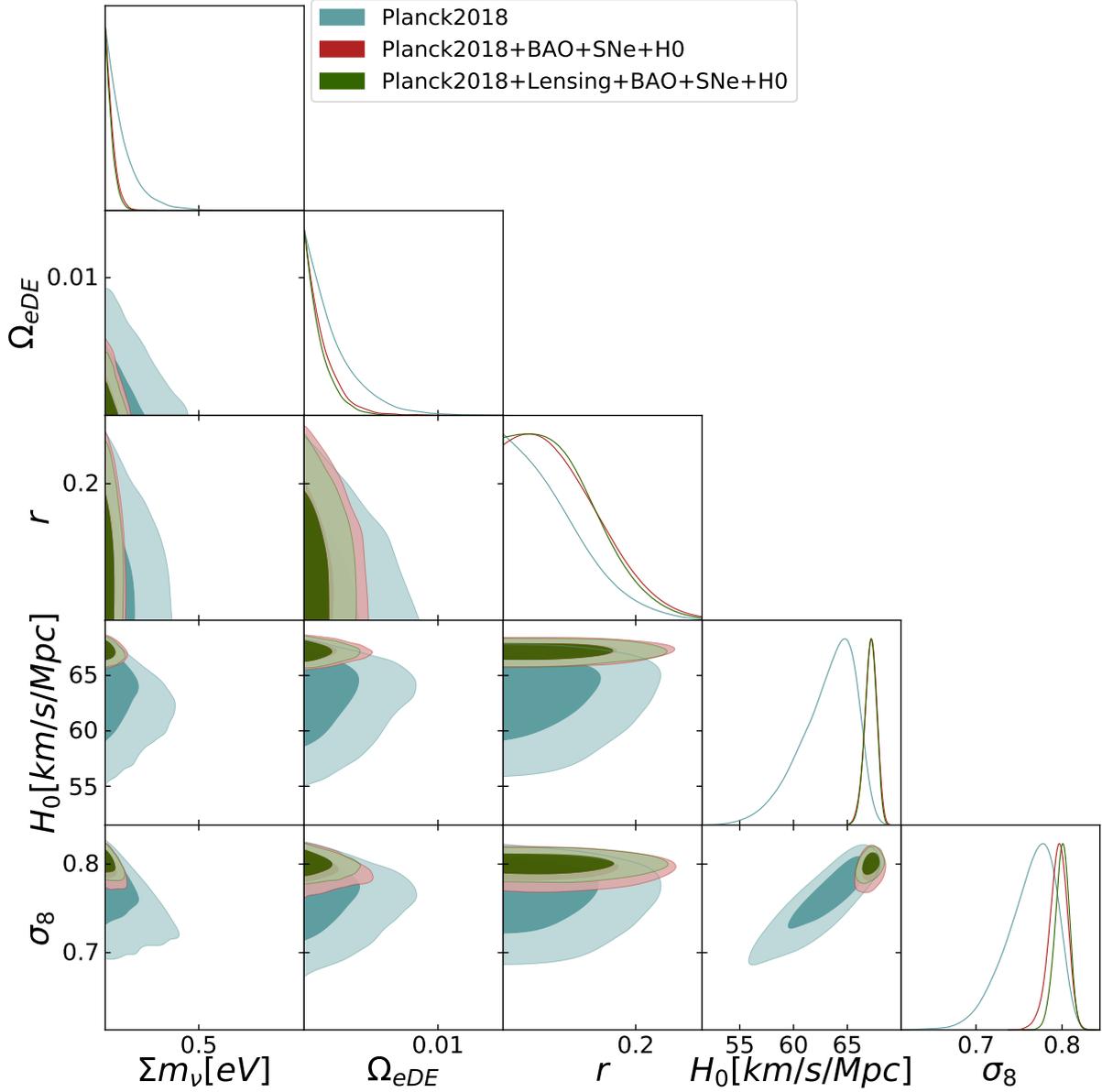
Figure 4.9: Marginalized 2D and 1D posteriors on $\sum m_\nu$, Ω_eDE, r, H_0, σ_8 including EDE. Blue contours show the 68% and 95% confidence level regions allowed from Planck (TT, TE, EE + lowE) 2.1 measurements. Contours in red include the BAO 2.3, JLA Supernova data set 2.3.1 and H_0 prior 2.3.3 as well. Green contours, in addition to the previous data set, also include the Planck Lensing 2.1.3 data. see the first 4 columns of Table 4.3 for the numerical values.

both $c_{\rm eff}^2$ and $c_{\rm vis}^2$ turn out to be always unconstrained; therefore we do not show them. The goal is to see the different constraining power of the combination of data sets, while we will address the role of priors in Sec. 4.4. We can clearly see that by adding BAO, SNe and also the H$_0$ prior to the Planck data set, the constraints become tighter (red contours), as expected. By including the lensing data we can see even tighter (green) contour plots, but always overlapping very well with the results by Planck only. The combination of data sets constraining the background and lensing pushes $\Omega_{\rm eDE}$ to lower values, passing from 0.0063 to 0.0039 ($\sim 40\%$ smaller 2σ upper limit) and by including the CMB lensing the 2σ upper limits decreases more to 0.0033 ($\sim 50\%$ smaller than the Planck-only case). The 2σ upper limit on the parameter w_0 becomes much tighter when background data are included, decreasing from -0.72 to -0.95 (about $\sim 30\%$), while including CMB lensing or QSO does not have a significant impact.

We can consider the anti-correlation between the equation of state today w_0 and the tensor to scalar ratio r at least when we are using only Planck data, which is reduced significantly by adding the BAO+SNe+H$_0$ data sets. There is also a degeneracy between $\Omega_{\rm eDE}$ and w_0 and between Σm_ν and the EDE parameters as well. A relative large value of the total neutrino mass ~ 0.3 eV would require a small value of $\Omega_{\rm eDE}$, but when background data are included the neutrino mass is much more constrained and somewhat larger values of $\Omega_{\rm eDE}$ can fit the data.

Overall, the main conclusion of this analysis is that the amount of EDE is bound to be well below 1% at $> 2\sigma$ confidence level. This confirms the limits found by the Planck collaboration [229], Table 3, where $\Omega_{\rm eDE} < 0.007$ at 2σ CL for fixed neutrino mass. Here we show that these bounds are robust against a variation of the neutrino mass summation, and that they improve once we include lensing and QSO. The degeneracies between the parameters describing the DE model (w$_0$, $\Omega_{\rm eDE}$) and the cosmological parameters that extend the simple vanilla 6-parameter space, neutrino mass and tensor to scalar ratio, are present but are not strong. It is also evident that the Hubble parameter is remarkably stable and constrained to be very close to its ΛCDM value. On the contrary, σ_8 inferred from Planck only within this EDE scenario is significantly lower than in ΛCDM, thus alleviating the tension with the low σ_8 values inferred from weak lensing [238]. However, the tension is fully restored once background data and CMB lensing are included.

Table 4.3: Mean values and 1σ marginalized error on the cosmological parameters. For Σm_ν, r, $\Omega_{\rm eDE}$, and w_0 we report the 95% upper limits.

Parameter	Planck	Planck+BAO+SNe+H$_0$	Planck+BAO+SNe+H$_0$+lensing	Planck+BAO+SNe+H$_0$+lensing+QSO
$\Omega_b h^2$	0.02226 ± 0.00016	0.02246 ± 0.00013	0.02245 ± 0.00014	0.02245 ± 0.00014
$\Omega_c h^2$	0.1213 ± 0.0015	0.1187 ± 0.0010	0.1190 ± 0.0010	0.1189 ± 0.0010
$100\theta_{MC}$	1.04071 ± 0.00034	1.04105 ± 0.00030	1.04103 ± 0.00030	1.04103 ± 0.00030
$\ln(10^{10} A_s)$	3.048 ± 0.016	3.048 ± 0.016	3.053 ± 0.015	3.054 ± 0.015
n_s	0.9622 ± 0.0048	0.9686 ± 0.0040	0.9680 ± 0.0039	0.9679 ± 0.0039
τ	0.0547 ± 0.0077	0.0574 ± 0.0079	0.0596 ± 0.0077	0.0597 ± 0.0078
Σm_ν [eV]	< 0.31	< 0.14	< 0.13	< 0.13
r	< 0.19	< 0.21	< 0.20	< 0.20
$\Omega_{\rm eDE}$	< 0.0063	< 0.0039	< 0.0033	< 0.0032
w_0	< -0.72	< -0.95	< -0.95	< -0.96
H_0 [km/s/Mpc]	63.13 ± 2.60	67.18 ± 0.58	67.16 ± 0.55	67.18 ± 0.55
σ_8	0.7632 ± 0.0294	0.7954 ± 0.0109	0.7999 ± 0.0085	0.8000 ± 0.0085

4.3 Including high redshift expansion tracers

We will now turn to the question of how much the EDE constraints obtained in the previous Section are affected by the inclusions of high redshift data tracing the cosmological expansion. Here we consider the effect of the Hubble diagram of QSOs which we discussed in Sec. 2.3.2. As shown in the last column of Table 4.3 there are no significant differences in the 2σ upper limits and also the mean value of the parameters, when QSOs are included. In Fig. 4.10 we compare the constraints on cosmological parameters using Planck2018+Lensing+BAO+SNe, with and without the QSO data set. We include the QSO data as well as the other high redshift tracer that we are using in all our analysis, the Lyman-α BAO data. As already mentioned, new BAO data at $z = 2.34$ were obtained from the auto-correlation of Lyman-α forest absorption in eBOSS Data Release 14 [215], as well as from the cross-correlation with quasars in eBOSS DR14 at $z_{\rm eff} = 2.35$ [216]. Therefore present high-redshift tracers of the cosmological expansion do not improve significantly the constraints on the EDE parameters of the EDE model. Similarly, no significant impact is observed for the other parameters. The rationale of including QSOs is that they are a very high redshift probe, and they are useful in particular in addressing their impact on massive neutrinos. Moreover, recent papers found that this quasar dataset prefers a Universe with no dark energy [239] (see also [240] where the same conclusion is reached in a model-independent way).
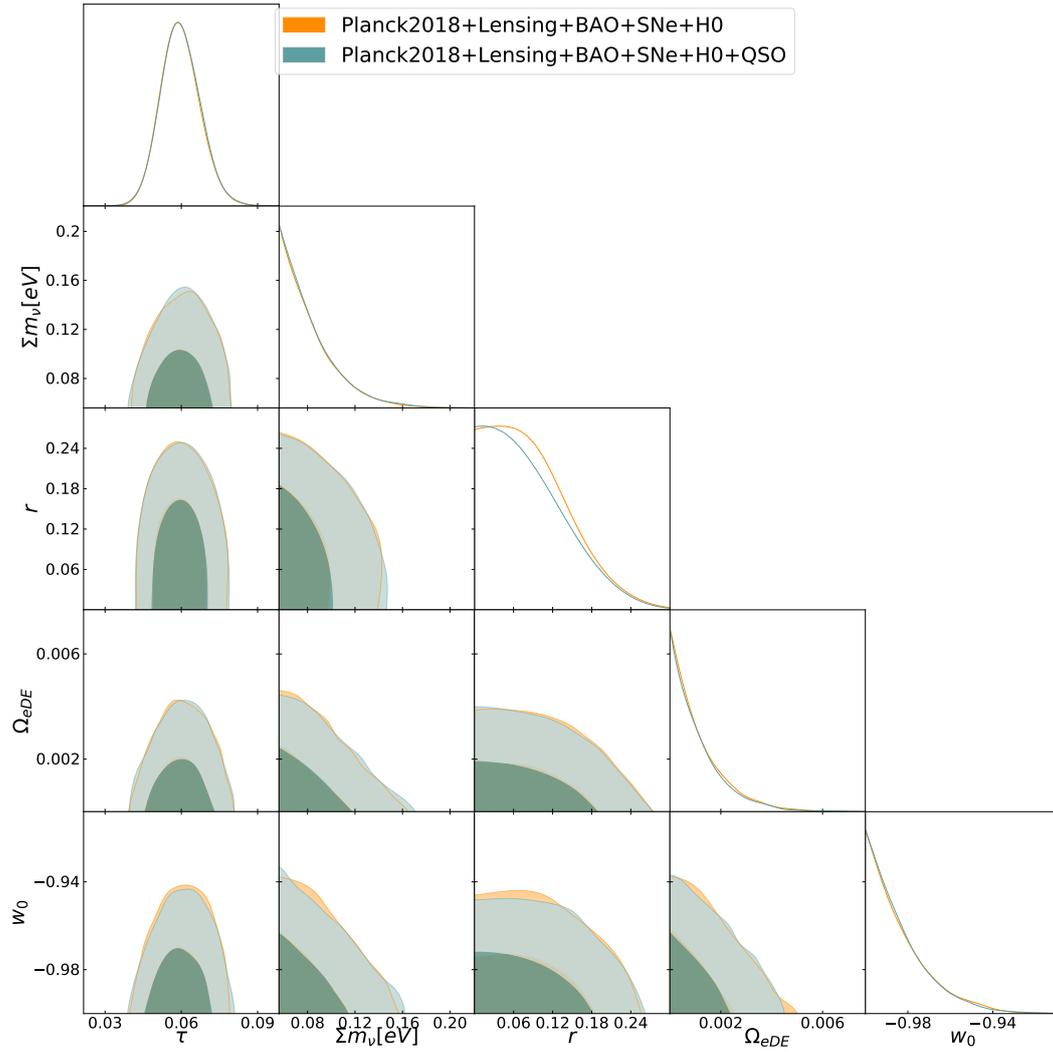
Figure 4.10: Comparison between cosmological constraints within EDE models with and without the QSO data, see the last column of the Table 4.3

4.4 Cosmological constraints from EDE to wCDM

We conclude our analysis by progressively simplifying our EDE models from the general ones to the simple constant equation of state of dark energy, which we refer to as wCDM. We make use of priors, listed in Table 4.4, while the ones in Table 4.2 are still adopted for the non-EDE cosmological parameters.

In the first EDE model, $\Omega_{\rm eDE}$ and w_0 are allowed to vary, as well as $c_{\rm eff}^2$ and $c_{\rm vis}^2$. The constraints are shown in Fig. 4.11 and listed in the EDE column of Table 4.5. In the second case, named "EDE fixed $c_{\rm eff}^2$ & $c_{\rm vis}^2$", perturbations parameters are set to the values $c_{\rm eff}^2 = 1$ and $c_{\rm vis}^2 = 0$, meaning that there is no anisotropic stress in the DE. The constraints are shown in Fig. 4.11 and Table 4.5 - "EDE fixed $c_{\rm eff}^2$ & $c_{\rm vis}^2$" column. In a third case, we let no EDE, i.e. $\Omega_{\rm eDE} = 0$, while the equation of state today varies as well as the effective and viscosity sound speed. The corresponding constraints are shown in Fig. 4.12 and the fourth column of Table 4.5, named wCDM. Finally, in the wCDM model, only the constant equation of state is allowed to vary, in the range $[-1, 0]$. Results are shown in Fig. 4.12 and the last column of Table 4.5.

Table 4.4: Priors for "EDE","EDE fixed $c_{\rm eff}^2$ & $c_{\rm vis}^2$", "wCDM varying $c_{\rm eff}^2$ & $c_{\rm vis}^2$" and "wCDM" models.

Parameter	EDE	EDE fixed $c_{\rm eff}^2$ & $c_{\rm vis}^2$	wCDM varying $c_{\rm eff}^2$ & $c_{\rm vis}^2$	wCDM
$\Omega_{\rm eDE}$	$[0, 0.1]$	$[0, 0.1]$	0	0
w_0	$[-1, 0]$	$[-1, 0]$	$[-1, 0]$	$[-1, 0]$
$c_{\rm eff}^2$	$[0, 1]$	1	$[0, 1]$	1
$c_{\rm vis}^2$	$[0, 1]$	0	$[0, 1]$	0

Table 4.5: Cosmological constraints for Planck2018+BAO+SNe+H0 prior for our 4 different possible scenarios listed in Table 4.4 .

Parameter	EDE	EDE fixed $c_{\rm eff}^2$ & $c_{\rm vis}^2$	wCDM varying $c_{\rm eff}^2$ & $c_{\rm vis}^2$	wCDM
$\Omega_b h^2$	0.02246 ± 0.00013	0.02246 ± 0.00014	0.02246 ± 0.00014	0.02246 ± 0.00014
$\Omega_c h^2$	0.1187 ± 0.0010	0.1187 ± 0.0011	0.1186 ± 0.0011	0.1186 ± 0.0011
$100\theta_{MC}$	1.04105 ± 0.00030	1.04102 ± 0.00030	1.04107 ± 0.00029	1.04107 ± 0.00029
$\ln(10^{10} A_s)$	3.048 ± 0.016	3.047 ± 0.016	3.047 ± 0.016	3.047 ± 0.017
n_s	0.9686 ± 0.0040	0.9684 ± 0.0040	0.9686 ± 0.0041	0.9685 ± 0.0041
τ	0.0574 ± 0.0079	0.0572 ± 0.0080	0.0570 ± 0.0079	0.0571 ± 0.0080
Σm_ν [eV]	< 0.14	< 0.14	< 0.14	< 0.14
r	< 0.21	< 0.20	< 0.21	< 0.21
$\Omega_{\rm eDE}$	< 0.0039	< 0.0034	0	0
w_0	< -0.95	< -0.95	< -0.95	< -0.95
H_0 [km/s/Mpc]	67.18 ± 0.58	67.16 ± 0.58	67.20 ± 0.59	67.22 ± 0.59
σ_8	0.7954 ± 0.0109	0.7952 ± 0.0109	0.7983 ± 0.0107	0.7985 ± 0.0106

By looking at results, we derive the following main conclusions. $\Omega_{\rm eDE}$ varies between 0.0034 to 0.0039 (2σ upper limit), with the tightest limit obtained in the absence of perturbations. The upper limits on the total amount of EDE tend to decrease with an increasing neutrino mass, while with a decreasing value of the tensor to scalar ratio we expect the same limits tend to become less tight.

The parameters describing the DE perturbations are unconstrained and with no significant degeneracy; including them in the analysis does not have any impact on the constraints for the wCDM scenario. The bound on the neutrino mass summation is not affected by any of the model extensions shown in Table 4.5. Finally, with respect to the Planck 2018 fit within ΛCDM the bounds on the tensor to scalar ratio are relaxed by a factor 2.

Finally, given that the perturbation parameters are unconstrained, we focus on the case "EDE fixed $c_{\rm eff}^2$ & $c_{\rm vis}^2$", and we check whether the varying the neutrino mass and/or the tensor to scalar ratio has any impact on the EDE parameters. The results are shown in Table 4.6. The limits on $\Omega_{\rm eDE}$ are slightly relaxed any time either Σm_ν or r or both are kept fixed. On the other hand, the bounds on w_0 are slightly more loose only if the neutrino mass summation is fixed. Finally, it is interesting to notice that the upper limit on the tensor to scalar ratio in our EDE scenario is quite stable whether the neutrino mass is varying or not. This indicates that the factor 2 in the upper limit on r with respect to ΛCDM is really induced by EDE.
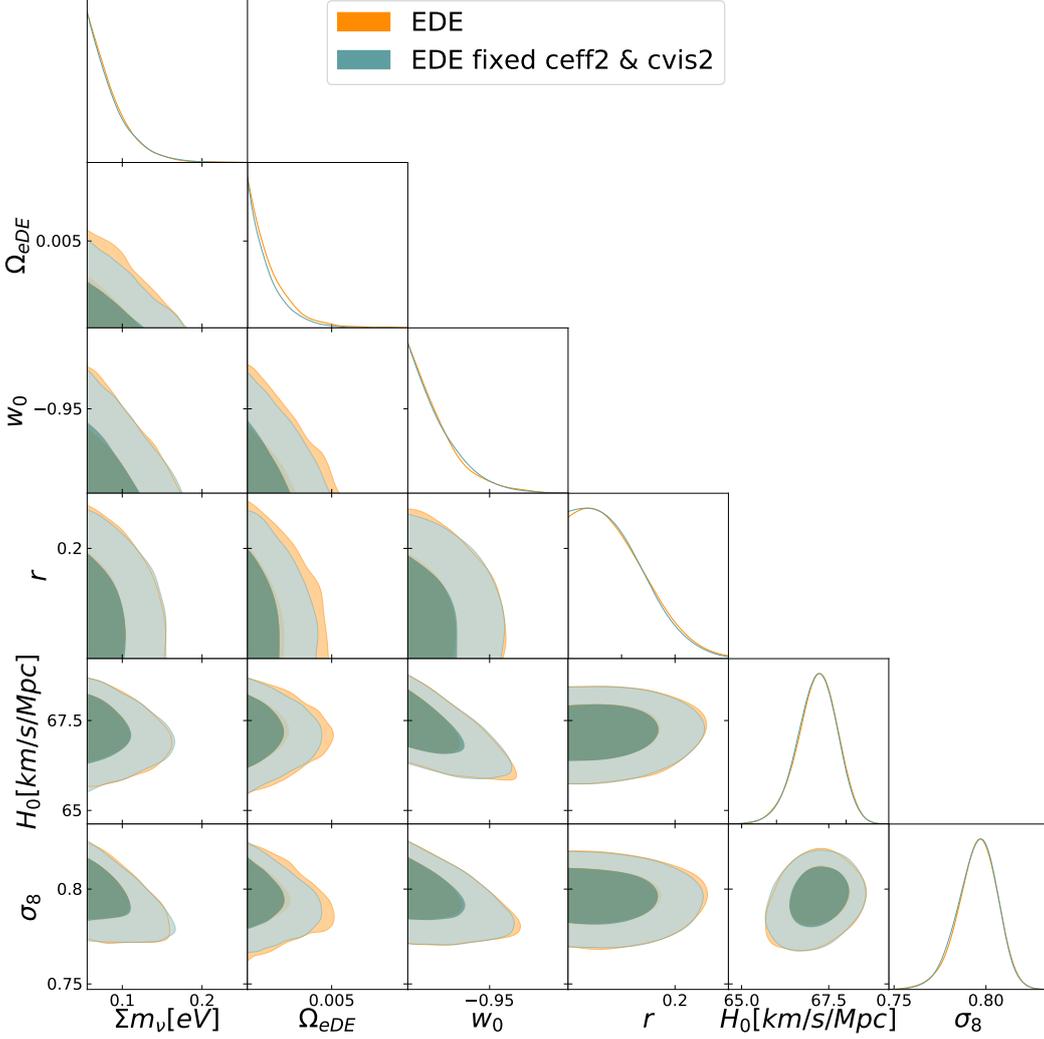
Figure 4.11: Constraints for EDE and EDE fixed c_{eff}^2 & c_{vis}^2 scenarios. Contours show the 68% and 95% confidence level regions allowed from Planck+BAO+SNe+H_0 measurements.

Table 4.6: Cosmological constraints for Planck2018+BAO+SNe+H0 prior for the case "EDE fixed c_{eff}^2 & c_{vis}^2" (second column of Table 4.5, corresponding here to the first column), fixing either Σm_ν (second column), or r (third column), or both (fourth column).

Parameter	EDE fixed c_{eff}^2 & c_{vis}^2	fixed $\Sigma m_\nu = 0.056$	fixed $r = 0$	fixed $\Sigma m_\nu = 0.056$ and $r = 0$
$\Omega_b h^2$	0.02246 ± 0.00014	0.02245 ± 0.00014	0.02246 ± 0.00013	0.02246 ± 0.00013
$\Omega_c h^2$	0.1187 ± 0.0011	0.1189 ± 0.0010	0.1189 ± 0.0010	0.1190 ± 0.0010
$100\theta_{MC}$	1.04102 ± 0.00030	1.04104 ± 0.00029	1.04099 ± 0.00031	1.04103 ± 0.00029
$\ln(10^{10} A_s)$	3.047 ± 0.016	3.047 ± 0.016	3.049 ± 0.016	3.048 ± 0.016
n_s	0.9684 ± 0.0040	0.9683 ± 0.0039	0.9669 ± 0.0040	0.9666 ± 0.0039
τ	0.0572 ± 0.0080	0.0570 ± 0.0079	0.0578 ± 0.0078	0.0571 ± 0.0079
Σm_ν [eV]	< 0.1389	$-$	< 0.1371	$-$
r	< 0.203	< 0.210	$-$	$-$
Ω_{eDE}	< 0.0034	< 0.0035	< 0.0036	< 0.0035
w_0	< -0.95	< -0.94	< -0.95	< -0.95
H_0 [km/s/Mpc]	67.16 ± 0.58	67.23 ± 0.61	67.13 ± 0.58	67.27 ± 0.56
σ_8	0.7952 ± 0.0109	0.8003 ± 0.0095	0.7966 ± 0.0110	0.8009 ± 0.0095
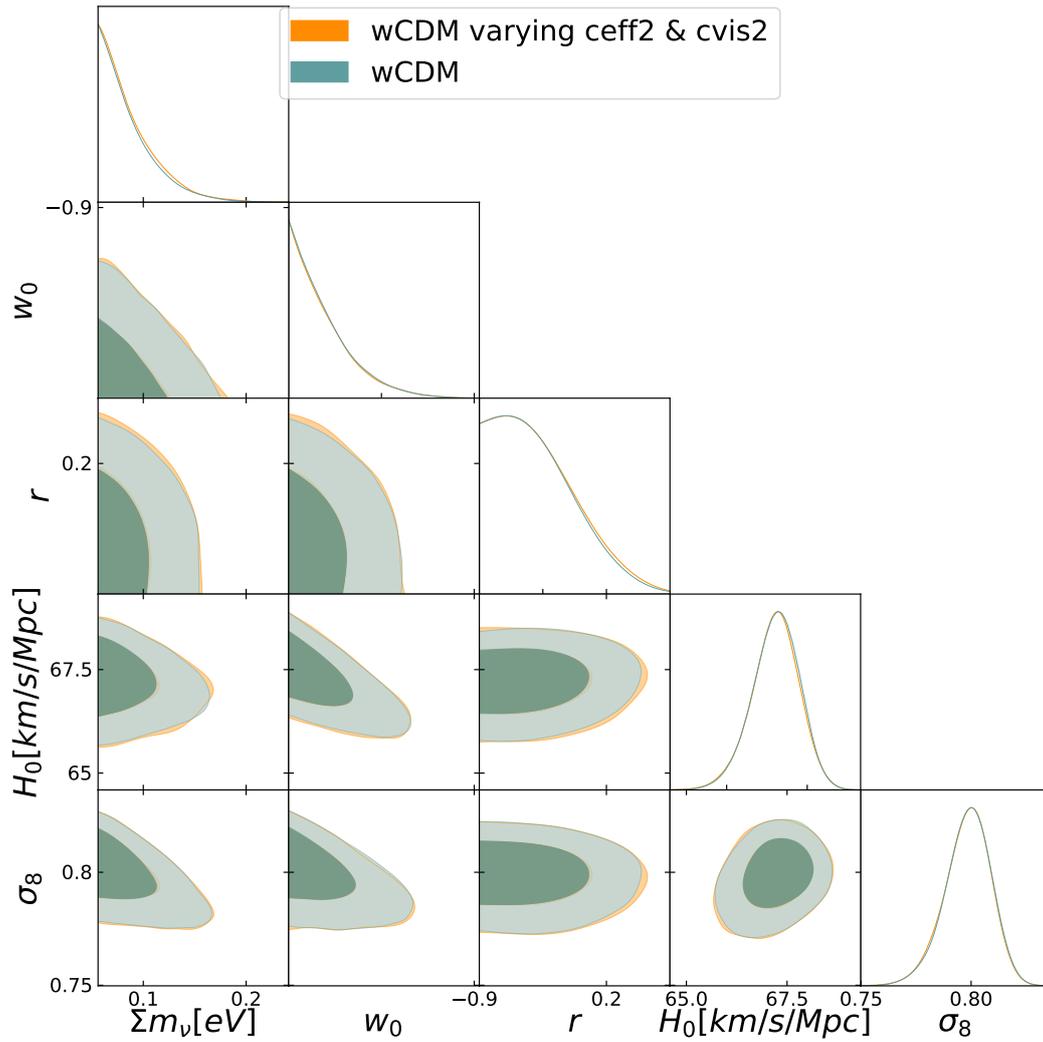
Figure 4.12: "wCDM varying c_{eff}^2 & c_{vis}^2" and "wCDM" models. Contours show the 68% and 95% confidence level regions allowed from Planck+BAO+SNe+H_0 measurements.

<div style="text-align: right; font-size: 3em;">**5**</div>

Quantifying the impact of baryon-CDM perturbations on halo clustering and baryon fraction

Baryons and cold dark matter (CDM) did not comove prior to recombination. This leads to differences in the local baryon and CDM densities, the so-called baryon-CDM perturbations δ_{bc}. These perturbations are usually neglected in the analysis of Large-Scale Structure data but taking them into account might become important in the era of high precision cosmology. Using gravity-only 2-fluid simulations we assess the impact of such perturbations on the dark matter halos distribution. In particular, we focus on the baryon fraction in halos as a function of mass and large-scale δ_{bc}, which also allows us to study details of the nontrivial numerical setup required for such simulations. We further measure the cross-power spectrum between the halo field and δ_{bc} over a wide range of mass. This cross-correlation is nonzero and negative which shows that halo formation is impacted by δ_{bc}. We measure the associated bias parameter $b_{\delta_{bc}}$ and compare it to recent results, finding good agreement. Finally we quantify the impact of such perturbations on the halo-halo power spectrum and show that this effect can be degenerate with the one of massive neutrinos for surveys like DESI[1][241].

This Chapter is organized as follows: In Sec. 1 we give a brief overview of baryon-CDM perturbations (Sec. 1.1) and how to measure $b_{\delta_{bc}}$ (Sec. 1.2). We then turn to a detailed description of our simulations in Sec. 2. We start by describing how we generate initial conditions in Sec. 2.1 and we give a few details of the simulations in Sec. 2.2. We then turn to various numerical tests to validate our setup (Sec. 2.3), and we describe the halo finding procedure in Sec. 2.4. We present our results in Sec. 3, focusing first on the baryon fraction in 3.1, and halo bias and power spectra in Sec. 3.2. At the end of this chapter we present a rapid overview of the separate Universe technique used in [242] to measure $b_{\delta_{bc}}$.

1 Theory

In this section we give a theoretical overview of baryon-CDM perturbations (Sec. 1.1), then in Sec. 1.2, we explain how to measure the associated bias, $b_{\delta_{bc}}$.

1.1 Baryon-CDM pertubations

In this section we summarize how baryon-CDM perturbations are generated in the early Universe, and we review the formalism to derive their evolution. All this was already discussed in details in [11, 242] so we stay concise and refer the interested reader to these papers.

We restrict ourselves to linear perturbation theory since we will focus on (very) large scales in this work.

We start by writing the Euler and Continuity equations (which we talked about them in chapter 3, Sec. 3.1) for the CDM and baryon components after decoupling.

---

[1]https://www.desi.lbl.gov

$$\frac{\partial}{\partial \tau} \delta_s(\mathbf{x}, z) = -\theta_s(\mathbf{x}, z), \quad s \in \{b, c\}$$

$$\frac{\partial}{\partial \tau} \theta_s(\mathbf{x}, z) + \mathcal{H}\theta_s(\mathbf{x}, z) = -\frac{3}{2}\Omega_m(a)\mathcal{H}^2\delta_m(\mathbf{x}, z) \tag{5.1}$$

Where $\delta_s \equiv \delta\rho_s/\bar{\rho}_s$ is the fractional density perturbations for different species (baryon and CDM denoted by b and c) and $\theta_s = \partial_j v_s^j$ is the fluid velocity divergence of the species. τ here denotes the conformal time so we have $dt = a d\tau$ and also $\mathcal{H} = d\ln a/d\tau$ shows the conformal Hubble rate.

Considering $\delta_m = (\Omega_b \delta_b + \Omega_c \delta_c)/\Omega_m$ and $f_b = \Omega_b/\Omega_m$, we express these in terms of the total and relative density perturbations

$$\delta_m = f_b \delta_b + (1 - f_b)\delta_c, \quad \delta_r = \delta_b - \delta_c, \tag{5.2}$$

Plugging baryon and CDM components in Eqs. 5.1 we have

$$\begin{cases} \dfrac{\partial}{\partial \tau}\delta_b = -\theta_b, \\ \dfrac{\partial}{\partial \tau}\theta_b + \mathcal{H}\theta_b = -\dfrac{3}{2}\Omega_m(a)\mathcal{H}^2\delta_m, \end{cases} \tag{5.3}$$

$$\begin{cases} \dfrac{\partial}{\partial \tau}\delta_c = -\theta_c, \\ \dfrac{\partial}{\partial \tau}\theta_c + \mathcal{H}\theta_c = -\dfrac{3}{2}\Omega_m(a)\mathcal{H}^2\delta_m, \end{cases} \tag{5.4}$$

Now for each component we plug θ_s from the first equation in to the second one:

$$\frac{\partial}{\partial \tau}\left(-\frac{\partial}{\partial \tau}\delta_b\right) + \mathcal{H}\left(-\frac{\partial}{\partial \tau}\delta_b\right) = -\frac{3}{2}\Omega_m(a)\mathcal{H}^2\delta_m. \tag{5.5}$$

$$\frac{\partial}{\partial \tau}\left(-\frac{\partial}{\partial \tau}\delta_c\right) + \mathcal{H}\left(-\frac{\partial}{\partial \tau}\delta_c\right) = -\frac{3}{2}\Omega_m(a)\mathcal{H}^2\delta_m. \tag{5.6}$$

Multiplying Eq. 5.5 by f_b and Eq. 5.6 by $(1 - f_b)$ and adding them together/ (Subtracting Eqs. 5.5 and 5.6, using the definition $\delta_r = \delta_b - \delta_c$) we have the first/ (second) following equations respectively:

$$\frac{\partial^2}{\partial \tau^2}\delta_m + \mathcal{H}\frac{\partial}{\partial \tau}\delta_m - \frac{3}{2}\Omega_m(a)\mathcal{H}^2\delta_m = 0,$$

$$\frac{\partial^2}{\partial \tau^2}\delta_r + \mathcal{H}\frac{\partial}{\partial \tau}\delta_r = 0. \tag{5.7}$$

As shown in [11] these two equations in term of these variables admit the following solutions

$$\delta_m(\tau) = A_+ D_+(\tau) + A_- H(\tau),$$

$$\delta_r(\tau) = R_+ + R_- D_r(\tau), \tag{5.8}$$

where $A_\pm$, $R_\pm$ are constants, $D_+(\tau)$ is the usual linear matter growth rate, and $D_r(\tau)$ can be approximated to $-2a^{-1/2}(\tau)$ during matter domination. Apart from the usual growing and decaying modes of δ_m, we are

interested in the two modes of δ_r. The first one is a constant mode δ_{bc}[2] of compensated perturbations with $\delta_m = 0$, i.e. $\delta\rho_b = -\delta\rho_c \Leftrightarrow f_b\delta_b = -(1 - f_b)\delta_c$, but $\delta_r \neq 0$. The second mode is a decaying one that can be shown to be related to the divergence of peculiar velocity perturbations between the two fluids, $\theta_{bc} = \theta_b - \theta_c$ (see [11]). We hence get

$$\delta_r(\mathbf{x}, \tau) = \delta_{bc}(\mathbf{x}) + \frac{\theta_{bc}(\mathbf{x}, z = 0)}{H_0}D_r(\tau). \tag{5.9}$$

In the scope of effective-field theory, the overdensity of tracers δ_h, such as dark matter halos, is expressed in terms of operators O constructed out of the *total* matter density field δ_m and tidal field K_{ij}, multiplied by numerical coefficients b_O, the bias parameters (see [97] for a very complete review)

$$\delta_h(\mathbf{x}, \tau) = \sum_O b_O(\tau)O(\mathbf{x}, \tau). \tag{5.10}$$

These two new perturbation terms (appeared in Eq. 5.9), must enter the bias expansion (Eq. 5.10) at linear order already with new associated bias parameters. They should normally be evaluated at the Lagrangian position $\mathbf{q}(\mathbf{x})$ (described in chapter 3, Sec. 3.1) corresponding to Eulerian position $\mathbf{x}$. However at linear order we can neglect this which allows us write the fractional halo density perturbation at linear order as

$$\delta_h(\mathbf{x}, \tau) = b_1(\tau)\delta_m(\mathbf{x}, \tau) + b_{\delta_{bc}}(\tau)\delta_{bc}(\mathbf{x}) + b_{\theta_{bc}}(\tau)\theta_{bc}(\mathbf{x}, \tau). \tag{5.11}$$

1.2   Measuring $b_{\delta_{bc}}$

We now turn to expressions for the halo cross-power spectra P_{hm} and P_{hbc}. As argued in [11, 242] the two last terms in 5.11 should be much smaller than the first one, which is why they are normally neglected. Furthermore, the last term proportional to θ_{bc} is expected to be much smaller than the second one proportional to δ_{bc}. We will hence neglect the velocity divergence term in what follows.

The cross-power spectrum P_{hbc} is defined as

$$(2\pi)^3\delta_D(\mathbf{k} + \mathbf{k}')P_{hbc}(k) = \langle\delta_h(\mathbf{k})\delta_{bc}(\mathbf{k}')\rangle, \tag{5.12}$$

and similarly for P_{hm}. Plugging 5.11 into 5.12 we get

$$P_{hm}(k) = b_1 P_{mm}(k) + b_{\delta_{bc}}P_{mbc}(k),$$
$$P_{hbc}(k) = b_1 P_{mbc}(k) + b_{\delta_{bc}}P_{bcbc}(k), \tag{5.13}$$

where P_{mm} is the usual matter power spectrum, while P_{mbc} and P_{bcbc} are the cross- and auto-power spectra of δ_{bc} with the matter field and itself respectively. Notice that if we neglect the last term in the second line, the ratio P_{hbc}/P_{mbc} should go to b_1. From these two equations we can get expressions for b_1 and $b_{\delta_{bc}}$ in the low k limit

$$b_1 = \lim_{k\to 0}\frac{P_{hm}(k) - b_{\delta_{bc}}P_{mbc}(k)}{P_{mm}(k)},$$

$$b_{\delta_{bc}} = \lim_{k\to 0}\frac{P_{hbc}(k)P_{mm}(k) - P_{hm}(k)P_{mbc}(k)}{P_{bcbc}(k)P_{mm}(k) - P_{mbc}^2(k)}. \tag{5.14}$$

We can hence get $b_{\delta_{bc}}$ from the second of these two equations and insert it in the first one to get b_1. It is very clear that the term proportional to $b_{\delta_{bc}}$ in the first line of 5.14 represents the deviation of b_1 from the traditional ratio P_{hm}/P_{mm} due to baryonic effects in 2-fluids simulations.

---

[2]We use the subscript m for total matter and reserve bc for the difference $b - c$.

Actually, a more practical way to obtain $b_{\delta_{bc}}$ with more constraining power is to obtain b_1 by the usual ratio P_{hm}/P_{mm} at low k in 1-fluid simulations, and to subtract $b_1 P_{mm}$ from the first line of 5.13 in order to detect any deviation in the usual relation in 2-fluids simulations, i.e

$$b_{\delta_{bc}} = \lim_{k \to 0} \frac{P_{hm}(k) - b_1^{1\mathrm{f}} P_{mm}(k)}{P_{mbc}(k)}, \tag{5.15}$$

where we have used the superscript "1f" to denote the linear bias as measured in 1-fluid simulations, i.e.

$$b_1^{1\mathrm{f}} = \lim_{k \to 0} \frac{P_{hm}^{1\mathrm{f}}(k)}{P_{mm}^{1\mathrm{f}}(k)}. \tag{5.16}$$

Finally we could do the same reasoning adding θ_{bc}. However we expect this term to be subdominant since it is a decaying one, and adding it would probably only diminish our constraining power on $b_{\delta_{bc}}$. We will however compare our results to ones obtained with the 1-fluid "separate Universe simulations" technique of [242] as outlined in Sec. 3.3. Since our 2-fluid simulations include by default θ_{bc}, which is not the case of the 1-fluid ones, any difference between the two measurements can be attributed to our neglection of θ_{bc} and would hence be an estimation of the magnitude of this term.

2   Simulations and halo finding

We now turn to a detailed description of our set of simulations. As was discussed in [64, 65, 10], obtaining the correct evolution of each species in 2-fluid simulations, even at linear order only, is already a nontrivial task which is why we thoroughly present a number of tests to validate our setup in this section. We present our tests results in 2.3. (for a basic explanation of the idea of simulation see chapter 3, Sec. 3.2)

Our fiducial cosmology is consistent with the Planck 2018 one [2], detailed as follows: $\Omega_m = 0.3111$, $\Omega_b = 0.0490$, $\Omega_c = 0.2621$, $\Omega_\Lambda = 0.6889$, $n_s = 0.9665$, $\sigma_8 = 0.8261$ and $h = 0.6766$. The box size is $L_{box} = 250\ h^{-1}\,\mathrm{Mpc}$ on each side for all our simulations.

We run two sets of simulations. The first one is a standard gravity-only one with one species of particles that we refer to as "1-fluid" (we explained this common approximation in N-body simulations in the second paragraph of chapter 3, Sec. 3). In addition to the fiducial cosmology we run two additional cosmologies with enhanced ("High") and lowered ("Low") Ω_b (while adapting Ω_c to keep Ω_m fixed) in order to compute $b_{\delta_{bc}}$ in the same fashion as [242]. We also use the fiducial simulation of this set to compute the linear bias b_1. Explicitly speaking, these simulations have only CDM particles, and we use 512^3 mass elements. The way we run this set of "separate Universe simulations of baryon-CDM perturbations" is identical to what was done in [242] and we refer the reader to their paper for more details.

The second set of simulations contains two distinct fluids representing baryons and CDM each with different primordial density and velocity fluctuations. Explicitly, we use two different transfer functions to initialize the two fluids that we then evolve jointly. Each fluid consist of 512^3 mass elements. We refer to this second set as "2-fluid". Furthermore we ran a hybrid version of the 2-fluid simulation where the two fluids are initialized with the same transfer functions, in order to check our numerical setup. We insist that we do not include any hydrodynamical effects for baryon evolution, and that all our simulations are gravity-only. Tab. 6.1 summarizes the varying parameters of our simulations. In the following sections we go into the details of our numerical setup for 2-fluid simulations and present some sanity checks.

2.1   Initial Conditions

We generate the initial conditions (initial position and velocities) for the particles in our simulations at an initial redshift $z_i = 49$ using the publicly available MUSIC code [243]. (for a basic idea of generating initial conditions see chapter 3, Sec. 3.2 and Fig. 3.10)

The 1-fluid simulations are initialized in the standard way by computing the primordial matter power spectrum using the CAMB code [4] at $z = 0$, and back-scaling it to the initial redshift assuming growing mode only for the specified cosmology. On the other hand, in the case of 2-fluid simulations we compute the transfer

Name	N_b	N_c	m_b	m_c	Ω_c	Ω_b	N_{real}	TFs
1-fluid Fid	0	512^3	–	1.0051	0.2621	0.049	16	–
1-fluid High	0	512^3	–	1.0051	0.2596	0.0515	16	–
1-fluid Low	0	512^3	–	1.0051	0.2645	0.0466	16	–
2-fluid-diff	512^3	512^3	0.1583	0.8468	0.2621	0.049	4	2
2-fluid-same	512^3	512^3	0.1583	0.8468	0.2621	0.049	4	1

Table 5.1: Summary of our sets of simulations. All simulations are gravity-only and have a box size of $250\,h^{-1}\,\text{Mpc}$ on each side. N_b and N_c are the number of baryonic and CDM particles respectively. m_b and m_c are the corresponding mass in units of $10^{10}M_\odot/h$. N_{real} corresponds to the number of realizations we ran of each simulations in order to build statistics, and "TFs" refers to the number of transfer functions used to initialize the two fluids (1 means that they are initialized with the same transfer function corresponding to the weighted total matter one).

functions for baryons and CDM from CAMB directly at $z = 49$. Notice that the total matter power spectrum is the same for both 1-fluid and 2-fluid simulations.

We then compute the displacement and velocity fields using the Zel'dovich approximation [100] (for a brief explanation of the Lagrangian perturbation theory approach see chapter 3, Sec. 3.1) (for simplicity we are not using the second-order Lagrangian Perturbation Theory (LPT) formalism here; notice however that it was recently figured out for two fluids in [67]). Furthermore we used the fixed mode amplitude technique incorporated in MUSIC of [244], in which the modulus of the white noise Fourier modes is set to unity in order to suppress the impact of cosmic variance. (for a short explanation of cosmic variance, see chapter 2, Sec. 5.1, Eq. 2.58)

2.2 N-body simulations details

Our simulations are performed with the cosmological code Gadget-2 [245] with a numerical setup very similar to the one of [64]. As stated before, we compute only gravitational interactions and neglect all hydrodynamical effects, implying that baryons behave like a collisionless fluid. This is because we are only interested in investigating the effect of baryon-CDM perturbations generated in the early Universe.

As discussed in detail in [64], one of the issues that needs to be addressed in these kind of collisionless simulations is the force resolution for the light fluid. Indeed a too high force resolution for the mass resolution could cause a spurious coupling between CDM and baryons affecting their clustering features and the growth of structures on all scales. A simple solution to remedy that is to make the baryon smoothing length unusually high. In fact, as shown in [64], the force softening must be of the order of the mean baryon inter-particle distance in order to recover the correct linear evolution. However, this is a problem since this length can become of the order of $1\,h^{-1}\,\text{Mpc}$ and small halos can have a final radius smaller than this implying that we would not simulate correctly structure formation at the small mass end of the mass function. (for a brief discussion on the softening length see chapter 3, Sec. 3.2 and Eq. 3.72)

Another solution discussed in [64], that we use in this work, is to use the adaptive gravitational softening (AGS) *for baryons only* [68], implemented in Gadget. This technique allows the softening length to vary in space and time according to the density of the environment. Specifically, in our case the force acting on baryonic particles is softened adaptively using an SPH kernel with a size set by the 28 closest neighbours (DesNumNgb=28 in Gadget). Furthermore we set a floor for the minimum softening length $\epsilon = 12.5h^{-1}\text{kpc}$ corresponding to roughly 1/40-th of the mean inter-particle separation of baryons. The CDM softening length is kept constant through space and time to $\epsilon = 12.5h^{-1}\text{kpc}$, also corresponding to 1/40-th of the mean CDM inter-particle separation. We present validating tests of this setup in the next section as well as the effect of varying specific details of the force softening in Sec. 2.3.

Finally let us note that recently several papers tackled the issue of the spurious coupling between the light and heavy particles without introducing a large softening length. Ref. [65] claimed that this can be done by using a Lagrangian glass for the baryon particles. The recent papers [10] and [67] formally generalized LPT to an arbitrary order n and use variations in particle masses to resolve the spurious deviations from expected perturbative results in baryon-CDM simulations. While we do not attempt to compare rigorously our setup with theirs, we compare the results for the baryon fraction F_b when using our one or the one of [10] in Sec. 3.1.
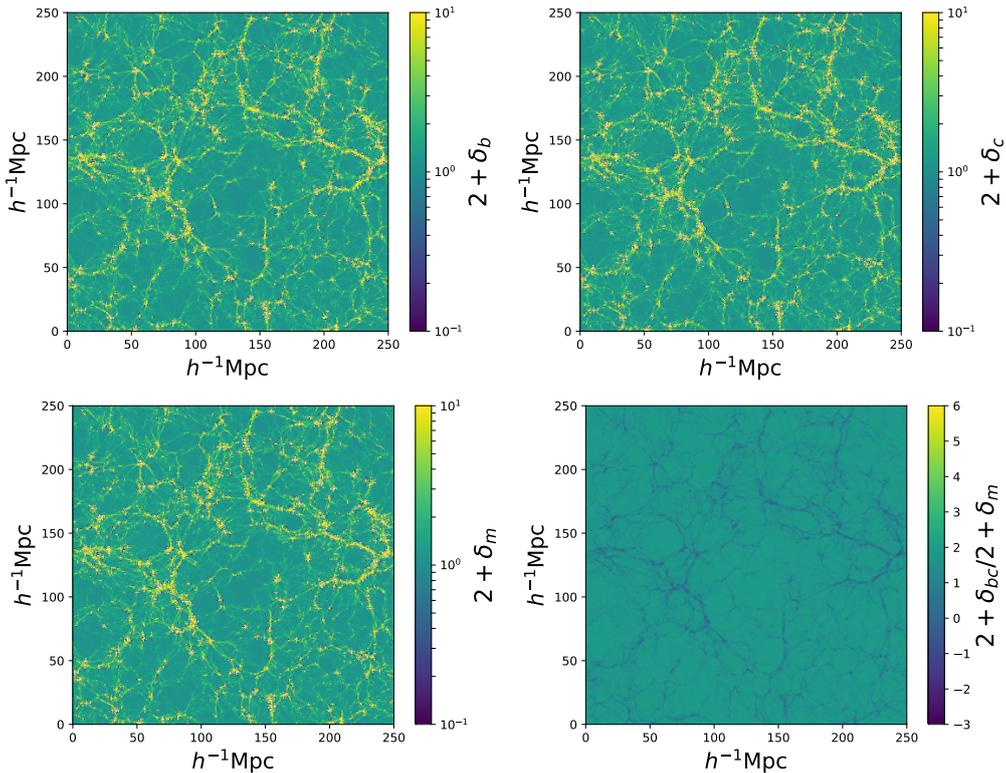
Figure 5.1: Maps of density fields in a slice of thickness 10 h^{-1} Mpc. The top left panel shows the baryon density field while the top right panel shows the CDM one. In the bottom left panel we present the total matter density defined as in Eq. 5.2. The bottom right one shows the relative baryon-CDM field denoted as $\delta_{bc} = \delta_b - \delta_c$ (if $\theta_{bc} \to 0$), normalized to δ_m. We show this ratio to allow for better visualization. We see that baryons do indeed trace the CDM fluid but with a small lag which makes δ_{bc} negative in high δ_m regions, while it is positive in low density ones.

2.3 Numerical tests

In this section, we present numerical tests to validate our 2-fluid simulations. All the tests presented here concern the 2-fluid-diff-TF set of simulations.

We start with the maps of the density fields in a slice of thickness $10\,h^{-1}$Mpc at $z = 0$ presented on Fig. 5.1. The top left panel represents the baryon density filed while the top right panel shows the CDM density field. Here one can see how baryons are following the CDM. In bottom the left panel we present the total matter density defined as $\delta_m = f_b \delta_b + (1 - f_b)\delta_c$. The bottom right panel shows the relative baryon-CDM field denoted as $\delta_{bc} = \delta_b - \delta_c$ (if $\theta_{bc} \to 0$), normalized to δ_m (to allow for better visualization). This figure gives us a visual validation of our simulations by confirming the presence of the usual structures in the cosmic web, the fact that baryons closely trace CDM, and allows us to visualize the relative baryon-CDM field showing that it is smaller in amplitude than the total one and has a negative value in average. Furthermore we see that baryons trace CDM with a small lag which makes δ_{bc} negative in high δ_m regions, while it is positive in low density ones.

We now turn to comparing the measured power spectra of the density fields shown on Fig. 5.1 with linear theory (for an explanation of the linear matter power spectra see: chapter 2, Sec. 6, Eq. 2.70 and Fig. 2.7) at different redshifts. The measurements were performed by mapping the particle distribution using a cloud-in-cell (CIC) scheme on to a 1024^3 grid and then Fast Fourier Transforming the field. (to have a brief explanation of the mass assignments, see chapter 3, Sec. 3.2) The top panel of Fig. 5.2 compares the measured growth of the total matter power spectrum ("+" signs) with the prediction of the linear perturbation theory, shown as solid lines, for different redshift. We can see that our simulations reproduces the expected linear growth from $z_i = 49$ up to $z = 0$, up to $k \sim 0.3\,h\,\mathrm{Mpc}^{-1}$ at $z = 0$. On small scales and at low redshift, the nonlinear growth of structures dominates which is why the measured power spectrum becomes higher than the linear one. This agreement supports the correctness of our numerical calculations. The two middle panels show the ratio between the baryon and CDM power spectra at various redshifts. Again measurements are represented by "+" while the solid lines show the linear prediction. The fact that the solid curves are systematically different from
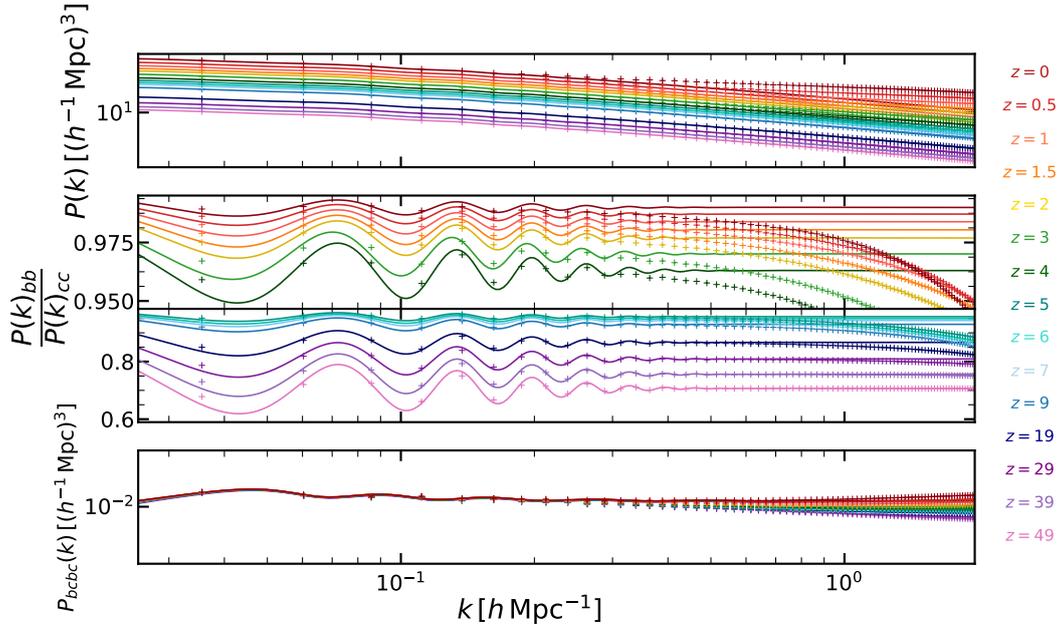
Figure 5.2: Power spectra as a function of wavenumber k for 14 different redshifts indicated by the color coding. Solid lines show the prediction of the linear perturbation theory and plus ("+") markers represent the results from simulations. The top panel presents the growth of the total matter field, the two middle ones show the ratio between the measured baryon to CDM power spectra, and the bottom one presents results for the relative perturbation δ_{bc} auto-power spectrum. We see that measurements agree with linear theory up to $k \sim 0.3\,h\,\mathrm{Mpc}^{-1}$ down to $z = 0$, as is expected. On the middle panels we see the BAO wiggles due to the fact that they are present only in the baryon power spectrum but not in the CDM one. We see that the difference between the baryon and CDM power spectra becomes of the order of 1% at $z = 0$. This is the impact of this difference on dark matter halos that we want to study in this work. The clear suppression of baryon perturbations compared to the CDM ones on small scales is due to our treatment of the force softening with AGS for baryons. We see no redshift dependency of P_{bcbc} on large scales on the fourth panel as is expected since the relative density δ_r can quickly be approximated by the constant mode δ_{bc} as $\theta_{bc} \to 0$ as explained in section Sec. 1.1. The good agreement between our measurements and linear theory for all curves validates our numerical setup.

76
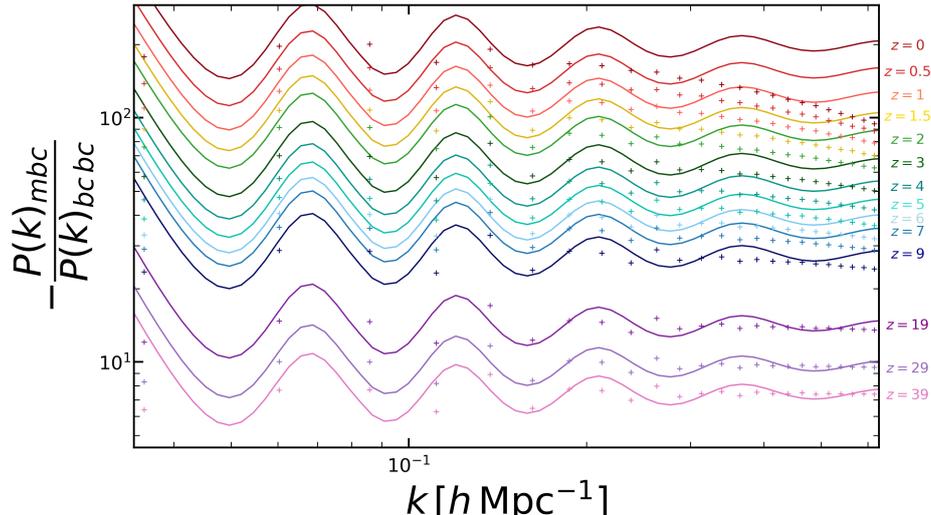
Figure 5.3: Comparison between the measurement of the ratio $-P_{mbc}/P_{bcbc}$ from the 2-fluid-diff-TF simulations with the CAMB prediction. We present results for 14 different output redshifts indicated by the color coding. Solid lines show the prediction of linear perturbation theory and plus ("+") markers represent simulations results. We find again overall good agreement validating once more our numerical setup. The suppression at small scales is due to our treatment of the force softening for baryons and nonlinear effects. The small inconsistency of our measurement with the theoretical prediction on the largest scale is due to the small number of modes in this first k bin.

unity implies that the overall shape of the power spectrum of baryons and CDM is different even on relatively large scales. The density perturbations in the baryon density field is smaller than the dark matter density field at all redshifts, this is because of the extra suppression produced by radiation pressure before recombination. In addition we clearly see the BAO wiggles due to the fact that BAOs are only present in the baryon power spectrum. These wiggles become less and less important at lower redshift as gravitational evolution slowly washes them away. We see that the difference between the two power spectra becomes of order 1% at redshift zero which is why simulations are usually initialized assuming the same power spectra for the two fluids. It is however precisely the impact of this difference on dark matter halos that we want to investigate in this chapter. The clearly visible suppression in P_{bb} compared to P_{cc} on small scales is due to our treatment of the force softening with AGS for baryons. Finally, the fourth panel of Fig. 5.2 presents the evolution of the relative baryon-CDM density field auto-power spectrum P_{bcbc} obtained from evaluating $\langle \delta_{bc}(\mathbf{k}), \delta_{bc}(\mathbf{k'}) \rangle$ with $\delta_{bc} \approx \delta_b - \delta_c$ (simulation results are once again shown by "+" markers and the solid lines comes from the linear perturbation theory by the CAMB code). We can see that there is no redshift dependency of P_{bcbc} which is expected since, as we showed in Sec. 1.1, the term proportional to θ_{bc} rapidly decays in δ_r leaving only the term δ_{bc} which is constant in time. The departure from this behaviour on small scales is probably due to nonlinear evolution as well as errors in our numerical setup due to the AGS. The results presented on this figure represent nontrivial tests of our simulations, and the good agreement with linear theory up to $k \sim 0.3 \, h \, \mathrm{Mpc}^{-1}$ at $z = 0$ allows us to validate them.

Before explaining our halo finding procedure, here we want to present some additional numerical tests to validate our numerical setup. We first compare our measurement of the ratio $-P_{mbc}/P_{bcbc}$ with the linear prediction from CAMB at various redshift on Fig. 5.3. The "+" markers shows the results from the 2-fluid-diff-TF simulations, and the solid lines show the expectation from linear perturbation theory. We find good agreement between the two at all redshift up to mildly nonlinear scales validating once more our setup. The small discrepancy observed on the largest scale is most likely due to the small number of modes in this lowest k bin. As can be seen on Fig. 5.3, the matter-bc cross spectra, P_{mbc} dominates the P_{bcbc} auto power spectra in all output redshift.

Finally, Fig. 5.4 shows the relative difference in the clustering of baryons and CDM as the ratio P_{bb}/P_{cc}. This corresponds to the middle panels of Fig. 5.2. The left panel shows the ratio at $z = 19$, and the right one at $z = 0.5$. Here we plot results from different runs featuring different force resolutions, with and without AGS for baryons. A redshift dependant versions of these plots shown in Fig. 5.5 Furthermore we show the impact of using different desired number of neighbours to setup the SPH kernel width (DesNumNgb), and different
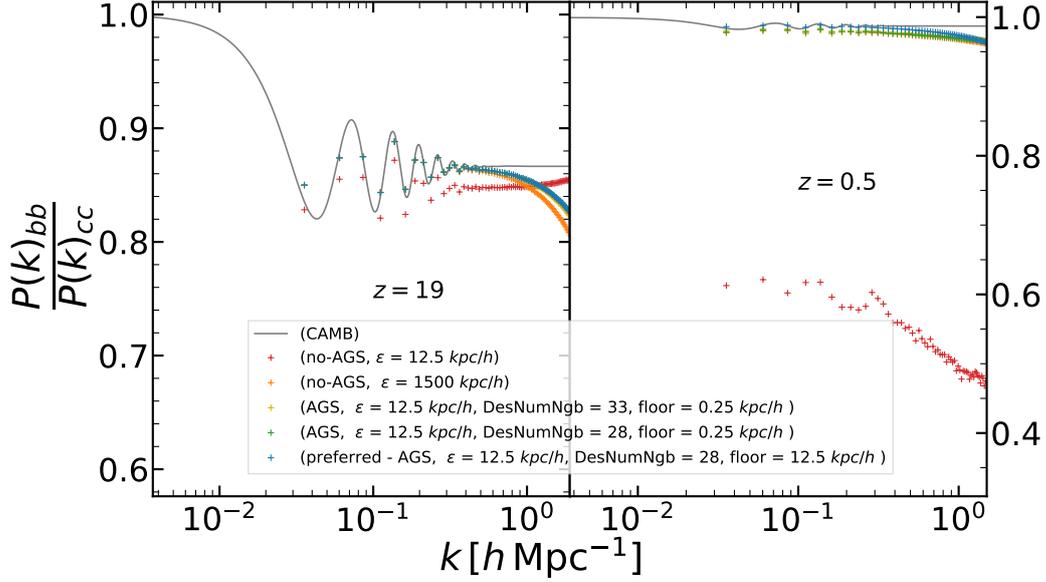
Figure 5.4: Numerical tests of the effect of the force resolution in 2-fluid simulations. We compare the ratio of the power spectrum of baryons and CDM for five different runs with different baryon softening lengths at $z = 19$ and $z = 0.5$ on the left and right panel respectively. The gray line displays the CAMB expectation. Red plus markers show a test run using a Plummer-equivalent softening length, ϵ, set to 1/40-th of the mean inter-particle separation corresponding to $\epsilon = 12.5 \mathrm{kpc}/h$. Orange plus markers show a similar test but using a very large softening length for baryons, set to 3 times the mean inter-particle separation corresponding to $\epsilon = 1500 \mathrm{kpc}/h$. These two test do not include AGS. Yellow plus markers show the results of another run using AGS, where forces between particles are softened adaptively using an SPH kernel with a width set by the distance to the 33rd closest neighbour. Green plus markers display a similar run but using the SPH kernel with a width set by the distance to the 28th neighbour. Both runs set a floor for the minimum softening length of baryon to 0.25kpc/h. The blue markers show our preferred setup where the SPH kernel width is set by the distance to the 28th neighbour but the minimum allowed SPH smoothing length is raised to 12.5kpc/h. All runs use a fixed softening length of 12.5kpc/h for CDM particles. We clearly see the need to use AGS with a reasonable softening length floor in order to recover the linear prediction from CAMB.
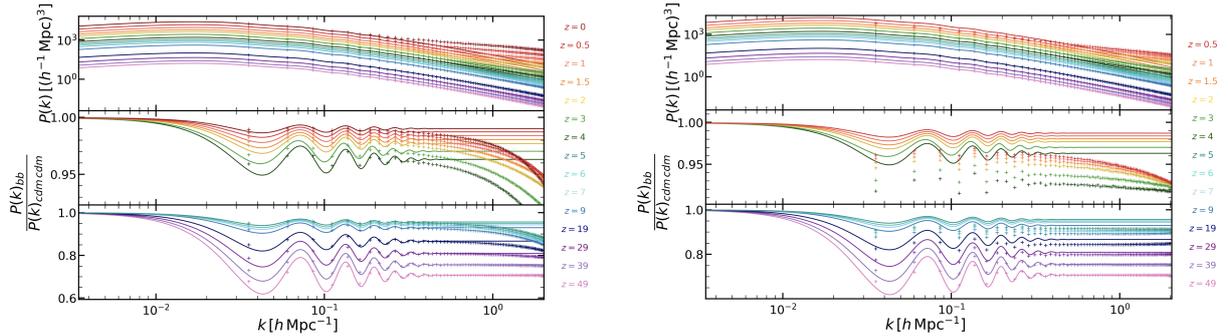
Figure 5.5: Numerical tests of the effect of Adaptive Gravitational Softening (AGS) for baryons in 2-fluid simulations for 14 different redshifts color-coded. The same as previous plots, solid lines indicate the prediction of the linear perturbation theory and the "+" markers represent the result of the simulation in each case (left and right panel) In the left panel simulations, we used AGS for baryons and in the right panel ones we did not. The top panel of each plot shows the growth of the total matter field. The middle and the last panel represent the ratio between the measured baryon to CDM power spectra. In the right panel, the non-AGS simulations, we see a discrepancy between the simulation results and the linear theory prediction at $z \sim 29$ growing by decreasing redshift to $z = 0$. This discrepancy would be solved up to $z = 0$ by using the AGS technique for baryons (right panel).

minimum allowed SPH smoothing length (floor) in case AGS is used. The desired number of SPH smoothing neighbours in the Gadget code represents the effective number of neighbours defined as the mass inside the kernel divided by the particle mass, and is kept constant very close to the desired value. A range can be defined to allow variation of the number of neighbours around the target value that we keep constant equal to 2. The floor is the minimum allowed SPH smoothing length which is used instead of the SPH kernel width in very dense regions. The color coding on Fig. 5.4 indicates the particular run, and the CAMB expectation is shown by the solid gray line. As discussed in detail in [64], using adaptive gravitational softening for baryons seems to correctly recover the relative large-scale clustering of baryons and CDM while avoiding the use of a very large smoothing length for baryons. This fact can be seen in the left panel of Fig. 5.4 at $z = 19$ for instance. The test run denoted with red plus ("+") markers shows results without using AGS but using a Plummer-equivalent softening length ϵ set to 1/40-th of the mean inter-particle separation. Although this is a standard value used in state-of-the-art simulations, in this case, it is underestimating the strength of the coupling between our two particle species yielding results completely inconsistent (about $\sim 5\%$) with the linear theory expectation. As can be seen in the right panel, this inconsistency becomes more and more important as we move to lower redshift (here for instance at $z = 0.5$ we can see $\sim 40\%$ discrepancy with the linear theory, to see the complete bahaviour in all output redshifts see Fig. 5.5). To remove this discrepancy at least at high redshift we can use a very huge softening length, ϵ, set to 3 times the mean inter-particle separation only for the baryon particles. For the CDM particles, we still set it to 1/40-th of the mean inter-particle separation. The result of this run is denoted in orange on Fig. 5.4. In spite of removing the discrepancy with the linear theory on large-scales, we can see a lack of power at smaller scales in this case. As can be seen on the right panel of Fig. 5.4, we still have about $\sim 3\%$ discrepancy with the linear theory expectation on large-scales at $z = 0.5$. The same occurs in another test, denoted in yellow, where we use AGS for baryons. In this test run, the forces between particles are softened adaptively using an SPH kernel with a width set by the distance to the 33rd neighbour, denoted as DesNumNgb=33. By decreasing the width of this kernel we can slowly reach the linear theory expectation even at lower redshift (green markers). Notice however that the minimum allowed SPH smoothing length (floor) was set to a very low values in these cases. By setting it to 1/40-th of the mean inter-particle separation and setting DesNumNgb=28 we can remove the discrepancy at all the redshift up mildly nonlinear scales. This is our final setting denoted with blue markers on Fig. 5.4. It is in agreement with linear theory up to $k \sim 0.3 \, h \mathrm{Mpc}^{-1}$. we refer to this setup as 2-fluid-diff-TF and use it to obtain the results presented in the main text.

2.4 Halo finding

Halos were identified using the Amiga Halo Finder (AHF) [119, 129], which identifies halos as spherical over-densities (SO) in the spatial distribution of particles in the simulations. (for a brief introduction on halos, see chapter 3, Sec. 3.3) The virial radius is defined as the radius within which the average density is given by $\bar{\rho}_{vir}(z) = \Delta_m(z) \, \rho_m(z)$ where ρ_m is the total matter background density (i.e. we identified halos using both
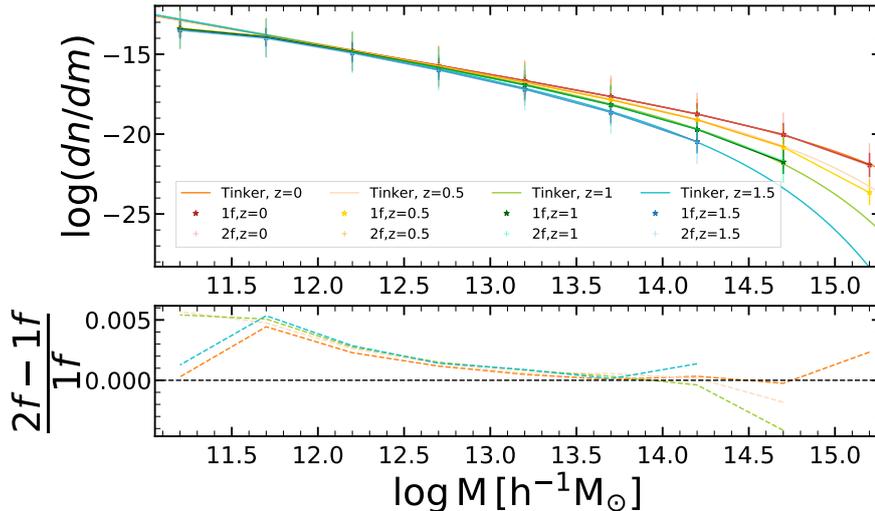
Figure 5.6: Halo mass functions for our 1-fluid and 2-fluid simulations at $z = 0$, $z = 0.5$, $z = 1$ and $z = 1.5$ indicated by the color coding. The symbols present our measurements while the solid lines are the Tinker 08 mass function. The shaded region show the 1σ errorbars obtained as the error on the mean over all realizations for each simulation. The lower panel shows the relative difference between the 1 and 2-fluid cases. We find good agreement with the Tinker fit as well as a subpercent difference between the two simulations sets as expected.

baryons and CDM particles), and we choose the overdensity threshold $\Delta_m = 200$. We refer the interested reader to [129] for more details. We set the minimum number of particles per halo to 20 and we use only main halos in this work (we discard subhalo). Finally, we identified halos at $z = 0$, $z = 0.5$, $z = 1$, $z = 1.5$, $z = 2$, and $z = 3$. We bin the mass range of halos in 8 tophat bins of width 0.5 in logarithmic scale centered from $\log M = 11.20$ to $\log M = 14.70$, where log is the base 10 logarithm, to ensure that we have enough halos in each bin.

To validate our simulations and our halo finding we present results for the halo mass function, $\mathrm{d}n(z)/\mathrm{d}M$ - i.e. the number of halos per unit volume per unit mass at redshift z - for our 1-fluid and 2-fluid simulations and a comparison with the well known Tinker halo mass function [246] on Fig. 5.6. We present results at $z = 0$, $z = 0.5$, $z = 1$ and $z = 1.5$ indicated by the color coding. The symbols present our measurements while the solid lines are the Tinker 08 mass function. The shaded region show the 1σ errorbars obtained as the error on the mean the over all realizations for each simulation. The lower panel shows the relative difference between the 1 and 2-fluid cases. We find good agreement with the Tinker fit, validating our halo catalogs, as well as a subpercent difference between the two simulations sets as expected since early baryonic effects should be at maximum of order 1% at low redshift. Furthermore since we keep the total matter power spectrum fixed on all scales we do not expect the halo abundance to depend strongly on the baryon-CDM perturbations.

3 Results

In this section we present results for the baryon fraction in halos, as well as the 2-fluid cross- and auto-power spectra constructed from δ_m, δ_{bc} and δ_h. We also show measurements of $b_{\delta_{bc}}$ as a function of halo mass M and linear bias b_1. In the last subsection we present the Baryon-CDM bias from 1-fluid "separate Universe simulations " as well.

3.1 Mean baryon fraction in halos

We first focus on the baryon fraction in halos F_b (normalized to the cosmic mean F_b^{cosmic}) as a function of total halo mass on Fig. 5.7. The red points in the upper panel show the baryon fraction of main halos identified in the first realization of our 2-fluid-diff-TF simulation (detailed in subsection Sec. 2.2). We compute the baryon fraction as
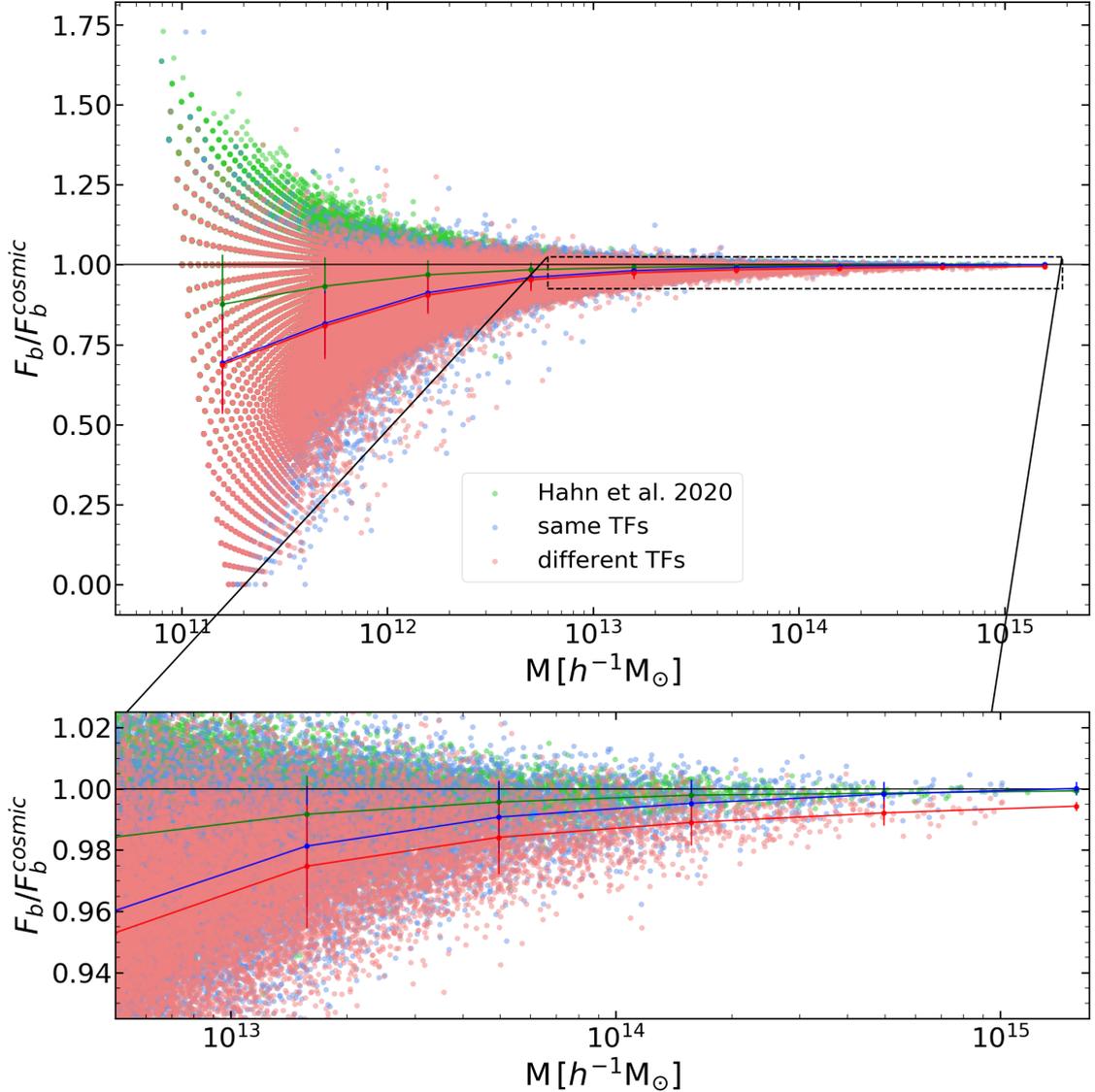
Figure 5.7: **Upper panel:** baryon fraction of halos in our 2-fluid simulations at $z = 0$ normalized to the cosmic mean, F_b/F_b^{cosmic}. Each point represents an individual halo while the linked points with errorbars show the mean and 1σ error on the mean in mass bins. We show results for the 2-fluid-diff-TF and 2-fluid-same-TF simulations in red and blue respectively. We see that in both cases F_b is consistent with the cosmic mean for well resolved halos of mass $M > 5 \times 10^{12} M_\odot/h$. We attribute the large downturn at small mass to AGS for baryons. To confirm that, in green we present results obtained from a simulation of [10] who do not use AGS. We see that in this case the mean baryon fraction stays consistent withe cosmic mean at all mass. see text for more details. **Lower panel:** Same as upper panel but zoomed on the region of well resolved halos with mass $5 \times M > 10^{12} h^{-1} M_\odot$. We see that for these objects F_b is indeed 95% of the cosmic mean, and reaches the cosmic value for all three numerical setups at very high mass, although a small $\sim 1\%$ difference remains for the 2-fluid-diff-TF setup.

$$F_b = \frac{M_{\text{baryon}}}{M_{\text{total}}}, \tag{5.17}$$

where M_{baryon} is the mass of baryons and M_{total} is the total mass of the halo (i.e. baryons + CDM). Notice that the mean cosmic baryon fraction for our cosmology is $F_b^{\text{cosmic}} = 0.1575$. We have a total sample of 282156 halos at z=0, represented by individual points. We further compute the mean and 1σ error on the mean in each of our 8 mass bins indicated by the linked points with errorbars. Noticeably this figure shows that for high mass objects the baryon fraction is independent of mass, and that the mean baryon fraction within the virial radius is approximately equal to the cosmic mean, with relatively small scatter (the standard deviation is less than 3% for halos of mass $M > 10^{13}\, h^{-1} M_\odot$), as can be seen on the lower panel of Fig. 5.7.

The baryon fraction in halos and clusters was extensively studied in hydrodynamical simulations (e.g. [247, 248, 249, 250, 251]). These works focused on late-time baryonic effects and we do not attempt to quantitatively compare our results with theirs here. However it is interesting to notice that all these works found a baryon fraction of roughly 0.9 of the cosmic mean in well resolved halos[3]. While we also find a baryon fraction slightly smaller than the universal one, in our case F_b/F_b^{cosmic} is greater than 0.95 for all halos with $M > 5 \times 10^{12} M_\odot/h$, and tends to unity at higher mass. This indicates that late-time baryonic effects are dominating over early ones for this quantity. However, depending on the precise considered mass, ignoring the effect of early baryon-CDM perturbations might lead to a non-negligible bias (up to $\sim 5\%$) when using the baryonic content of cluster to, e.g., infer the cosmic baryon fraction.

A noticeable decrease in F_b is observed below $M \approx 5 \times 10^{12}\, h^{-1} M_\odot$. However this may be ascribed to our use of AGS for baryons and not actually be physical. Indeed the number of particles in these halos is of the order of the number of neighbors used to establish the smoothing length for baryons, and hence these halos fall in the regime where the computation of the trajectories of the baryon particles is not accurate. This may lead to an underestimation of the density of baryons in these halos. Notice that such a downturn was already observed in [250] who also attributed it to the poor resolution of low mass halos. To further test this we ran a hybrid test of 2-fluid simulations with the same initial transfer functions for different species (2fluid-same-TF simulations), showed in blue on Fig. 5.7. Again points show our individual halos in our total sample of 281309 halos in this simulation, while the linked points with errorbars show the mean and error in each mass bin. We see that the mean value is not affected and that the scatter around the mean is also very similar in both cases which comforts our idea that the departure from the cosmic mean is not physical but merely a numerical artifact due to AGS, and gives us an indication of the limitations of our numerical setup. Finally, in order to confirm our hypothesis we use a simulation from [10] (Hahn et al. (2020) in the following) who vary individual particle masses to circumvent the use of AGS. Their cosmology is very similar to ours[4], and we use a simulation with the same box size, number of particles, and initialized with the Zel'dovich approximation too. The results for 369933 distinct halos as well as the mean relation are shown in green on the upper panel of Fig. 5.7. We can see that the deviations from the universal baryon fraction is much less than in our simulations, while the scatter stays approximately the same, confirming that the large downturn at low mass is due to AGS for baryons.

Although on average the baryon fraction in halos is not affected by baryon-CDM perturbations, deviations with respect to the mean could be related to such fluctuations. Thus, we next turn to investigating the observed scatter in F_b and its correlation with the local baryon-CDM perturbation δ_{bc}. To do this we smooth δ_{bc} with a tophat filter on a scale of $20\, h^{-1}$ Mpc and interpolate it at the halo positions. This yields the environment baryon-CDM density of halos δ_{bc}^R that we can then plot against F_b. Results are presented on Fig. 5.8 for individual halos, color coded by mass. Star markers show the mean in each mass bin. The left panel represents the $F_b - \delta_{bc}^R$ plane for our 2-fluid-diff-TF simulation, the middle one displays results for the case of our 2fluid-same-TF simulation, and the right panel comes from the 2-fluid simulation communicated by Hahn et al. 2020 [10] who use varying individual particle masses instead of AGS.

We see that the scatter in δ_{bc}^R is relatively independent of the halo mass in all cases. The values taken by δ_{bc}^R are in majority negative, as expected from Fig. 5.1, and they are much smaller in the case of the same transfer functions simulation, which is also expected. The fact that δ_{bc}^R is not exactly zero in that case results from numerical imprecision during the N-body evolution, and shows us what fraction of δ_{bc} is actually due to primordial baryonic effects compared to simple time evolution. We further notice that the scatter in F_b seems independent of the one in δ_{bc}^R. This indicates a small correlation between the local large-scale density δ_{bc} and the resulting baryon fraction in halos, confirming that the large scatter in F_b at low mass is rather unphysical (i.e. δ_{bc}^R only weakly affects the measured F_b) but is due to numerical effects of poorly resolved halos. The

---

[3]Notice however that when feedback is included (which was not the case in these works) this fraction drops to roughly 0.6 of the cosmic mean, see e.g. figure 4 and 5 of [252].

[4]Only $\Omega_b = 0.04897$ and $\sigma_8 = 0.8102$ differ from our parameters
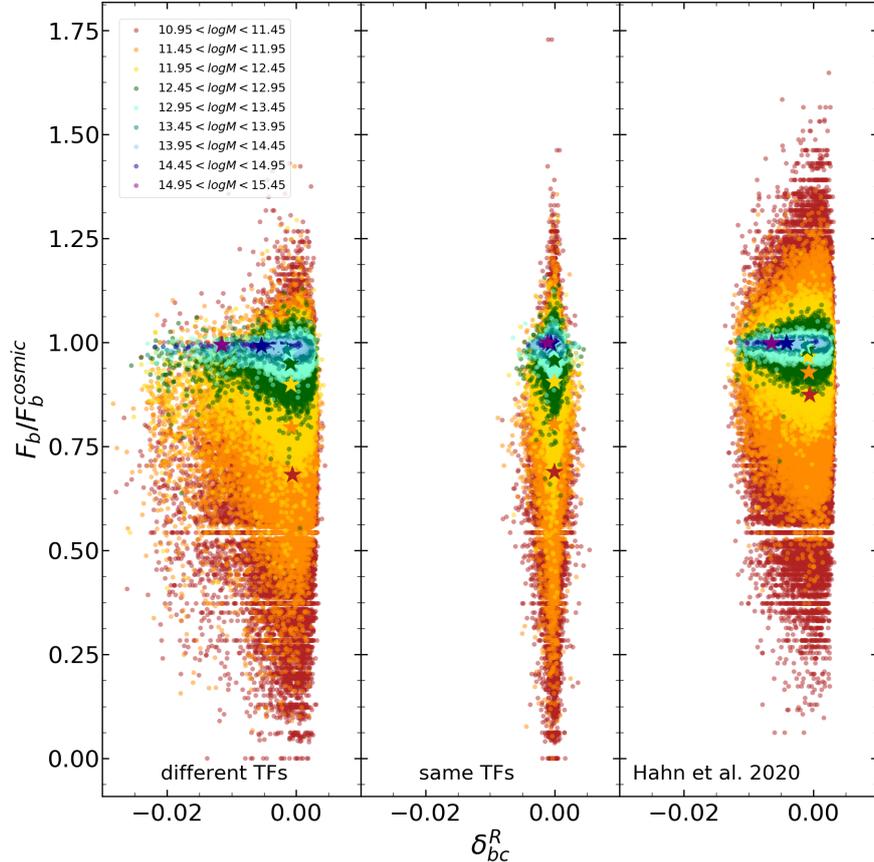
Figure 5.8: The baryon fraction as a function of the baryon-CDM density smoothed on a scale $R = 20\,h^{-1}\,\mathrm{Mpc}$, δ_{bc}^R, at $z = 0$. The color code indicates different mass bins. The left, middle and right panels show results for the 2-fluid-diff-TF, 2-fluid-same-TF, and Hahn et al. 2020 simulations respectively. Each point represents a halo while star markers show the mean value. We see that the scatter in δ_{bc}^R is relatively independent of mass and that the one in F_b is independent of δ_{bc}^R. The values of δ_{bc}^R are much smaller in the case of same transfer functions, and the mean values are close to zero, as expected. The fact that δ_{bc}^R is not exactly zero in that case results from numerical imprecision during the N-body evolution. The scatter in Hahn et al. (2020) seems roughly twice smaller than for our 2-fluid-diff-TF simulations. We however checked that this is only due to a few outliers with a high δ_{bc}^R, while the mean values are similar between the rightmost and leftmost panels.

scatter in Hahn et al. (2020) seems roughly twice smaller than for our 2-fluid-diff-TF simulations. We however checked that this is only due to a few outliers with a high δ_{bc}^R, while the mean values are similar between the rightmost and leftmost panels.

3.2   2-fluids power spectra and halo bias

We now turn to results for the impact of baryon-CDM perturbations on halo clustering. We first focus on the correlations between δ_m, δ_h, and δ_{bc} before presenting measurements of the associated bias $b_{\delta_{bc}}$ obtained as explained in Sec. 1.2 (Eq. 5.15).

Fig. 5.9 shows the cross- and auto-power spectra between the halo field δ_h, the matter field δ_m, and the baryon-CDM perturbation one $\delta_{bc} \approx \delta_b - \delta_c$ measured from our 2-fluid-diff-TF simulations at $z = 0$. Each panel presents a different mass bin. P_{hh}, P_{mm} and P_{bcbc} are the halo, total matter, and relative baryon-CDM density perturbation auto-power spectra respectively, while P_{hm}, P_{mbc}, and P_{hbc} are the halo-matter, matter-δ_{bc}, and halo-δ_{bc} cross-spectra, respectively. We show P_{mm}, P_{mbc} and P_{bcbc} on all panels even though they do not depend on mass to allow for a better comparison of the evolution of the halo power spectra with mass. The lines show the mean of all realizations while light shaded areas show the 1σ error on the mean. We note that since P_{hbc} and P_{mbc} are negative, we plotted $-P_{hbc}$ and $-P_{mbc}$. These negative values reflect the anticorrelation between δ_{bc} and δ_h, or δ_{bc} and δ_m coming from the fact that photon pressure in the early Universe prevents the collapse of baryons at early times, which in turn acts against the collapse of dark matter and the formation of halos. The fact that P_{hbc} is nonzero at all masses demonstrates that baryon-CDM perturbations δ_{bc} affect the clustering of halos even in the low-redshift Universe, and might have to be taken into account in the studies of structure formation. To our knowledge this is the first time that such correlations are shown.

Turning to the mass evolution of the spectra involving δ_h, P_{hh}, P_{hm}, and P_{hbc}, we note that all increase with increasing halo mass. In particular P_{hm} and P_{hh} have the well-known behaviour of going from values smaller than P_{mm} at low mass to dominating all the other spectra at high mass, reflecting the fact that low mass halos have a linear bias $b_1 < 1$ while more massive, rarer objects are highly biased (e.g. [253] and references therein). We also see that at high mass the halo auto-power spectrum becomes roughly constant over the whole k range indicating that it is dominated by the shotnoise term which is more important for these bigger and sparser objects. Since we have seen that the baryon-CDM field δ_{bc} is smaller than the total matter one we expect P_{bcbc}, P_{hbc} and P_{mbc} to be smaller than P_{hh}, P_{hm} and P_{mm}. In particular we expect P_{bcbc} to have the smallest value in comparison to the others. It is indeed what we observe on Fig. 5.9, and it is interesting to see that P_{bcbc} is nonzero, while P_{hbc} follows a similar mass evolution with respect to P_{mbc} than P_{hm} with respect to P_{mm}.

From the results presented on Fig. 5.9 we can obtain the bias associated to baryon-CDM perturbations $b_{\delta_{bc}}$ using Eq. 5.15. In order to maximize the signal to noise we follow the following procedure. We first obtain b_1 from our fiducial 1-fluid simulation by fitting a second order polynomial to the ratio in Eq. 5.16, i.e.

$$\frac{P_{hm}^{1f}(k)}{P_{mm}^{1f}(k)} = b_1^{1f} + A_1 k^2, \tag{5.18}$$

up to a maximum wavenumber $k_{\max}$. Here A_1 is simply the amplitude of the k^2 term we add to push the fit to higher $k_{\max}$. We do this for each of the first 4 realizations and for each mass bin. We then insert the result in the ratio Eq. 5.15 that we also fit with a second order polynomial, i.e.

$$\frac{P_{hm}(k) - b_1^{1f} P_{mm}(k)}{P_{mbc}(k)} = b_{\delta_{bc}} + A_{bc} k^2, \tag{5.19}$$

where again A_{bc} is an amplitude that we do not try to constraint. Importantly we use the result of b_1^{1f} in realization i to obtain $b_{\delta_{bc}}$ in realization i before averaging over all realizations and obtaining errorbars, and we do this for each mass bin. This allows us to cancel some of the cosmic variance since the CDM particles positions in our 1- and 2-fluid simulations are initialized with the same random seed. Since our simulations are of modest size (the fundamental mode k_F is $0.025\,h\,\mathrm{Mpc}^{-1}$) we choose a $k_{\max} = 0.21\,h\,\mathrm{Mpc}^{-1}$. We have tested the stability of our results under a change of this value and found it to be the optimal choice to maximize our signal to noise ratio while still obtaining unbiased results. Finally, in order to optimize the fit we put a loose mass-dependent flat prior on the value of $b_{\delta_{bc}}$ of roughly ten times its amplitude at a given mass.

The results for $b_{\delta_{bc}}$ as a function of halo mass are presented on Fig. 5.10 in green for various redshifts between 0 and 3. We further compare them to the results obtained from the 1-fluid "separate Universe simulations" in the manner of [242] in red. Since these authors described their procedure in great details, we only do a brief recap of it in Sec. 3.3. For each set of points the errorbars show the 1σ error on the mean obtained from all
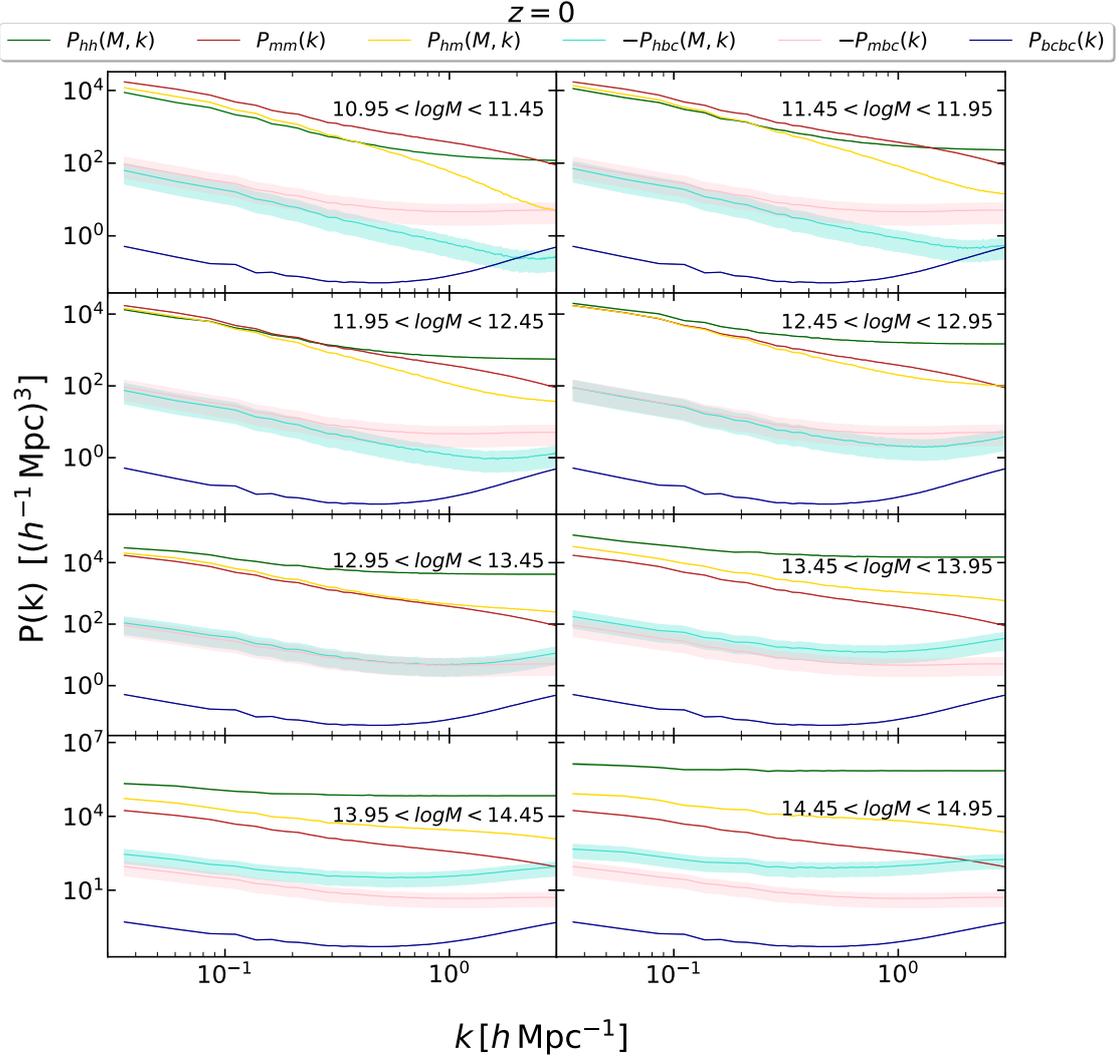
Figure 5.9: 2-fluid auto- and cross-power spectra constructed from δ_m, δ_{bc} and δ_h for all halo mass bins at redshift $z = 0$. We plot $-P_{\mathrm{mbc}}$ and $-P_{\mathrm{hbc}}$ since these quantities are negative, reflecting the anticorrelation between δ_{bc} and δ_h. The curves show the mean value and the shaded area show the 1σ error over all realizations of the 2-fluid-diff-TF simulations. We plot P_{mm}, P_{mbc} and P_{bcbc} on all panels even though they do not depend on mass to allow for a better visualization of the evolution of the halo power spectra with mass. The fact that P_{hbc} is nonzero at all masses demonstrates that baryon-CDM perturbations δ_{bc} affect the clustering of halos even in the low-redshift Universe. See text for more details.
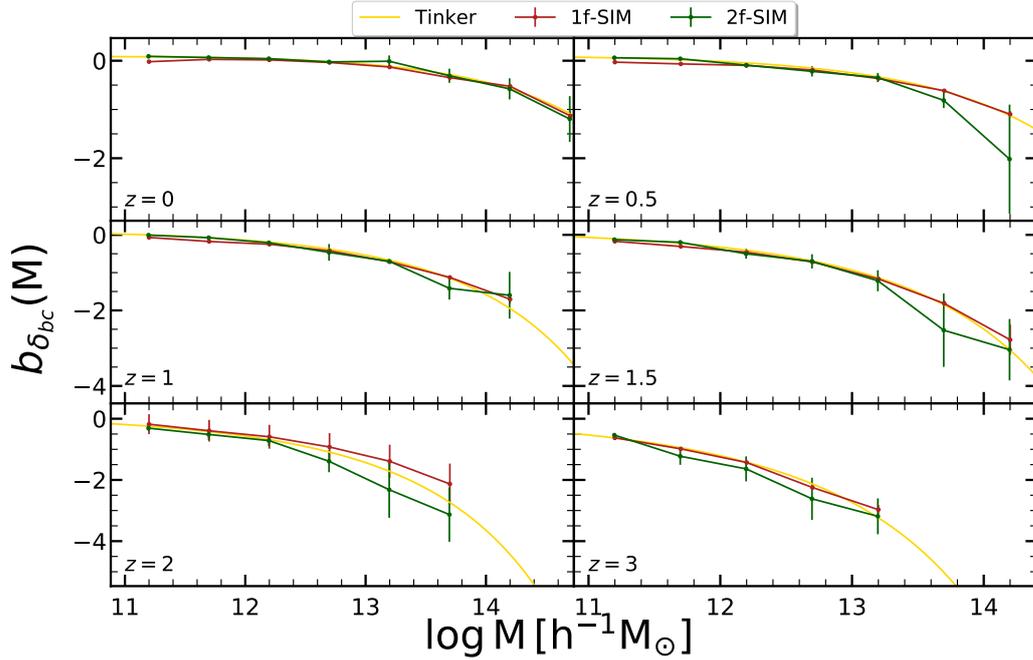
Figure 5.10: Baryon-CDM perturbations bias $b_{\delta_{bc}}$ as a function of total halo mass for different redshift. The green dots joined by the solid line indicate $b_{\delta_{bc}}$ as measured in 2-fluid simulations using Eq. 5.15. The red dots joined by the solid line show the same measurement from 1-fluid separate Universe simulations using Eq. 5.26. The errorbars show the 1σ error on the mean. The yellow solid line represents the Tinker prediction obtained using Eq. 5.20. We see that we get good agreement between the two methods as well as with Tinker. We find the same behaviour for this parameter as in previous work, i.e. it is overall negative with decreasing amplitude as a function of redshift, and more massive halos are more biased than small ones. see text for more details.
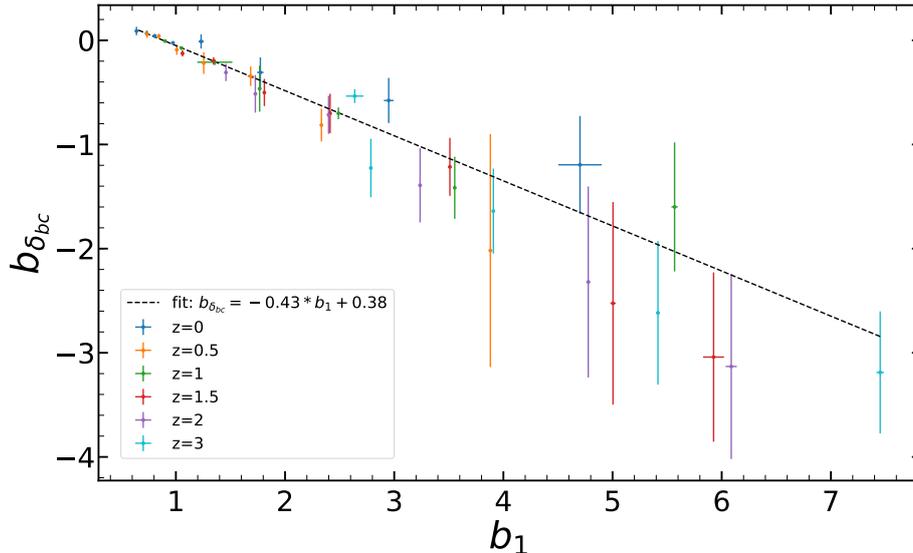
Figure 5.11: Baryon-CDM perturbations bias $b_{\delta_{bc}}$ as a function of the linear bias b_1 at different redshift indicated by the color coding. We see no clear trend with redshift although results at $z = 0$ seems to be systematically higher than the other ones. The dashed line is a linear fit obtained by fitting all points simultaneously.

realizations. The yellow solid line represents the prediction obtained using the universal Tinker mass function [246] as

$$b_{\delta}^{bc,univ}(z, M) = \frac{1}{\delta_{bc}} \left[\frac{n^{\text{SepUni,univ}}(z, M)}{n^{\text{Fiducial,univ}}(z, M)} - 1 \right], \tag{5.20}$$

where $n^{\text{SepUni,univ}}$ and $n^{\text{Fiducial,univ}}$ are the universal halo mass function predictions computed with the Tinker fitting function and the linear matter power spectrum of the fiducial and separate Universe cosmologies, respectively, and δ_{bc} is obtained with Eq. 5.25. Notice that Eq. 5.20 simply corresponds to Eq. 5.27 applied to the Tinker mass function. (see sec. 3.3)

Our results from the 2-fluid-diff-TF simulations are in overall very good agreement with the 1-fluid ones. It is an important cross-check since the two methods are completely independent and this parameter was only measured once before in [242]. We observe the same behaviour as these authors, i.e. that $b_{\delta_{bc}}$ is negative over most of the mass and redshift range and is a decreasing function of halo mass. It can be seen that $b_{\delta_{bc}}$ decreases more with halo mass at higher redshift. $b_{\delta_{bc}}$ is negative at $z > 1$, the only positive values for $b_{\delta_{bc}}$ being at $z = 0$ for halo masses between $10^{10.95} < M < 10^{12.45} \, [h^{-1} M_\odot]$, and at $z = 0.5$ for halo masses between $10^{10.95} < M < 10^{11.95} \, [h^{-1} M_\odot]$. The fact that $b_{\delta_{bc}}$ is negative again reflects the fact the baryon-CDM perturbations make halo formation more difficult. Finally, our results are also in agreement with the curves derived from the Tinker mass function.

We also note that even though we neglected the velocity bias $b_{\theta_{bc}} \theta_{bc}$ in Sec. 1.2, our measurements of $b_{\delta_{bc}}$ from 2-fluid simulations are in complete agreement within the errorbars with those from 1-fluid separate Universe simulations. This means that indeed θ_{bc} is subdominant at low redshift, as could be expected since it is a decaying term, and as was already pointed out in e.g. [11, 242]. This also mean that in order to try and measure this term from 2-fluid simulations one would need a much larger volume in order to obtain a detection. This goes beyond the scope of this work and we defer it to future work.

Finally on Fig. 5.11 we present $b_{\delta_{bc}}$ as a function of the linear bias b_1 at different redshift indicated by the color coding. We see no significant trend with redshift although results at $z = 0$ seems to be systematically higher than the other ones, especially at high mass as we noticed on Fig. 5.10. This figure shows us an approximately linear behaviour between these two quantities but while $b_1(z, M)$ is always positive $b_{\delta_{bc}}(z, M)$ is majoritarly negative. This motivated us to provide a fitting formula for this relation which might prove useful for accurate modeling of the halo power spectrum without introducing a new free parameter. We choose a linear fit, shown as the dashed line on Fig. 5.11, given by

$$b_{\delta_{bc}}(b_1) = -0.43 b_1 + 0.38. \tag{5.21}$$
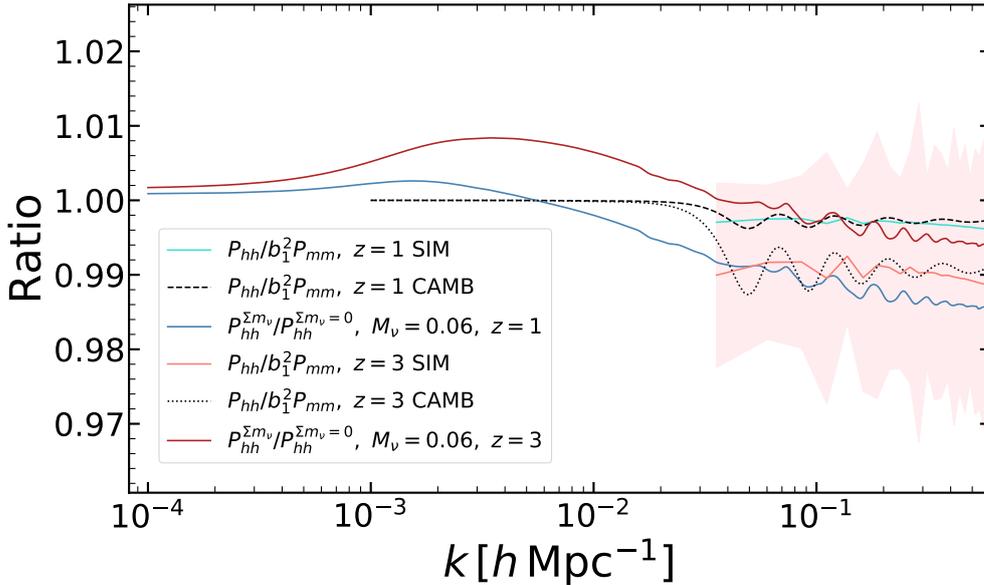
Figure 5.12: Comparison between the effect of early baryon-CDM perturbations and the effect of massive neutrinos on the halo power spectrum P_{hh} for a Euclid-like survey at $z = 1$, and Quasars from a DESI-like survey at $z = 3$. The light blue and light red curves with shaded regions for errorbars represent the ratio of the two models in Eq. 5.22 from our simulations at $z = 1$ and $z = 3$ respectively. The dashed and dotted black curves show the same quantity when we use power spectra from CAMB at $z = 1$ and $z = 3$ respectively in Eq. 5.22. The blue and red curve show the effect of massive neutrinos of total mass $\Sigma M_\nu = 0.06\,eV$ for a Euclid-like and DESI QSO-like survey at $z = 1$ and $z = 3$ respectively. We see that while the effect of neutrinos clearly dominates at lower redshift, it is subdominant compared to that of δ_{bc} at $z = 3$.

From the results on Figs. 5.9 and 5.11 we can assess the importance of the contribution of δ_{bc} to the halo power spectrum by comparing its value when including it or not, i.e

$$P_{hh} = b_1^2 P_{mm} \quad \text{vs} \quad P_{hh} = b_1^2 P_{mm} + 2b_1 \delta_{\delta_{bc}} P_{mbc} + b_{\delta_{bc}}^2 P_{bcbc}. \tag{5.22}$$

We do this comparison at $k = 0.1\,h\,\mathrm{Mpc}^{-1}$ and at $z = 0$ for all mass bins and find a maximum effect of roughly 0.3% for the highest mass halos[5]. Notice that at low mass the contribution due to δ_{bc} is negative while it becomes positive at high masses. For a more concrete case we consider a Euclid-like survey at redshift $z = 1$ for which the linear bias should be of 1.46. We can then use Eq. 5.21 to infer the value of $b_{\delta_{bc}} = -0.248$. Plugging everything into Eq. 5.22 we get a relative difference of 0.3%. Doing the same exercise for a higher redshift sample such as the DESI QSO one (centered around $z = 3$ with $b_1 = 3$ and hence $b_{\delta_{bc}} = -0.91$) yields a 1% impact of baryon-CDM perturbations on P_{hh} at the same scale.

We put these values in contrast with the effect expected from massive neutrinos for these two samples. We consider a total neutrino mass of $\Sigma M_\nu = 0.06eV$ consistent with the lower limit set by neutrino oscillations experiments (e.g. [254] and references therein). We consider the scale-dependent effect induced by neutrinos both on the linear matter power spectrum P_L and the linear bias $b(k)$ as parametrized in [255], i.e

$$\frac{P_{hh}^{\Sigma M_\nu}(z,k)}{P_{hh}^{\Sigma M_\nu=0}(z,k)} = \frac{[1 + (b_1(z) - 1)f(k)]^2 P_L^{\Sigma M_\nu}(z,k)}{b_1(z)^2 P_L^{\Sigma M_\nu=0}(z,k)}, \tag{5.23}$$

where the superscripts "ΣM_ν" and "$\Sigma M_\nu = 0$" stand for the model with and without neutrinos respectively, and $f(k)$ represents the transition between scales larger and smaller than k_{fs} (the neutrino free-streaming scale).

---

[5]Notice that for a pure affine relation between $b_{\delta_{bc}}$ and b_1 one could reabsorb the effect of δ_{bc} on the halo power spectrum in a scale dependent linear bias (by neglecting the last term in Eq. 5.22) as is done for neutrinos. The fact that $b_{\delta_{bc}}$ is nonzero at $b_1 = 0$ prevents from doing so.

It can be approximated by a step-like tanh function with a width of Δ_q [256]

$$f(k) = \left[1 + \frac{3\Delta_L}{2} \times \left(\tanh\left(\frac{q}{\Delta_q} \right) + 1 \right) \right] \tag{5.24}$$

where $q = \log(5k/k_{\rm fs})$, $\Delta_L = 0.55 f_\nu$ with $f_\nu = 0.045$ the neutrino fraction for our choice of Ω_m and ΣM_ν, and we used $\Delta_q = 0.6$. Finally we compute the ratio of the linear power spectra from CAMB by keeping the total matter density Ω_m, the baryon fraction f_b and the initial amplitude of perturbation A_s fixed.

Result are presented on Fig. 5.12 for a Euclid-like galaxy sample and a DESI QSO one. The light blue and light red curves with shaded regions for errorbars represent the ratio of the two models in Eq. 5.22 from our simulations at $z = 1$ and $z = 3$ respectively, while the dashed and dotted black curves show the same quantity when we use power spectra from CAMB. The blue and red curve show the effect of massive neutrinos of total mass $\Sigma M_\nu = 0.06\,eV$ at those two redshifts. We see that while the effect of neutrinos clearly dominates at lower redshift, it is subdominant at $z = 3$. Hence the impact of δ_{bc} has to be taken into account if one is to put solid constraint on neutrino mass from DESI QSO-like sample, which is a fundamental point for future surveys.

3.3 Baryon-CDM bias from 1-fluid "separate Universe simulations"

Here we describe in summary the separate Universe technique for the case of baryon-CDM perturbations in 1-fluid simulations. As already discussed in [242], the effects of baryon-CDM density perturbations on structure formation can be mimicked by a change in the baryon density Ω_b and in the CDM one $\tilde{\Omega}_c$, keeping the total matter density Ω_m constant. These changes can be described by a parameter Δ_b as $\tilde{\Omega}_b = \Omega_b[1 + \Delta_b]$ and $\tilde{\Omega}_c = \Omega_c[1 - f_b\Delta_b]$, where the tilde indicates the baryon and CDM density in the modified cosmology, and f_b is here the ratio of the baryon density over the CDM density in the fiducial cosmology $f_b = \Omega_b/\Omega_c$[6]. By using the following relation and a Taylor expansion we get

$$1 + \delta_{bc} = \frac{\tilde{\Omega}_b / \tilde{\Omega}_c}{\Omega_b / \Omega_c} = \frac{1 + \Delta_b}{1 - f_b\Delta_b} \approx 1 + (1 + f_b)\Delta_b. \tag{5.25}$$

We follow the same procedure as [242] and consider three different cosmologies dubbed Fiducial, High and Low with the following parameters: $\Delta_b^{\rm High} = 0.05$, $\Delta_b^{\rm Low} = -0.05$ and $\Delta_b^{\rm Fiducial} = 0$, corresponding to $\tilde{\Omega}_b^{\rm High} = 0.0515$, $\tilde{\Omega}_c^{\rm High} = 0.2596$, and $\tilde{\Omega}_b^{\rm Low} = 0.0466$, $\tilde{\Omega}_c^{\rm Low} = 0.2645$ while we keep Ω_m and all other cosmological parameters fixed in all three different cosmologies. We measure the baryon-CDM density bias $b_{\delta_{bc}}$ from 16 realizations of 1-fluid simulations as follows:

$$b_{\delta_{bc}}(z, M) = \frac{b_{\delta_{bc}}^{\rm High}(z, M) + b_{\delta_{bc}}^{\rm Low}(z, M)}{2}, \tag{5.26}$$

where

$$b_{\delta_{bc}}^{\rm High}(z, M) = \frac{1}{\delta_{bc}^{\rm High}} \left[\frac{N^{\rm High}(z, M)}{N^{\rm Fiducial}(z, M)} - 1 \right],$$

$$b_{\delta_{bc}}^{\rm Low}(z, M) = \frac{1}{\delta_{bc}^{\rm Low}} \left[\frac{N^{\rm Low}(z, M)}{N^{\rm Fiducial}(z, M)} - 1 \right]. \tag{5.27}$$

Here $\delta_{bc} = (1 + f_b)\Delta_b$, and $N(z, M)$ denotes the number of halos found in the corresponding cosmology at redshift z in some mass bin M. We refer the interested reader to [242] for more details about this procedure.

---

[6]Note that it is not the baryon fraction here but we follow the notation of [242]

6

Cosmic voids and BAO with relative baryon-CDM perturbations

In the previous chapter we introduced the gravity-only 2-fluid N-body simulation we have performed to quantify the impact of baryon-CDM perturbations on halo clustering and baryon fraction. In this chapter by reviewing the aforementioned simulations including baryons and cold dark matter (CDM) initialized with two different transfer functions, we study the statistics of various large-scale structure tracers in configuration space. In particular, we focus on the statistics of cosmic voids, as well as on the matter and halo real-space 2-point correlation function and baryon acoustic oscillations (BAO) peak. We find that the void size function is affected at the 1-2% level at maximum, and that the impact is more important at higher redshift, while the void density profile and void bias are roughly unaffected. We do not detect a sizeable impact of relative baryon-CDM perturbations on the real-space correlation functions of matter and halos or the BAO peak, which is in line with results from previous works. Our results imply that it would be hard to use voids or real-space correlation functions to constrain baryon-CDM relative perturbations, but also that we might not have to include them in models for the analysis of future cosmological surveys data.

This chapter is organized as follows. Sec. 1 is devoted to a detailed description of our numerical arrangement, including details of our simulation setup and the halo finding procedure (Sec. 1.1), as well as the void finding algorithm (Sec. 1.2). In Sec. 2, we explore the impact of relative baryon-CDM perturbations on the void size function (VSF), while we study the void correlation functions, void profile, and void bias in Sec. 3. We then focus on the BAO signal in our simulations in section Sec. 4, showing results for the full shape 2PCF in Sec. 4.1, as well as for the position of the BAO peak in Sec. 4.2.

1   Numerical setup

We begin with describing our numerical setup. The following subsections represent the simulations suit we performed to investigate baryon-CDM perturbations together with the halo finding method as well as the void identification algorithm.

1.1   Simulations and halo finding

Our N-body simulation suite (see chapter 3, Sec. 3.2 for a brief introductory explanation on N-body simulations) is based on the one presented in [12] (see chapter 5), and consists of

1. a set of collisionless gravity-only simulations in which baryons and CDM are evolved as two distinct fluids initialized from two distinct primordial power spectra as predicted by early universe physics. We refer to this set of simulations as "2-fluid".

2. a set of a standard gravity-only simulation in which the baryons and CDM are considered as perfectly comoving and are hence simulated as one total matter field. We refer to this set as "1-fluid".

Name	L_{box}	N_b	N_c	m_b	m_c	N_{real}
–	[Mpc/h]	–	–	[$10^{10} M_\odot/h$]	[$10^{10} M_\odot/h$]	–
1-fluid	500	0	512^3	–	1.0051	8
2-fluid	500	512^3	512^3	0.1583	0.8468	8

Table 6.1: Principal parameters of our numerical setup. L_{box} denotes the length of the side of the box, N_b and N_c are the number of baryon and CDM particles respectively, m_b and m_c denote their corresponding mass in units of $10^{10} M_\odot/h$, and N_{real} is the number of realizations.

Our cosmology is consistent with Planck 2018 [2] ΛCDM, namely: $\Omega_m = 0.3111$, $\Omega_b = 0.0490$, $\Omega_c = 0.2621$, $\Omega_\Lambda = 0.6889$, $n_s = 0.9665$, $\sigma_8 = 0.8261$ and $h = 0.6766$. In this work, we enlarge our previous simulation box size to $L_{\mathrm{box}} = 500 \, h^{-1}$ Mpc on each side to be large enough for void finding. We perform 8 realizations of each types of simulations (1-fluid/2-fluid) with 512^3 particles of each species. The details of the simulations are given in Tab. 6.1.

To generate the initial conditions of the density and velocity perturbations (see chapter 3, Sec. 3.2, Fig. 3.10 for a brief overview on the initial conditions) we used the publicly available initial condition code "MUSIC" [243] at redshift $z_i = 49$. For the 1-fluid case, we compute the matter power spectrum at $z = 0$ using the publicly available Boltzmann code CAMB [4] and back-scale it to the initial redshift, while for the 2-fluid scenario we use the two different transfer functions for baryons and CDM directly at $z_i = 49$. We use the first order Lagrangian perturbation theory (see chapter 3, Sec. 3.1), Zel'dovich approximation, [100] to estimate the velocity as well as the density fields. In order to reduce the effect of cosmic variance, we use the fixed-mode amplitude technique implemented in the MUSIC code [244]. Importantly, we keep the total matter power spectrum the same for the 1-fluid and 2-fluid scenarios, and we use the same random seeds to initialize 1-fluid simulations and CDM particles in the 2-fluid case.

We perform our simulations using the cosmological N-body code GADGET-II [245]. In the case of 2-fluid simulations, we use adaptive gravitational softening (AGS) [68]) for baryons only to alleviate the spurious coupling arising between CDM and baryon particles, as was discussed in [64, 12], and we refer the interested reader to these references for more details (see chapter 3, Sec. 3.2 and Eq. 3.72 for a brief discussion on softening length). Finally, we insist again that since we are interested in computing the effect of *early* baryon-CDM perturbations on LSS, we neglect the late-time impact of baryonic processes and do not include hydrodynamical forces in the simulations. We refer the reader to [12] (see chapter 5) for all the details and validating tests of our numerical setup.

We use the spherical overdensity (SO) algorithm (for a brief introduction on halos, see chapter 3, Sec. 3.3) Amiga Halo Finder (AHF) [119, 129] to identify halos. The definition of the virial radius is the one of a sphere in which the average density is given by $\bar{\rho}_{vir}(z) = \Delta_m(z)\,\rho_m(z)$ where ρ_m is the background total matter density. We chose the overdensity threshold as $\Delta_m = 200$, and set the minimum number of particles per halo to 20. For this work, we only used main halos and discarded subhalos from the catalogues. We identify halos at redshift $z = 0$, $z = 0.5$, $z = 1$, $z = 1.5$, $z = 2$, and $z = 3$. In the case of 2-fluid simulations, we use both CDM and baryon particles to identify halos. We compared the halo mass function in the 1-fluid and 2-fluid simulations and found good agreement (see Fig. 5.6 of chapter 5).

1.2 Void Finder

In chapter 3, Sec. 3.4, we brought an overview on cosmic voids and we introduced the void finding algorithm we use in the following analysis, here we describe the methods in more details, then we turn to our statistical analysis. We use the publicly available REVOLVER (REal-space VOid Locations from surVEy Reconstruction)[1] void finder to build our void catalogues with the ZOBOV (ZOnes Bordering On Voidness) algorithm [9], which is a 3D void finder and has been widely used both in simulations and observed catalogues [191, 192, 169, 192]. The ZOBOV algorithm performs a Voronoi tessellation of a set of points, identifies depressions in the density distribution of these points, and merges them into group of Voronoi cells using a watershed transform [189] without pre-determined assumptions about voids shape, size or mean underdensity, which is the most appealing aspect of the watershed method. Here we briefly outline the basic steps of the void-finding technique in ZOBOV and we refer the interested readers to the main ZOBOV paper [9] for a detailed description. One can describe the ZOBOV mechanism with the four following main steps:

1. *Voronoi Tessellation Field Estimator [190]*: the algorithm divides the space into cells around each tracer

---

[1] https://github.com/seshnadathur/Revolver

i (halos or particles in this work) in which the region inside the cell is closer to particle i than to any other one. It then estimates the density of each Voronoi region using the volume of each cell $1/V(i)$.

2. *Definition of the minimum density*: after estimating the density in each cell in the first step, the algorithm finds the minimum density cells, defined as Voronoi cells with a density lower than all their neighboring ones.

3. *Formation of basins*: the algorithm then joins adjacent higher-density cells to the minimum-density cell until no neighbor cell with a higher density can be found. It means that the void finder links all the particles to their minimum density neighbour. This procedure defines basins as the zones of these cells. At this point, these basins themselves could be considered as voids because they are depression regions in the density field, but one single basin may also arise from spurious Poisson fluctuations due to the discreteness of the particles.

4. *Watershed transform*: the last step is when these basins are joined together using a watershed algorithm [189]. For each basin b, the "water" level is set to the minimum density of b. It is then slowly elevated so that it can flow to the neighbor basins, joining all of them to basin b. The process stops when the "water" flows into a basin with a lower minimum, which defines the final void distribution.

Void centers are then defined as the center of the largest sphere completely empty of tracer that can be inscribed within the void. Indeed, this is the best predictor of the location of the minimum of the matter density field [257]. The effective radius of the void, R_v, is computed using the total volume of the underdense region and assuming sphericity

$$V_{\text{void}} \equiv \sum_{i=1}^{N} V_i^t = \frac{4}{3}\pi R_v^3, \tag{6.1}$$

where V_i^t is the volume of the Voronoi cell of the ith tracer, and N represents the number of points that are included in the void.

We run the ZOBOV algorithm for all realizations of our 1-fluid and 2-fluid simulations presented in Sec. 1.1 at redshift $z = 0$, $z = 0.5$, $z = 1$, $z = 1.5$, $z = 2$, and $z = 3$ for two tracers:

- Halos

- Dark matter particles.

In order to better handle the computational cost of running the void finder in the particle field, we have made a down-sampling routine to randomly select CDM particles of the simulation snapshots down to a constant average density of 6.71×10^6 particles per cubic box-size ($500\,h^{-1}\,\text{Mpc}$), which corresponds to 5% of the particles at each redshift, and insures us to be conservative with the density. We have verified that the different void statistics we study here were not affected when using a different random sample. We note that in the case of the 2-fluid simulation scenario, even if we have both types of particles (baryons & CDM) in the simulation, we only used the down-sampled positions of CDM particles. We should in principle select voids in the total matter density field, including baryons, however, the ZOBOV algorithm can not discriminate between different populations of particles with different masses. Therefore we must identify the voids in one of the two density fields only. Since CDM particles are much more massive than baryons, they are more representative of the underlying total mater field, and are the stronger contributor to the evolution of cosmic structures. We emphasize that we do not expect the inclusion of baryons or not in the void finding procedure to strongly affect our results.

We note that the total number of voids identified in the particle-field is significantly greater (from ~ 20 times for $z = 0$ to ~ 200 times for $z = 3$) than the number of voids in the halo field due to the difference in the mean tracer densities [168]. Moreover, for both types of simulations when one uses halos as tracer of the matter field, the total number of voids gradually decreases with increasing redshift (for instance for the 1st realization of our 1-fluid simulation we found 2085, 1950, 1621, 1225, 860 and 289 voids at $z = 0$, $z = 0.5$, $z = 1$, $z = 1.5$, $z = 2$ and $z = 3$ respectively) which is due to the fact that the number of halos formed at higher redshift is smaller than the ones at lower redshift which decreases the tracer density at higher redshift, and consequently the number of voids. On the other hand, in the case where CDM particles are used as tracer, the total number of voids increases as the redshift increases since we kept the tracer density constant at all redshift in this case (for example for the same 1st realization of the 1-fluid simulation in the particle field we found 32544, 42208, 52188, 61642, 70076 and 83430 voids at $z = 0$, $z = 0.5$, $z = 1$, $z = 1.5$, $z = 2$ and $z = 3$ respectively). In order to understand these features in the statistics of the voids in a better way, we will look at the distribution in size of cosmic voids in the next section.
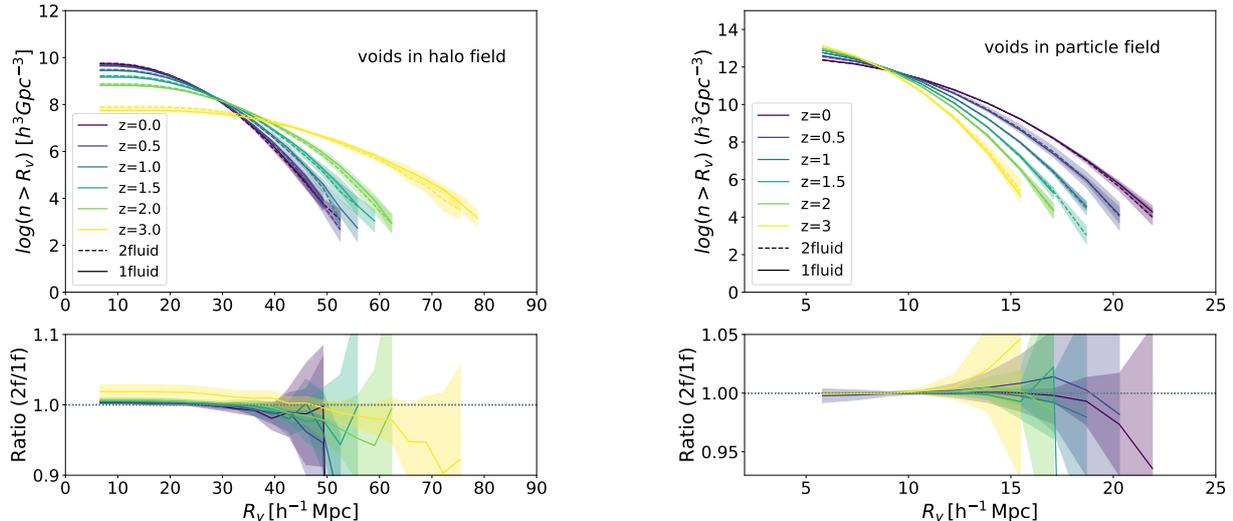
Figure 6.1: Cumulative void size function (number density of voids with radii above R_v) in the 2-fluid simulations in dashed and 1-fluid simulations in solid lines in the **halo field** (left panel) and in the **particle field** (right panel) for different redshift illustrated by the color bar. The lower panels show the ratio of the VSF as "2-fluid/1-fluid" to see the difference better. The gray dotted line in the lower panels stand for the situation in which the VSF is equal in both types of simulations. The shaded area in each case depicts the 1σ error on the mean obtained from the 8 different realizations.

2 Void Size Function

The Void Size Function (VSF), or void abundance [258, 259] is the number of voids in a given radius bin at a given redshift. The VSF is a relatively recent tool that nowadays is becoming promising to probe dark energy [260, 261] as well as constraining neutrino masses [167, 168, 169]. In addition to that, some recent works have also explored the differences between VSF in the concordance model of cosmology ΛCDM and modified gravity theories (see [162]), Galileon or non-local gravity (see [163]), or the possibility of couplings between CDM and dark energy (see [165]). Here we will present the comparison between the VSF in 1-fluid and 2-fluid simulations to assess the impact of baryon-CDM relative perturbations on these statistics. Each time we focus on voids identified both in the CDM density field (down-sampled) and in the distribution of collapsed halos to highlight how the use of different tracers with different bias might result in a different relative behaviour in the VSF. Notice that the impact of these perturbations has been studied in [12] for key observables of overdense regions of the density field (halo mass function and power spectrum, and the contribution of the baryon-CDM perturbation bias term to the halo power spectrum was found to be at maximum 0.3% at $k = 0.1 \, \mathrm{h \, Mpc^{-1}}$, at $z = 0$). However, they remain unexplored for underdense regions observables.

2.1 VSF in the halo field

The left panel of Fig. 6.1 shows the void size function of voids identified in the halo field both for the 1-fluid (solid line) and 2-fluid (dashed line) simulations. Based on the fact that no relevant differences have been observed between the halo mass function of the two types of simulations (see figure 3 of [12]), we are not expecting the void size function to be strongly affected either. We do however notice that the number of small voids identified in the halo field in the 2-fluid simulation is higher than the one in the 1-fluid simulation for all redshifts considered, while for larger voids ($R_v \gtrsim 40 \mathrm{Mpc/h}$) we can see the opposite trend (we identified more large voids in the 1-fluid simulation rather than the 2-fluid one). Nevertheless, these differences are relatively small and almost remain inside the errorbars (which shows the error on the mean obtained from the 8 different realizations). This can be seen more directly in the lower left panel of Fig. 6.1 that shows the ratio of the void size function in the 2-fluid and 1-fluid simulations. We see the most significant difference between 1-fluid and 2-fluid simulations for small voids at $z = 3$, where we observe more small voids in 2-fluid simulations with a significance of roughly 1.5σ. We see the opposite effect for larger voids but with larger errorbars and consistent with 1. We emphasize the fact that the observed trend is something that we are expecting, since clustering
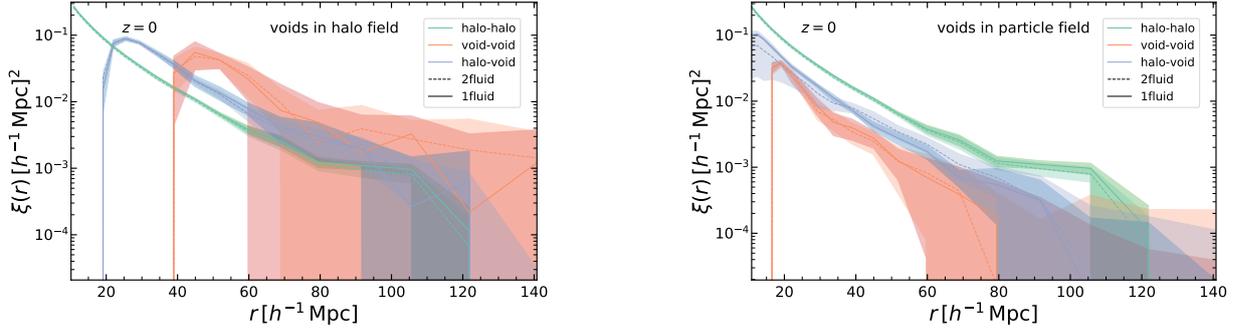
Figure 6.2: Void-void (orange), halo-halo (green) and halo-void (blue) 2PCF as a function of separation r using voids found in the **halo field** (left panel) and the **particle field** (right panel) at $z = 0$. The results for the 1- and 2-fluid cases are shown by solid and dashed lines respectively. The shaded area in each case shows the 1σ error obtained from 8 realizations.

is slightly diminished in 2-fluid simulations. Indeed, in [12], figure 9, we have shown that the amplitude of the ratio of the halo-halo power spectrum in 2-fluid simulation over the 1-fluid case is below 1, confirming the fact that baryon-photon coupling in the early Universe decreases the clustering in 2-fluid simulations. Hence, we expect to have more small voids and less large voids in 2-fluid simulations, and we expect this effect to be more important at higher redshift since gravitational evolution washes out relative baryon-CDM perturbations after decoupling. We also note that the effect of baryon-CDM perturbations on the cumulative VSF is smaller than the effect caused by massive neutrinos (see for instance figure 2 of [167] in which the authors observed an impact due to neutrino masses up to $\sim 30\%$ for $\sum m_\nu = 0.6\,\mathrm{eV}$ at $z = 0$). Finally, the left panel of Fig. 6.1 shows that in both types of simulations, ZOBOV found more small voids at lower redshift and more large voids at higher redshift as can be seen in the redshift trend shown by the color bar. This is also something that we expect, as discussed at the end of Sec. 1.2.

2.2   VSF in the particle field

The right panel of Fig. 6.1 presents the VSF for voids found in the particle field. While we found more large voids and less small voids with increasing redshift in the case of halo field voids, here we see that we find more small voids at higher redshift (and symmetrically less large voids). The redshift trend, in this case, is hence different than for halo field voids for which we recall that the average density of tracers in the box is evolving with redshift which is not the case for particles. This confirms, as was shown in various previous works, that the void population depends on the tracer type one is using, in particular on the tracer density and tracer bias (see for example [262, 263]). The particle field voids are smaller and found in greater numbers than the voids in the halo field. This is due to the fact that the distribution of collapsed halos is sparser than that of cold dark matter particles. These results are again expected, as we discussed at the end of Fig. 1.2.

For particle field voids, the difference in the number of voids found in the 1-fluid and 2-fluid simulations is even reduced compared to the halo field void case, and we do not observe any redshift evolution trend of the effect on these VSF caused by the 2-fluid formalism. Hence baryon-CDM relative perturbations impact the VSF of voids identified in halos more importantly, which suggests that they might also impact the VSF of voids found using luminous tracers (such as galaxies) in observations.

3   Voids 2-point statistics

In chapter 3, Sec. 2.2, we briefly explained the 2-point statistic as a well-known method to describe clustering. Here in this section, considering the voids in particular, we bring our results of the full correlation functions in Sec. 3.2, we devote Sec. 3.2 to void density profiles and we present our results of the void bias in Sec. 3.3.

3.1   Full correlation functions

The 2-point correlation function (2PCF) of a set of objects is a measurement of the degree of clustering of the considered objects defined as the excess probability of finding an object at a given distance from another one

with respect to a homogeneous distribution of objects. Estimators of the 2PCF, $\xi(r)$, in which r denotes the comoving separation, have been studied by various authors (see for example [264, 265, 266, 267, 95]). Among them, we use the so-called "natural" estimator [264] which has been implemented in the nbodykit pipeline[2] [268] to measure the void and halo auto/cross 2PCF in our simulation boxes.

$$\xi(r) = \frac{DD(r)}{RR(r)} - 1, \tag{6.2}$$

which calculates the 2PCF using a data catalogue D, and a synthetic random catalogue R. $DD(r)$ and $RR(r)$ represent the pair counts with separation r in the data and random catalogues respectively. Notice that, in order to reduce computational cost, nbodykit analytically estimates the random pairs $RR(r)$ in the case of uniform periodic randoms such as for simulated boxes.

Fig. 6.2 shows the void-void, halo-void and halo-halo auto(cross)-correlation functions at $z = 0$ in 1- and 2-fluid simulations (solid and dashed lines respectively), for voids identified both in the halo field (left panel) and in the CDM particle field (right panel). These correlations are computed for all halos and voids without binning in size. For all cases, the 2PCF is monotonically decreasing as a function of distance. In both panels, the amplitude of the halo-void correlation function stands between the halo-halo and void-void ones for all separations r. The halo-halo correlation function (green curve) is the same in both panels (since it obviously does not depend on the tracer used to find voids), and serves as reference to compare the two cases. In the case of voids in the halo field, the amplitude of the halo-void and void-void cross/auto correlations is higher than the halo-halo case, while in the particle field, the halo-void and void-void 2PCFs are lower than the halo-halo one. This demonstrates that, as expected, voids identified in the halo field are more correlated with halos than the voids found in the particle field. Another important feature here is that since voids are larger in the halo field than in the CDM particle field, the void-void 2PCF (orange curve) in the left panel starts to be nonzero at larger separation than the one in the right panel due to the exclusion effect. Indeed, since voids are low-density regions extending several tens of megaparsecs (hence with little amount of tracers inside them), the signal at scales inside the void radius becomes really low (or even zero) when computing the correlation function (or power spectrum) due to the lack of objects inside the voids, (see for instance [269, 270, 271]). This also has for effect to increase the amplitude of the correlation on larger scales in the halo field since larger voids (corresponding to a merging of small ones) can form in the halo field. Finally, we further note that due to the much larger number of halos in comparison to voids (~ 150 times larger) the signal to noise is much higher for the cross correlation than the auto correlation of voids. This for instance will have a consequence on the precision of the void bias estimation (see Sec. 3.3).

We now inspect in more details the impact of baryon-CDM relative perturbation on the 2PCFs by comparing results in the 1- and 2-fluid cases (solid versus dashed lines). We see that all differences are very small and well within 1σ errorbars. The largest difference is seen in the case of the halo-void correlation function for voids identified in the particle field (blue lines in the right panel), with the 2PCF computed in the 2-fluid case being slightly smaller at small scales and slightly larger at larger scales. Moreover, we see a small trend on the halo-halo 2PCF, where the 2PCF computed using 2-fluid simulations seems always slightly below the one computed from 1-fluid simulations. This suggests that baryon-CDM relative perturbations tend to lower the clustering, which is in agreement with the expectation of baryon-photon coupling slowing down the clustering process (as discussed in [12]). However, this effect is quite small and still within our errorbars. Note that this effect is also in agreement with the one we mentioned in Sec. 2.1 for the VSF, regarding the fact that since we have less clustering in 2-fluid simulations we identify more small voids and less large ones.

3.2 Density profiles

Cosmic voids are underdense regions close to their center with an overdense compensation wall at $r \sim 2R_v$, r being the radial distance from the center of the void. Moreover, the deepness of the void center, as well as the amplitude of the compensation wall have been shown to strongly depend on the void population considered (see for example [142, 144, 143]). The density profile of voids encodes the same information as the void-tracer cross correlation function since the radial profile of voids is indeed equal to the way that we count the number of tracers at distance r from the center of the void (see [150, 272] for a detailed explanation). In more details, the average radial number density of tracers at distance r from the void center, $\rho_{vt}(r)$ (also known in the literature

---

[2]https://github.com/bccp/nbodykit
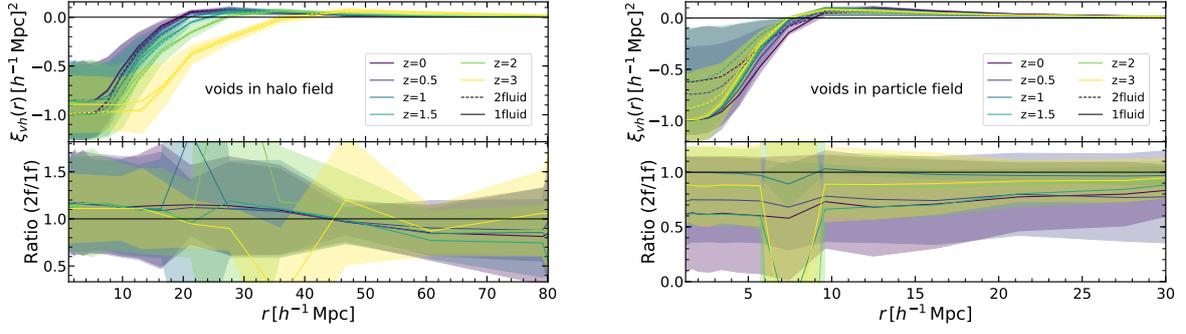
Figure 6.3: Halo-void cross-correlation corresponding to the void stacked profile for voids in the **halo field** (left panel) and in the **particle field** (right panel) at different redshift, and for 1-fluid and 2-fluid simulations (solid and dashed lines). We computed ξ_{vh} for all voids in our catalogues (i.e. without applying any cut in radius). Lower panels are showing the ratio of the 2-fluid simulations over the 1-fluid case. Note that the curves at $z = 0$ are equivalent to the blue curves in Fig. 6.2, with a vertical axis in linear scale.

as the *void stacked profile*), can be written as

$$\frac{\rho_{vt}(r)}{\langle \rho_t \rangle} = \frac{1}{N_v} \sum_i \frac{\rho_{vt}^i(r)}{\langle \rho_t \rangle}$$

$$= \frac{1}{N_v} \sum_i \frac{1}{N_t} V \sum_j \delta^D(x_i^{\text{center}} - x_j^t + r)$$

$$= V \sum_{i,j} \int \frac{1}{N_v} \delta^D(x_i^{\text{center}} - x) \frac{1}{N_t} \delta^D(x - x_j^t + r) d^3 x$$

$$= \frac{1}{V} \int \frac{\rho_v(x)}{\langle \rho_v \rangle} \frac{\rho_t(x+r)}{\langle \rho_t \rangle} d^3 x = 1 + \xi_{vt}(r), \tag{6.3}$$

where N_v and N_t are the number of voids and tracers respectively (with $\langle \rho_v \rangle$ and $\langle \rho_t \rangle$ their respective mean density), V is the total observed volume, x denotes the position (we use the index i to run over voids and j to run over tracers), and δ^D is the Dirac delta function. We have used the definition of the density of tracers within the void as a sum of Dirac deltas in the second equality, which can then be written as a convolution of the number density of the center of the voids ρ_v and the number density of tracers ρ_t (third and fourth equality), which is finally the definition of the void-tracer cross correlation function $\xi_{vt}(r)$.

We use this definition and compute the mean void profile as the halo-void cross correlation function for voids identified both in the halo and particle field. The void density profile for different redshift and different simulations scenarios (1-fluid and 2-fluid) are presented in Fig. 6.3. The left and right panels display the density profiles of the voids identified in halo and particle fields respectively. Note that Fig. 6.3 is similar to the blue curve in Fig. 6.2 but with a linear vertical axis, and for different redshift represented by the color bar. In Fig. 6.3, we can distinguish 3 different scales with 3 different behaviors in the density profile:

1. The innermost scales ($\sim r < \bar{R}_v / 2$) ($\bar{R}_v$ is the mean void radius) in which ξ_{vh} approximately tends to -1 at the void centers. Note that since the central part of voids is not totally empty, the cross correlation is not exactly equal to -1.

2. The intermediate scales ($\sim \bar{R}_v / 2 < r < 2\bar{R}_v$) or the void profile regime, on which we can see the compensation wall of the voids, which is a positive correlation around the void at all redshift. Notice that for voids identified in the halo field (left panel) the compensation wall moves to higher scales with increasing redshift. This is caused by the fact that the VSF at higher redshift is shifting towards larger radius voids (see Fig. 6.1, left panel). On the contrary, in the case of particle field voids (right panel), we see that the compensation wall moves towards lower scales with increasing redshift, which corresponds to the fact that the VSF of particle field voids at higher redshifts is shifting towards smaller radius voids (Fig. 6.1, right panel).

3. The linear regime ($\sim r > 2\bar{R}_v$) in which we see that the compensation wall disappears and $\xi_{vh} \to 0$. This is the regime in which we will compute the void bias in Sec. 3.3.
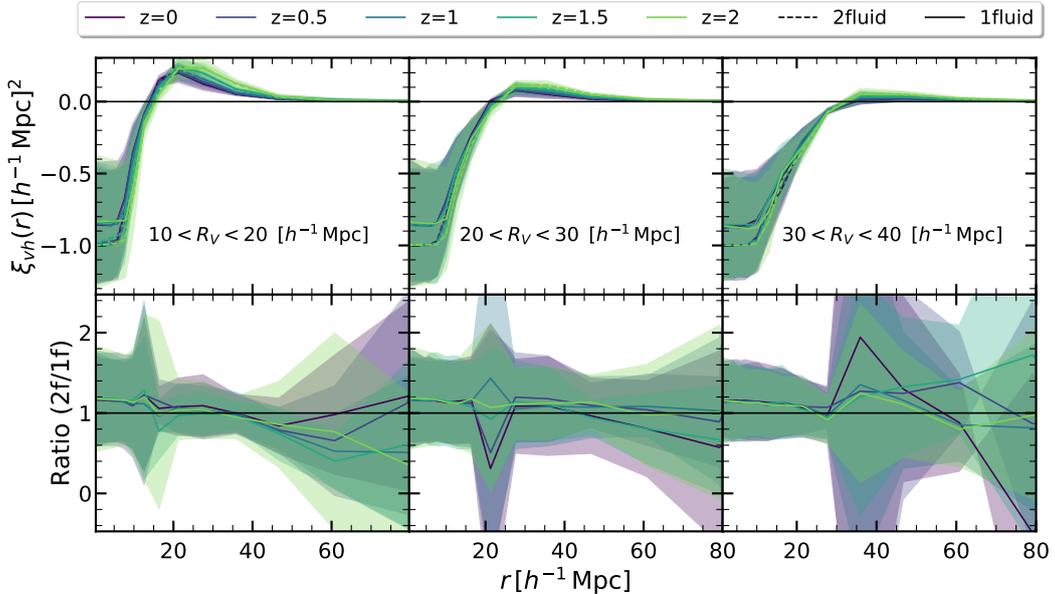
Figure 6.4: Void profile of **halo field voids** for 3 different bins of void radius R_v for 1-fluid (in solid line) and 2-fluid (in dashed line) simulations at 5 different redshift (color coded). In each bin and at all redshift the cross correlation approaches -1 close to the center of the void ($\sim (r < R_v/2)$). On scales $\sim (R_v/2 < r < 2R_v)$, the void profile shows a prominent compensatory ridge of halos for smaller voids $10 < R_v < 20$ h^{-1} Mpc, which disappears for the largest voids $30 < R_v < 40$ h^{-1} Mpc. In each bin, this compensation wall moves to lower radius (smaller voids) with decreasing redshift, which is the same behavior as we noticed in Fig. 6.3. Lower panels show the ratio of 2-fluid over the 1-fluid simulation results for each void size bin, where we see that voids in the 2-fluid case tend to be slightly less dense in their center.

Comparing the left and right panels of Fig. 6.3, we can also see that halo field voids have a much larger mean size than that of the particle field ones. This behaviour is confirmed by the VSF in Fig. 6.1. The bottom panels of Fig. 6.3 present the ratio of the density profile computed from the 2-fluid simulations over the 1-fluid scenario. We see that for the halo field voids, for small scales that are inside the void radius, the ratio is slightly greater than 1 at all redshift, suggesting that 1-fluid voids are somewhat smoother (recall that the density is negative on those scales). This effect is within errorbars but can be seen for the mean value of the ratio for halo field voids. Moreover, this effect is not seen in particle field voids (right panel) due to the fact that the signal is more noisy since we correlate particle field voids with halos. Finally, we note that errorbars in the void center are quite large due to the low-density definition of voids, and thus the lower amount of halos to compute the correlation.

The density profile of voids has been shown to depend on the void size (see for example [142, 263]), and we next explore whether or not the effect due to baryon-CDM perturbations could also vary with voids size. We divided our catalogues of void identified in the halo field in 3 different radius bins: $10 < R_v < 20$ h^{-1} Mpc, $20 < R_v < 30$ h^{-1} Mpc, , $30 < R_v < 40$ h^{-1} Mpc, and the catalogues of void identified in the particle field in 4 radius bins: $1 < R_v < 5$ h^{-1} Mpc, $5 < R_v < 10$ h^{-1} Mpc, $10 < R_v < 15$ h^{-1} Mpc, $15 < R_v < 20$ h^{-1} Mpc. The void profile (i.e. the void-halo cross correlation function) for each radius bin for each type of voids and at different redshift are shown in Fig. 6.4 (for halo field voids) and Fig. 6.5 (for particle field voids). In Fig. 6.4, we do not show results at $z = 3$ since the number of voids is quite small and the cross correlation signal becomes too noisy. For both types of voids and for all different void size bins, we observe the same 3 different regimes mentioned above (the innermost scale, the intermediate scale and the linear regime). We note that for all types of voids (found in halos or particles) the compensation wall found at intermediate scales (the void profile regime) is more pronounced at smaller radius: in Fig. 6.4, we see a clear positive bump in the first panel for smallest halo field voids, and as we move to the second and third panels (to larger voids), the bump becomes less prominent and it disappears in the last panel for the largest voids. We observe the same behaviour in Fig. 6.5 for particle field voids. The results found here are qualitatively in agreement with [273, 142] and [274].

In the same manner as for the void profiles of all voids (without classifying them by their radius), we show
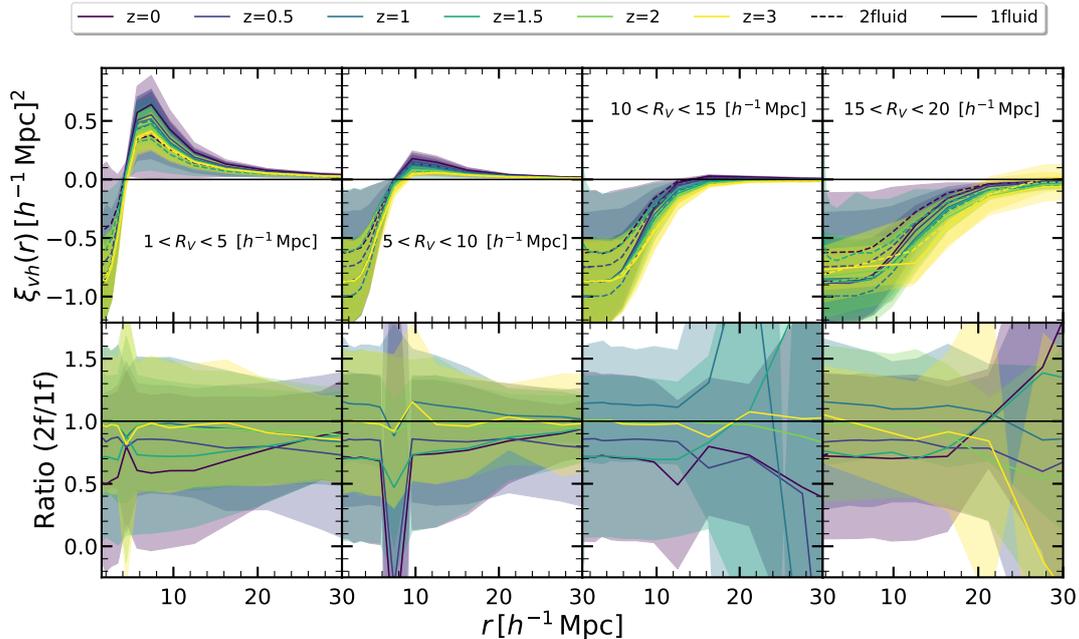
Figure 6.5: Same as Fig. 6.4 but for **voids in the particle field** in 4 different bins of void radius R_v. Since the number of voids identified in the particle field is larger than in the halo field, we divided them into more radius bins than halo field voids. The results of the 1-fluid scenario are shown by the solid lines and the ones of the 2-fluid by the dashed lines. The color bar represent different redshift. The void profile shows a sizable compensation wall for the voids in the smallest size bin ($1 < R_v < 5 \ h^{-1} \, \mathrm{Mpc}$). When moving to larger voids this structure becomes less prominent. The lower panels show the ratio of 2-fluid over 1-fluid results.

the ratio of the results from the 2-fluid scenario over the 1-fluid one in the lower panels of Figs. 6.4 and 6.5. For halo field voids (Fig. 6.4), inside the voids, we see that $\xi_{vh,\mathrm{2f}}/\xi_{vh,\mathrm{1f}} \gtrsim 1$ at all redshift which tells us again that 1-fluid voids are slightly smoother. We do not observe this for particle field voids (Fig. 6.5), because the signal is much more noisy again. We note that the effect of baryon-CDM perturbations on void profiles does not seem to depend on the void radius as we observe that the ratio seems to be similar inside the voids in all panels. Finally, we emphasis that these ratio are always compatible with 1 within 1σ errorbars, therefore we conclude that there are no significant differences between void profiles in 1-fluid and 2-fluid simulations, and hence that baryon-CDM relative perturbations due to photon pressure do not significantly impact this quantity. The results in each radius bin in Fig. 6.4 and Fig. 6.5 are compatible with those obtained for all voids without binning in size (Fig. 6.3).

3.3 Void bias

In addition to the different void observables presented above, we also aim in this work to quantify the impact of baryon-CDM perturbations on the linear bias of cosmic voids. Indeed, the estimation of the clustering bias of cosmic voids is an essential element to achieve competitive cosmological inference from voids, in the same way as galaxy bias in the case of galaxies ([97, 275, 276] and references therein). In this perspective, the interest in understanding it and modeling is raising [258, 269, 270]. Moreover, the possibility of using void bias directly to constrain cosmology is also recently gaining interest (see for example [277, 278, 279]). Here, we will measure the bias of our voids following the methodology described in [274], for both 1-fluid and 2-fluid simulations at different redshift, and considering both voids identified in the halo field and in the particle field. Similarly to [274], we define the void bias using two different expressions, the first one using the halo-void cross-correlation as

$$b_v^{\mathrm{cross}} = \frac{\xi_{vh}}{b_h \, \xi_{mm}}, \tag{6.4}$$
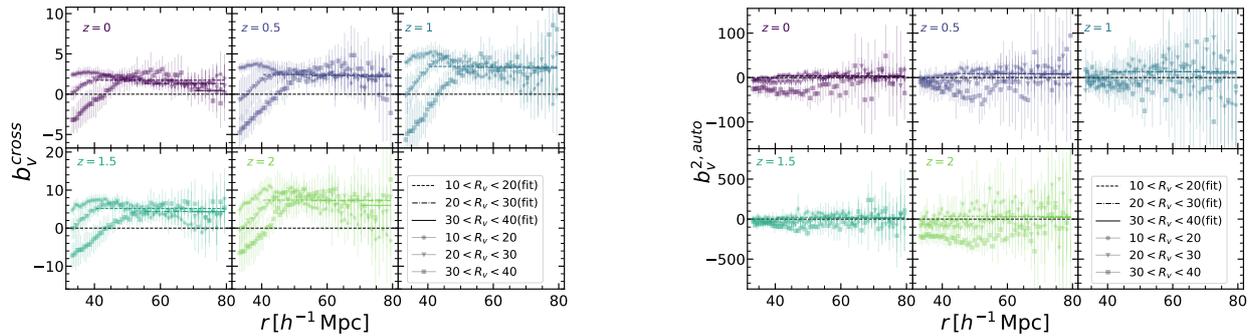
Figure 6.6: b_v^{cross} (left) and $(b_v^{\mathrm{auto}})^2$ (right) as a function of scale (Eqs. 6.5 and 6.6) and an example of the fit with a zeroth order polynomial to obtain the mean void bias value. Both panels present results from 2-fluid simulations for voids in the halo field. Each subpanel with different color presents results at a different redshift. Different markers ans line styles show the measurement and associated fit at different void radius R_v. The errorbars show the 1σ error on the mean obtained from 8 realizations. Since the number of voids is roughly ~ 150 times less than the number of halos, we have very large errors when computing $(b_v^{\mathrm{auto}})^2$.

in which the halo bias can be obtain using the halo auto-correlation signal as $b_h \equiv \sqrt{\xi_{hh}/\xi_{mm}}$. Thus one can rewrite Eq. 6.4 as

$$b_v^{\mathrm{cross}} = \frac{\xi_{vh}}{\sqrt{\xi_{hh}\,\xi_{mm}}}. \tag{6.5}$$

The second definition uses the void-void auto-correlation as follows

$$b_v^{\mathrm{auto}} = \pm\sqrt{\frac{\xi_{vv}}{\xi_{mm}}}, \tag{6.6}$$

where in all the above equations ξ_{mm} is the matter-matter auto correlation function measured directly from the simulation snapshots (using only CDM in 2-fluid simulations), and ξ_{hh} is the halo-halo auto correlation function shown in Fig. 6.2. For b_v^{auto}, we first measure the bias squared and then we chose the sign of the square root using the sign of b_v^{cross} (identically to what has been done in [274]). Considering the number of voids in each bin, we expect b_v^{auto} to be much more noisy. Nevertheless, it is interesting to cross-check to see if both bias measurements give comparable values.

Fig. 6.6 presents, as an example, the cross (left panel) and auto (right panel) bias as a function of scale for halo field voids at various redshift and void radius. Each small panel with different color presents a different redshift. Considering the few number of voids identified at $z = 3$ and the low signal to noise ratio resulting, we do not show the bias analysis results at $z = 3$. We use different markers for different void size bins. The markers here show the mean value of the bias and the errorbars are the 1σ error over 8 realizations. As expected, in the linear regime both bias are showing a constant behaviour. We then obtained the values for b_v^{cross} and $(b_v^{\mathrm{auto}})^2$ as a function of redshift and void size by fitting a zeroth order polynomial on linear scales (horizontal lines in the figure). In both cases, we use only scales between $2R_v < r(\mathrm{Mpc/h}) < 80$ for the fit. The lower limit assures that we are using only pairs of distinct voids, and the upper limit assures us to avoid the BAO scale on which dividing by ξ_{mm} would create a high noise. We use different line styles to show the fit in different size bins, and we show here the fit over the mean values taking into account the errorbars over different realizations. We also did the same fit for each of the realization to find the errorbars over the mean value of the bias from 8 realizations. As expected, we observe a higher amount of noise in $(b_v^{\mathrm{auto}})^2$ than in b_v^{cross} (notice the difference in y-axis range) due to the fact that the pair counts in ξ_{vv} are much smaller than ξ_{vh}. In addition, the errorbars are increasing with redshift due to the smaller amount of voids found at higher redshift. Regarding the values of b_v^{cross} and $(b_v^{\mathrm{auto}})^2$, since the linear bias of halos is increasing with redshift (e.g. [280]), one can expect the voids identified with this tracer to also become less biased as time evolves, which is indeed what we observe. We also see that the void bias slightly decreases with increasing void size which is in agreement with the results in [274, 269].

We then show in Fig. 6.7 the mean void bias as a function of the void radius integrated over the scales mentioned above (i.e. the value of the fits obtained on scales $2R_v < r < 80\,h^{-1}\,\mathrm{Mpc}$). We show both void bias results from cross-correlation, b_v^{cross}, and auto-correlation b_v^{auto} using different tracers (in left panels we present
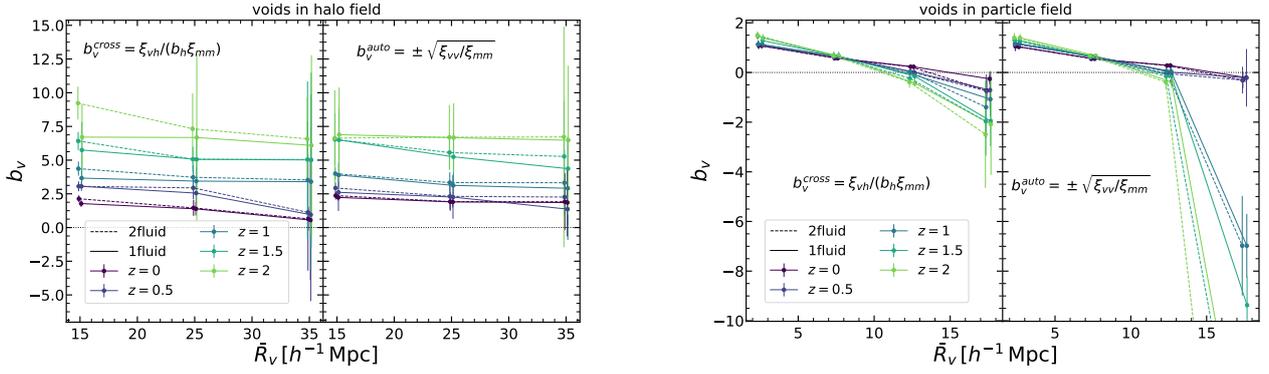
Figure 6.7: Mean, scale-independent void bias as a function of mean void radius $\bar{R}_v$ obtained from the fits on Fig. 6.6. **Left panels:** b_v^{cross} and b_v^{auto} obtained using voids identified in the halo field for all different redshift (color coded), for both 1-fluid (solid lines) and 2-fluid (dashed lines) simulations. We see that both b_v^{cross} and b_v^{auto} slightly decrease with increasing void size, and that both increase with increasing redshift. **Right panels:** same as the left panels but for voids found in the CDM particle field. In this case, we see that both biases depend more strongly on the void radius, and larger voids become negatively biased at all z. We further see that all voids become more positively bias *and* more negatively biased with increasing redshift. We observe only small differences that are all within 1σ errorbars between the void bias measured from 1- and 2-fluid simulations.

results in the halo field and the right panels show results in the particle field). Different colors show different redshift as before. Since in the particle field we have a much larger number of voids, the errorbars are quite smaller compared to the halo field results. Moreover, the number of voids in both halo and matter fields drops significantly with increasing redshift, resulting in more noisy correlation measurement, and consequently, the errorbars of our void bias measurement are also increasing with redshift. This is the main reason why we do not show results at $z = 3$.

Inspecting Fig. 6.7 in more details, we see that measurements of the bias from the two definitions, i.e. using either the auto (Eq. 6.6) or the cross (Eq. 6.5) correlation signals, are broadly consistent for all void size bins considered, both for halo field and for particle field voids, except for the highest radius bin of the particle field voids. However, this is probably due to the fact that the signal in this case is really noisy due to the low number of objects, which affects our measurements and might lead to a slight underestimation of the errorbars. A detailed investigation is beyond the scope of this work in which we focus on the comparison between 1- and 2-fluid simulations. If we now inspect the difference between halo field and particle field voids, we see that choosing different tracer significantly affects the void bias: voids identified in the halo field are more biased than the particle field voids which is something expected since dark matter halos are biased themselves. Furthermore, we find that in the case of the voids in the halo field, the mean value of the void bias is a slightly decreasing function of the void size (almost consistent with a constant considering the errorbars), while for the particle field, the void bias is a decreasing function as the size of the voids is increasing. In the right panel of Fig. 6.7, we observe that the particle field void bias changes sign at a specific "turning scale", which is a similar behaviour as observed by [274], with however a different turning scale. This turning scale is roughly at $\sim 15\,h^{-1}\,\mathrm{Mpc}$ for our voids in the particle field while roughly at $\sim 25\,h^{-1}\,\mathrm{Mpc}$ for SDSS voids in [274]. However, we do not expect to observe the change of sign at the exact same scale since these authors find voids in a different tracer field using a different void finder.

Comparing the void bias from 1-fluid and 2-fluid simulations (solid versus dashed lines), we see that voids from the 2-fluid simulations are slightly more biased for both voids from the halo field and the particle field. This difference is within 1σ errorbars, but the trend of the 2-fluid simulation bias being slightly larger is expected: since the 2-fluid halo-halo 2PCF (the green curves in Fig. 6.2) is showing less clustering than in the 1-fluid scenario, the linear halo bias b_h is expected to be smaller in 2-fluid simulations as well. Then we can see from Eqs. 6.4 and 6.6 that the void bias should be slightly larger in the 2-fluid case.

4 Baryon acoustic oscillations

In chapter 3, Sec. 2.5 (see Fig. 3.7), we described the BAO feature in power spectrum and 2-point correlation function as a powerful tool in precision cosmology. In this section, we extend the computation of the real-space

2-point correlation function in 2-fluid simulations from voids to each component of the simulations, i.e. total matter, baryons only, CDM only, baryon-CDM relative perturbations (δ_{bc}), and halos. In particular, we focus on modulations of the BAO feature and BAO peak position by comparing our results for the total matter and halo fields in 1- and 2-fluid simulations. This is a direct extension of our previous work [12] (chapter 5) where we focused on Fourier space quantities.

Relative velocity perturbations between baryons and CDM can possibly shift the BAO scale because they are sourced by the same physical effect which imprinted the BAO peak itself. The shift in the BAO scale is crucial for cosmology since it could lead to a potential systematic shift in measurements of the angular diameter distance $D_A(z)$, the Hubble factor $H(z)$, and the growth factor $f\sigma_8$ [69, 70, 72, 73]. This effect might also be important to obtain unbiased results when one is investigating the effect of massive neutrinos on the BAO scale [281] or when one is using reconstruction methods to measure the BAO location in 21 cm intensity mapping surveys [282, 283].

4.1 Full-shape correlation function

In this subsection, we first focus on the full shape of the 2-point correlation function. To compute the 2-point correlation function in real-space, we use the Fast Fourier Transform (FFT) estimator introduced in [284] in which the density field is computed on a grid in Fourier space, squared, inverse Fourier transformed, and averaged in radial bins

$$\xi_{\text{SIM}}(r) = \frac{1}{N_{\text{modes}}} \sum_{r_{\min} < |r| < r_{\max}} \text{FFT}^{-1} \left[|\delta(\text{k})|^2 \right] (\text{r}), \tag{6.7}$$

where the sum runs over all radii r in the bin and N_{modes} is the number of modes in the bin. We use the Cloud-In-Cell (CIC) mass-assignment (explained in chapter 3, Sec. 3.2, Fig. 3.14) scheme to compute the density field on the Fourier grid $\delta(k)$. To compute the total matter field δ_m in 2-fluid simulations is given by the weighted sum of the CDM field δ_c, and the baryon field δ_b, as δ_m as $\delta_m = f_b \delta_b + (1 - f_b)\delta_c$, where $f_b = \Omega_b/\Omega_m$. Moreover, we choose the edges of the bins $r_{\min}$ and $r_{\max}$ such that each bin as a width given by the mean interparticle separation, which in our case is $512 \, h^{-1} \, \text{Mpc}/500 \approx 1 \, h^{-1} \, \text{Mpc}$. We take advantage of the fact that this estimator is implemented in the PYLIANS library [3], which we use to obtain our results. Finally, we restrict ourselves to the real-space 2-point correlation function in 1-fluid and 2-fluid simulations without considering redshift space. The estimator introduced here to calculate the 2PCF is much faster than the natural estimator we were using to compute the void correlation functions and density profiles in Sec. 3. There we were using the natural estimator since the sparsity of voids and exclusion effects introduce large noise which prevented us to use the Taruya estimator to obtain the void profile. Here, since we are interested in the correlation function of particles and halos, which are by far more numerous, we can use the FFT estimator to significantly reduce the computation time while keeping a high-level of accuracy.

Fig. 6.8 shows the total matter-matter 2PCF computed in 1- and 2-fluid simulations (solid and dashed lines respectively) using Eq. 6.7 for different redshift (color coded). We recognize the standard shape of the correlation function which decreases as r increases, as well as the BAO peak at around $r \sim 105 \, h^{-1} \, \text{Mpc}$. We also see that both the correlation and the BAO peak increase with decreasing redshift since the clustering becomes more important at lower redshift. We observe small differences between the two cases with the correlation function being slightly lower on smaller scales in 2-fluid simulations, while on the scales of the BAO peak, the 2-fluid simulations give us a higher value of the 2PCF, and the effect is more important at low redshift due to nonlinear evolution (recall that the total matter linear power spectrum is kept constant between 1- and 2-fluid simulations). Notice that these differences are within 1σ errorbars obtained over different realizations on all scales. These results confirm that baryon-CDM relative perturbations have a rather small impact on the matter clustering (under the detection threshold corresponding to our simulation volume) as was already pointed out in [64, 12].

We now turn to a more detailed investigation of the cross-correlation of each fluid component in 2-fluid simulations in Fig. 6.9. The top and middle panels show the two different component of the matter field (baryon and CDM) 2PCF divided by the square of the linear growth factor D^2. In case of baryons, we can see that the correlation function exhibits a strong BAO peak at high redshift, and that then the amplitude of the peak decreases with redshift due to gravitational interactions with CDM particles (note that with decreasing redshift the 1σ error on the mean value increases). We can also see a small scale-dependent suppression of the correlation function at scales $r \lesssim 80 \, h^{-1} \, \text{Mpc}$ to accommodate for the growing peak. We see a somewhat different behaviour for CDM in the middle panel of Fig. 6.9: from $z = 39$ to $z = 7$, we see the BAO peak

---
[3]https://github.com/franciscovillaescusa/Pylians
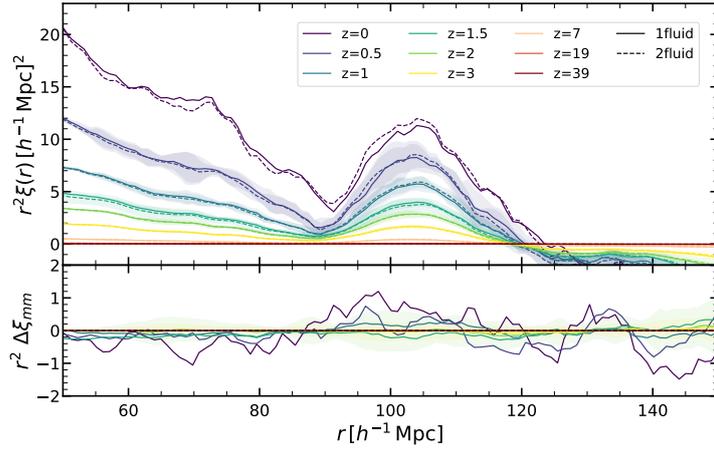
Figure 6.8: The redshift evolution of the full-shape total matter 2-point correlation function in 1-fluid (solid) and 2-fluid (dashed) simulations in real space as measured by Eq. 6.7. We multiply the 2PCF by r^2 to see the BAO peak better. The shaded area show the 1σ errorbar on the mean obtained from the standard deviation over all realisations. The lower panel presents the difference between 2-fluid and 1-fluid sets: $r^2 \, \Delta\xi_{mm} = r^2 \, (\xi_{mm}^{2f} - \xi_{mm}^{1f})$. We see that any small difference between the two cases is within the errorbars on all scales.
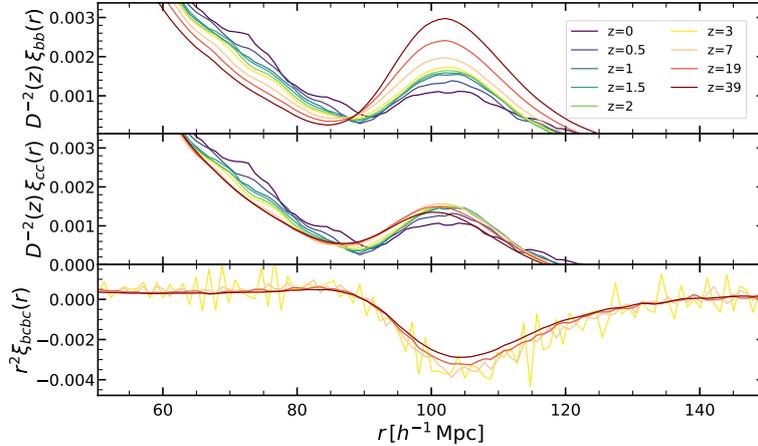
Figure 6.9: **Top panel:** Baryon 2-point correlation function measured in the 2-fluid simulations at different redshift (color coded). We see clearly how the BAO peak diminishes with time in this field. **Middle panel:** Same as top panel but for the CDM fluid. In this case and in this range of z, the BAO peak remains roughly constant. **Bottom:** The δ_{bc} relative perturbation auto-correlation function. In this case, we show results down to $z = 3$ only since the noise becomes too important at later times. The BAO feature is clearly visible and is negative in this field. Furthermore, we see no redshift evolution, which is consistent with the fact that δ_{bc} is constant in time, as discussed in e.g. [11, 10, 12]. Note that the two upper panels are divided by the square of the growth factor $D^2(z)$ to see the difference in evolution of BAO in baryons and CDM, while in the bottom panel we multiplied the 2PCF of δ_{bc} by r^2 in order to show the BAO feature better. Shaded area on each curve represent the 1σ error, and we see that with increasing redshift the error becomes less prominent.
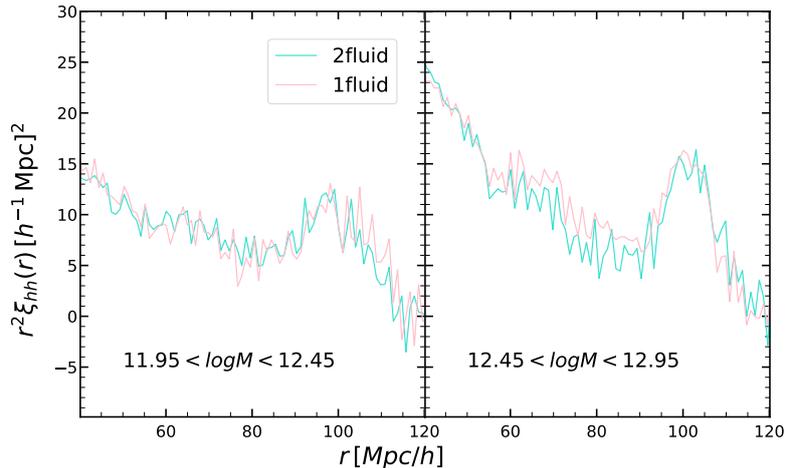
Figure 6.10: Halo-halo 2-point correlation function at $z = 0$ for 1-fluid (in pink) and 2-fluid (in blue) simulations for two different halo mass bins. The shaded area show the 1σ error over 8 realizations of each set of simulations. Again we multiply the 2PCF by r^2 to better see the BAO feature. We do not detect any impact of baryon-CDM relative perturbations on this 2PCF either.

slightly increasing as CDM particles fall in the baryon potential well on these scales, imprinting the feature from the baryon field into the CDM field gradually (note that we observe the same position of the peak in baryons and CDM). The peak reaches its maximum relative amplitude at roughly $z = 7$, the moment at which mild nonlinear effects appear. At redshift lower than $z = 7$, we observe then a small decrease in the peak amplitude. On small scales, we note the same scale-dependent suppression for CDM fluctuations that appeared as well in the baryon fluctuations. The results here are in agreement with the ones in figure 9 of [64]. In addition, as we saw for the halo-halo 2PCF in Fig. 6.2 (green curves), and also for the halo-halo power spectrum in figure 9 of [12] (chapter 5, Fig. 5.12), baryon-CDM relative perturbations tend to diminish the clustering. We however observed a slight increase of clustering on scales around the position of the BAO peak in the matter-matter 2PCF in Fig. 6.8. We can now understand this in light of Fig. 6.9: the pronounced baryon BAO feature increases the total matter BAO peak in 2-fluid simulations.

Finally, we compute the 2-point correlation function of the baryon-CDM perturbation field δ_{bc} in the bottom panel of Fig. 6.9. We show this 2PCF only down to redshift $z = 3$ because the noise increases as we reach lower redshift, and the 2PCF becomes consistent with zero on all scales. We see that this 2PCF is roughly constant close to zero, except for the BAO feature which is a BAO dip instead of the BAO peak in this case. This is because the BAO feature in the baryon field gradually imprints itself into the CDM field, which creates a skewed distribution of CDM with a sharp fall inside the BAO scale but with a larger tail on scales slightly larger than the BAO one (even though the position of the BAO peak is observed to be identical for baryons and CDM). Therefore we expect to observe an anti-correlation signal for δ_{bc} on scales slightly larger than the BAO scale (δ_{bc} is too small) in a skewed way, as can be seen in the bottom panel of Fig. 6.9. We do not observe any notable redshift evolution for this 2PCF which is consistent with the fact that δ_{bc} itself is constant in time, as discussed in e.g. [11, 10, 12]. Notice that this kind of correlation function was also predicted using 2LPT in [61]. While we do not conduct a detailed quantitative comparison of their prediction with our results, we note that they found the same kind of dip for correlation functions including relative baryon-CDM density perturbations.

Finally, we investigate the halo-halo 2PCF at redshift zero in Fig. 6.10. We present results for two halo mass bins centered around $\log M = 12.2 M_\odot/h$ and $\log M = 12.7 M_\odot/h$. Recall that halos in the 2-fluid simulations are identified by considering both types of particles (baryons and CDM). As we see the halo-halo 2PCF is more noisy than the one obtained from particles due to the lower number of halos in comparison to particles. We see that results obtained in the 2-fluid simulations are fully consistent with the ones from 1-fluid simulations. This once again reflects the smallness of the impact of baryon-CDM perturbations on galaxy clustering at low redshift and implies that these effects will probably not need to be included in the modeling of correlation functions for the analysis of future surveys BAO peak estimation. This has a positive impact for such analysis since it will reduce the number of free parameters entering the model. These results are in line with previous results in the literature: [73, 12] estimated that the impact of baryon-CDM perturbations on the late-time halo power spectrum should not exceed 1 – a few percent; [72] conducted an analysis of the BOSS DR12 data with a model including baryon-CDM relative density and velocity perturbations, and obtained results for the

2fluid	redshift	CDM	baryon	δ_{bc}	total matter	halo (12.45 < logM < 12.95)
	$z = 0$	102.0 ± 0.9	102.0 ± 1.0	102.3 ± 5.7	102.0 ± 0.9	101.2 ± 3.5
	$z = 0.5$	102.8 ± 1.5	102.8 ± 1.5	99.7 ± 7.5	102.9 ± 1.5	101.4 ± 1.8
	$z = 1$	102.9 ± 1.2	103.0 ± 1.1	105.2 ± 2.8	102.9 ± 1.2	104.5 ± 4.9
	$z = 1.5$	102.9 ± 0.9	102.9 ± 0.8	106.5 ± 2.0	102.9 ± 0.9	104.9 ± 2.3
	$z = 2$	102.7 ± 0.7	102.8 ± 0.7	104.7 ± 1.5	102.7 ± 0.7	107.4 ± 2.1
	$z = 3$	102.5 ± 0.5	102.5 ± 0.5	106.2 ± 2.6	102.5 ± 0.5	105.9 ± 2.8
1fluid	redshift	CDM	baryon	δ_{bc}	total matter	halo (12.45 < logM < 12.95)
	$z = 0$	–	–	–	102.4 ± 1.0	99.1 ± 9.7
	$z = 0.5$	–	–	–	103.0 ± 1.8	102.2 ± 1.8
	$z = 1$	–	–	–	102.9 ± 1.2	101.9 ± 3.5
	$z = 1.5$	–	–	–	102.8 ± 0.8	104.9 ± 1.9
	$z = 2$	–	–	–	102.7 ± 0.7	105.6 ± 2.1
	$z = 3$	–	–	–	102.5 ± 0.5	106.4 ± 1.9

Table 6.2: Position of the BAO peak of the halo and matter fields in 1-fluid and 2-fluid simulations for different redshifts. In the case of 2-fluid set, we also compute the position of the peak for CDM, baryons, and the δ_{bc} fields separately. We see that any shift in the peak position is within 1σ errorbars.

bias parameters associated with such perturbations consistent with zero, indicating an effect too small to be detected; finally, using 2LPT, [61] showed the effect to be at most one order of magnitude smaller than the halo 2PCF itself.

4.2 Position of the BAO peak

We now focus more specifically on the position of the BAO peak estimation for our two sets of simulation. [285] showed that the position of the BAO linear point, namely the midpoint scale between the peak and the dip of the 2PCF, can be extracted from the 2PCF measured in N-body simulations or galaxy data sets in a model-independent way by introducing a polynomial function to smooth the 2-point correlation function, and using a root-finding algorithm to estimate the zero-crossing of the first derivative of the 2PCF. To measure the linear point one needs to estimate the position of the BAO peak as well as the BAO dip through this modeling, but here we will just focus on the maximum of this fit. We use the following polynomial fit

$$\xi^{\text{fit}}(r) = \sum_{n=0}^{N} a_n r^n. \tag{6.8}$$

Following [285], we obtain the best fit parameter for the degree of the polynomial N by minimizing the χ^2. We use scales in the range $85 - 115\,h^{-1}\,\text{Mpc}$, and we choose $N = 7$, which allows us to obtain good fits in the sense that the reduced χ^2 is close to 1 for all correlation functions we consider here while avoiding overfitting. We have also checked that the results for the position of the BAO peak depend only weakly on the degree of the polynomial (for example, the results for the matter-matter correlation function are consistent for polynomials of degree 4 to 8). Having the polynomial fit, to identify the peak position, we find the point the fit where the first derivative of the 2PCF is equal to zero, and the second derivative is negative. Fig. 6.11 illustrates this process by showing the matter-matter, CDM-CDM and baryon-baryon correlation functions, and the related position of the BAO peak in each case (dotted-dashed vertical line with 1σ error) at $z = 0$. Each time the solid line shows the measurement while the dashed line shows the fit. We see that the position of the peak extracted from the baryon-baryon and CDM-CDM 2PCFs align with each other and with the total matter one in 2-fluid simulations. The position of the peak in 1-fluid simulation is slightly higher but the difference between the two cases lies within the 1σ errorbars. As we explained in the discussion of Fig. 6.9, this is expected since the BAO feature originates in the baryon field through baryon oscillations sourced by photon pressure, and then is imprinted into the CDM field with the same position but a lightly skewed distribution towards higher values. This results is in a slightly overestimated position of the peak when assuming that the two fluids perfectly comove as is done in 1-fluid simulations.

The values of the position of the BAO peak for each fluid and for several redshift are reported in Tab. 6.2 in details. Since the position of the BAO peak in all cases remains the same within errorbars (at least with the 8 realizations that we used here), we can argue that the BAO peak remains a standard ruler even in the presence of baryon-CDM perturbations. Notice that to decrease the errorbars by at least a factor of 5, we would need at least 900 realizations of each types of simulations but this would still not assure that we would see any significant differences.
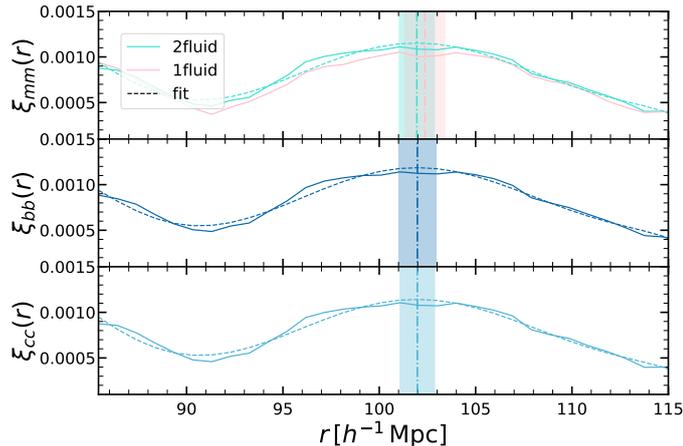
Figure 6.11: **Top panel:** Comparison between the matter-matter 2PCF in 1-fluid (pink) and 2-fluid (cyan) simulations. **Middle and bottom panels:** baryon-baryon and CDM-CDM 2-point correlation functions respectively in 2-fluid simulations. Each time we show the 2PCF in real space at $z = 0$, computed using Eq. 6.7. The shaded area show the 1σ error over 8 realizations in each simulation. The vertical dotted-dashed lines show the position of the BAO peak obtained from a 7^{rmth} degree polynomial fit of the form of Eq. 6.8. The polynomial fit is plotted in each case with the dashed line style and the same color for each type of correlations. The position of the BAO peak in each case for all different redshifts are presented in Tab. 6.2.

Using the results in Fig. 6.10, we also computed the BAO peak position in the halo-halo 2-point correlation function for 2-fluid and 1-fluid simulations for the high mass bin. The results are shown in the last column of Tab. 6.2. As was already visible in the left panel of Fig. 6.10, the positions of the peak are compatible within errorbars between the two cases.

Finally, we compute the position of the BAO feature in the bc-bc cross-correlation function from Fig. 6.9. In this case we use the same polynomial fitting formula (Eq. 6.8) but looking now for the minimum of our fit. As we mentioned before, results at low z become noisy which is why the errorbars on the peak position increase. The results are shown in the fourth column of Tab. 6.2. We do observe a somewhat higher value of the scale of the BAO dip with respect to that of the BAO peak of all other fields we consider (except halos), which is expected as explained before.

To conclude, from Tab. 6.2, we do not detect any significant impact of relative baryon-CDM perturbations on the BAO peak position measured from the matter or halo correlation function. This is in line with results from the previous section where we found no evidence for a change in the broadband correlation function from such perturbations. This is also again in line with previous results from [72] who found no evidence for nonzero bias parameters associated to these perturbations from the BOSS galaxy power spectrum. Furthermore, [73] also forecasted that the BAO peak position should be shifted by less than 1% for halo samples similar to the one we consider here (their section 4).

We end this section by a small word of caution. In this work we only considered the effects of baryon-CDM relative perturbations generated by baryon-photon coupling prior to recombination. However, as we already mentioned in the introduction, compensated isocurvature perturbations (CIP) can also be generated in some Inflation scenarios. As was discussed in [15, 286], such CIPs can also locally affect the position of the BAO peak or the galaxy power spectrum, and these statistics could hence be used to constrain them as well as inflationary scenarios. A direct measurement of the impact of CIPs on the BAO peak position could be done using 1-fluid separate universe simulations as described in [73, 12], but this is beyond the scope of this work.

7

Concluding remarks

In Chapters 4, 5 and 6, we described the main results of this thesis. Here we give a brief summary and draw our conclusions (Sec. 1) including a future outlook on the main lines of investigations we would pursue in the near future (Sec. 2).

1 Summary and Conclusions

Over the past decades, we have witnessed dramatic progress in exploring the cosmos. A wealth of cosmological observations has allowed to establish the concordance ΛCDM model. Nevertheless, the main components of this model still pose a challenge for modern cosmology due to the lack of understanding of their nature. There are still plenty of open issues to be addressed. Planned surveys in the near future will probe structure formation with unprecedented precision to constrain cosmological parameters and deepen our understanding of the Universe. Moreover, cosmological N-body simulations play a crucial role in connecting the initial perturbations, analytically predicted by theory and measured in the CMB, to the collapsed structures in the late time Universe we observe today. With increasingly larger and more accurate surveys in the future, the need for more detailed and accurate modeling rises in order to estimate covariances, study systematics, and test data pipelines.

In this thesis, we focused on some of the challenges mentioned above, depicted by the behavior of the dark components of the Universe, energy, and matter relative to the baryons.

In this context, the first step that we took in this thesis has a general phenomenological study of the DE perturbation dynamics, by considering a perturbed EDE models, which involve sound speed and anisotropic stress, and implemented in the latest version of the Boltzmann equation solver CAMB [4]. The focus has been on a quantitative exploration of an extended parameter space, obtained by considering simultaneous variation of 12 cosmological parameters (including perturbed EDE parametrization as well as the mass of neutrinos and tensor to scalar ratio). We used the publicly available MCMC package COSMOMC [90], exploiting several of the most important and recent data sets: the 2018 CMB data from Planck in combination with other astrophysical data sets like BAO, Type Ia SNe, the Hubble diagram, Quasar and the Lyman-α forest data sets. The main findings of this study can be summarized as follows:

- By considering this generalized context, we found $\Omega_{\mathrm{eDE}} < 0.0039$ at 95% confidence level when Planck, BAO, SNe and the H_0 prior are considered and in particular EDE perturbations are also included in the modelling.

- By fixing the perturbation parameters to the default values returned $\Omega_{\mathrm{eDE}} < 0.0034$ at 95% confidence level, we have thus a $\sim 15\%$ tighter limit for the Ω_{eDE} parameter, in comparison with the case in which perturbations parameters are varied.

- We also found that c_{eff}^2 and c_{vis}^2, and their possible departure from standard values, respectively 1 and 0, for a minimally coupled scalar field, are not yet constrained by present data.

- In addition, we explored some cases with fixed values of Σm_ν and r and found that the upper limits on Ω_{eDE} and w_0 were remarkably stable even when Σm_ν and the r parameter are varied. We also observed a moderate degeneracy Σm_ν or r and Ω_{eDE}.

- We also compared the inferred cosmological parameters within the EDE scenario with those obtained in a model with wCDM plus perturbations, and with a constant equation of state and found a non negligible impact on the inferred upper limits on the total neutrino mass and DE equation of state.

- Moreover, we investigated the impact of EDE as well as the perturbations on cosmological observables such as: CMB power spectrum, the linear matter power spectrum, ISW contribution and the CMB lensing potential. We found that in the absence of EDE component in the early Universe, the effect of perturbations are only confined at relatively large scales due to the ISW effect while when the EDE component is present, the effect is also visible on smaller scales due to the fact that EDE influences directly the recombination process so the EDE can affect on the evolution of the acoustic oscillations before recombination which happens at higher ℓs (smaller scales).

- In the absence of an EDE component, the effects of perturbation components on the matter power spectrum is mostly seen at large scales which are affected by cosmic variance. On the other hand, by including the EDE component, the effect of perturbations would be distinguishable at smaller scales where non-linearities are present in the matter power spectrum.

- We also observed that in absence of perturbation parameters, since there is no friction in the growth of perturbations, the lensing potential is significantly enhanced with respect to the case in which we include perturbation parameters ($\sim 45\%$ in presence of EDE and 10% in the absence of EDE component, each of them w.r.t. the case when perturbations are included).

This work demonstrated the capability of CMB and LSS probes to constrain the total amount of EDE well below the percent level. The detailed explanations and the results of this study was presented in Chapter 4.

At this point, we considered the physics of the matter components of the Universe in order to obtain a direct estimate of gravitational coupling of baryons and CDM. We relied on N-body simulations to quantify the impact of relative baryon-CDM density δ_{bc} perturbations on large-scale structures. Before recombination, baryons and photons were coupled together, so that baryons and CDM were not co-move and they had different evolution. Even if the gravitational evolution slowly reduces this difference (since baryons can fall in CDM potential wells), we found that this effect is not totally negligible at $z = 0$ and can reduce the clustering of LSS. We assessed the impact of baryon-CDM perturbations by performing gravity-only 2-fluid cosmological N-body simulations where the baryons and CDM fluids are initialized with different transfer functions. We generated our initial conditions using the MUSIC code [243] and performed our simulations using the GADGET-II code [245]. Performing these types of gravity-only 2-fluid simulations was not a trivial task, since the N-body set-up contains two types of particles with different masses and a high force resolution for the mass resolution would cause a spurious coupling between baryons and CDM. In order to alleviate this, we used an AGS scheme for baryons that has been implemented in the Gadget code by [68] to retrieve agreement with linear theory on large scales. We identified halos using AHF [119, 129] which relies on as spherical over-density criterion in the spatial distribution of particles in the simulations. In this study, we focused on the baryon fraction in halos F_b as a function of halo mass and local baryon-CDM relative density, and the cross-correlation between the halo field and the baryon-CDM perturbation field. We also measured the associated baryon-CDM bias parameter $b_{\delta_{bc}}$. We also assessed the impact of δ_{bc} on the halo power spectrum for a Euclid-like galaxy sample and a DESI quasar one and compared it to that of massive neutrinos. Our main findings can be summarized as follows:

- We found that the baryon fraction in halos was slightly smaller than the universal one and weakly dependent on the halo mass. For masses $5 \times M > 10^{12} h^{-1}(M_\odot)$ we found it to be larger than 95% the cosmic mean with relatively small scatter. We measured a noticeable downturn in lower mass bins for both our 2-fluid-diff-TF and 2-fluid-same-TF simulations, which we ascribe to our softened forces with AGS.

- In order to further study the scatter in F_b we looked at its correlation with the local large-scale baryon-CDM relative density, and found it to be small . This confirmed that the large scatter in F_b at low halo mass was rather unphysical (i.e. δ_{bc}^R only weakly affects the measured F_b), and was rather due to numerical effects of poorly resolved halos.

- The halo-baryon-CDM cross-spectra $P_{h\,bc}$ has been shown to be non-zero, demonstrating that baryon-CDM perturbations affect the clustering of structures even at low redshifts. P_{hbc} and P_{mbc} were negative reflecting the anti-correlation between δ_{bc} and δ_h, or δ_{bc} and δ_m. P_{bcbc} had the smallest value in comparison to the other power spectra but it is nonzero. To our knowledge this was the first time such correlations were measured.

- Our results for $b_{\delta_{bc}}$ from the 2-fluid simulations were in agreement with those from 1-fluid separate Universe simulations. It was an important cross-check since the two methods are completely independent and this parameter was only measured once before in [242]. This parameter was negative reflecting again the anticorrelation between δ_{bc} and halo formation.

- Even though we neglected the effect of relative velocities to measure $b_{\delta_{bc}}$ from 2-fluid simulations we found perfect agreement with previous results within the error-bars. This has confirmed that the term proportional to θ_{bc} was indeed sub-dominant in the bias expansion.

- We found a linear relation between $b_{\delta_{bc}}$ and b_1 at all z and provided a fit. We found the contribution of terms proportional to $b_{\delta_{bc}}$ to the halo power spectrum to be at maximum 0.3% at $k = 0.1\,h\,\mathrm{Mpc}^{-1}$ at $z = 0$. We compared this to the effect of massive neutrinos for a Euclid-like galaxy sample at $z = 1$ and a DESI-like quasar one at $z = 3$, finding the impact of δ_{bc} to be dominant in the latter case which is a fundamental point for future surveys.

To sum up, the 2-fluid numerical simulations have proven to be an effective tool to study the impact of baryon-CDM relative perturbations on LSS. Several works have now worked out near optimal setups for these simulation [65, 10] and have opened the way to study how various LSS observables are affected. While we showed that halo clustering is only weakly affected and that measuring the relative velocity bias $b_{\theta_{bc}}$ would require a very large simulation volume. We also showed that the baryon fraction in halo deviates from the cosmic mean by a non-negligible amount even with respect to late-time hydrodynamical effects for some halo masses. Finally we also showed that the effect of δ_{bc} could be degenerate with that of neutrinos for some future surveys. A complete description and the results of this study can be found in Chapter 5.

Following the previous project, generalizing our 2-fluid simulations mentioned above, we assessed the impact of baryon-CDM density perturbations on some very promising cosmological probes: cosmic voids and the position of the BAO peak. This project indeed had a two-fold path, the first one on the investigation of the cosmic voids and the second on baryon acoustic oscillations, in which we studied them using 2-fluid simulations and compared the results with those obtained in a standard gravity-only 1-fluid simulation. We highlighted that in this project we worked in configuration space, in contrast with the previous one where we worked in the Fourier space. Thus, considering voids, we studied the impact of baryon-CDM relative perturbations due to photon pressure prior to recombination on void statistics, void-tracer cross-correlations, the density profile of voids and their linear bias. Regarding BAOs, we calculated the 2-point correlation function and position of the BAO peak in real space of various fluid components. In order to identify voids in our simulations we used the publicly available REVOLVER void finder to build our void catalogues with the ZOBOV algorithm [9], which is a 3D void finder and has been widely used in simulated and observed catalogues. We identified the voids in two different density tracer types, DM halos and particles. We computed void size function, void-tracer cross-correlations, the density profile of the voids and the linear void bias in both different tracer types. The main findings of our studies here can be summarized as follows:

- The VSF depends strongly on the tracer used to identify voids (there are more small voids and less large ones in the particle field than in the halo field). The VSF of particle field voids is unaffected by baryon-CDM relative perturbations, while the VSF of halo field voids is affected at $1 - 2\%$ level: smaller voids are more abundant in presence of such perturbations and larger voids less, which is a consequence of the fact that these perturbations act against clustering (chapter 6, Fig. 6.1).

- We did not detect any statistically significant impact of baryon-CDM relative perturbations on the void, matter or halo auto- and cross- 2PCF. We found hints that these perturbations diminish the clustering on scales smaller than the BAO one, and enhance the BAO peak amplitude (chapter 6, Fig. 6.2 and Figs. 6.8 and 6.10), which is in agreement with our expectations.

- The density profiles of voids in halo and particle fields display the three known regimes (negative deep inside the void followed by the void profile regime with the positive compensation wall, and the linear regime where the halo-void correlation function becomes zero), and voids in the halo field are larger on average. We found no significant impact of baryon-CDM relative perturbations on any of the profiles, but a hint for voids in 2-fluid simulations to be emptier (chapter 6, Figs. 6.3 and 6.5).

- The void bias depends significantly on the tracer used to find voids (the bias is almost constant over void size for halo field voids but it decreases for larger voids in the particle field), but we fond consistent results for bias obtained from cross- and auto- correlation functions. Again we did not find any significant difference for the bias in 1- and 2-fluid simulations, but found hints that it is slightly larger in the latter case, as we expect (chapter 6, Fig. 6.7).

- The amplitude of the BAO peak in the baryon 2PCF decreases with time due to gravitational evolution. It is gradually imprinted in the CDM 2PCF where the amplitude of the peak grows down to $z \sim 7$ and then decreases down to $z = 0$ due to nonlinear effects (chapter 6, Fig. 6.9).

- The relative density perturbation δ_{bc} auto-correlation function presents a dip as BAO feature on scales slightly larger than the BAO peak, which is consistent with the fact that on these scales CDM particles lag behind baryons (chapter 6, Fig. 6.9).

- We directly measured the impact that baryon-CDM perturbations have on the BAO peak position of halo and matter for the first time to our knowledge, and found no evidence for a statistically significant impact (chapter 6, Fig. 6.11 and Tab. 6.2), which is in agreement with previous works ([72]).

The halo field VSF is the only quantity that we found to be affected with statistical significance by baryon-CDM relative perturbations due to photon pressure prior to recombination. This effect might hence also affect the VSF of voids obtained from galaxy fields in observational data, and this statistics could hence be used to constraint such perturbations. We note however that the effect remains quite small. Our results for the matter-matter and halo-halo 2PCF added to ones from previous works confirm that the impact of baryon-CDM perturbations on cosmological constraints from the BAO feature in current and future galaxy surveys should be negligible at low redshift ($z \leq 3$). This has important consequences for future galaxy clustering surveys since it means that these effects will not have to be included in the modeling of leading-order quantities used for the analysis of their data. The results of this study are presented in Chapter 6.

2   Future perspectives

Each of the three main projects presented in this thesis provide several interesting and promising possibilities for future scientific exploitation.

- One of the main targets would be to consider the high redshift regime, both in terms of data available and, theoretically, for models that do present an impact at such early epochs. The motivation comes from the fact that BAO measurements have been made in several large samples of galaxies at different redshifts, including six-degree-Fields Galaxy survey (6dFGS) [210], at $z = 0.106$, SDSS [32, 33, 34, 35] at $z = 0.15 - 0.6$, and DES yr 1 [287] at $z = 0.81$. At higher redshifts, the BAO are measured using the correlation of Lyman-α forest flux [213, 214, 215, 216] at $z = 1.5$, $z = 2.33$ and $z = 2.4$. The expansion rate measured from Lyman-α BAO is currently in moderate tension with the standard model (see Fig. 7.1). Therefore it is essential to measure the expansion rate at higher redshifts precisely and compare with the existing measurements to see if the tension remains or not. This could be possible in the light of e.g. Lyman-α emitting galaxy data from experiments like HETDEX[1][288, 289]. One interesting possibility would be to measure the BAO scale and also to compute accurate constraints on the EDE models relying on such high-z data sets.

- In these context, we would need to consider the interplay between reionization, neutrinos and DE. Reionization is believed to have occurred when the first generations of stars and quasars produced sufficient amounts of UV and X-ray radiation to ionize the vast majority of neutral hydrogen in the Universe. We have a limited understanding of how and exactly when cosmological reionization happened based on current observational data. The Epoch of Reionization (EoR), mentioned briefly in chapter 3, Sec. 1, is the key event to understand the Inter-Galactic Medium (IGM) evolution and subsequent structure formation. The reionization history often assumed in CMB analyses has a parametric form that matches physical expectations reasonably well, involving a single smooth step from an almost fully neutral Universe. The free electron fraction, x_e, is modelled as a tanh function and describes the EoR as a step-like transition between an essentially vanishing ionized fraction x_e at early times, to a value of unity at low redshifts. The optical depth to reionization is defined by

$$\tau = \int_0^{\eta_0} d\eta \, a n_e^{reion} \, \sigma_T \,, \tag{7.1}$$

in which n_e^{reion} represents the number density of free electrons produced by reionization at conformal time η, η_0 is the conformal time today, σ_T is the Thomson scattering cross-section and a is the scale

---

[1]The Hobby-Eberly Telescope Dark Energy Experiment is a unique spectroscopic survey that will map the evolution of Dark Energy using Lyman-α emitting galaxies in the high redshift Universe ($1.9 < z < 3.5$).
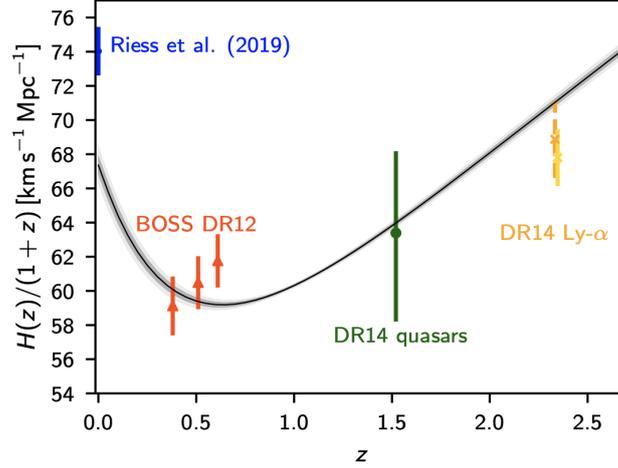
Figure 7.1: Measurements of the expansion of the Universe from standard candles (data points at $z \approx 0$) and the BAO standard ruler in the galaxy distribution (points at higher redshifts). The line shows the best-fitting ΛCDM model to the CMB and BAO measurements at $z < 1$. HETDEX will precisely measure BAO at redshifts $1.9 < z < 3.5$ and can add new data points to this plot, with a few percent error at $z \sim 2.2$ and 2.7, being crucial in order to see if at these redshifts the measurements of the expansion of the Universe is compatible with the concordance ΛCDM model of cosmology or not. Credit: [2].

factor. n_e is parameterized as $n_e \propto (1+z)^3 x_e(z)$, in which x_e represents the number of free electrons per hydrogen atom. Using the fact that reionization is expected to happen during matter domination, we have

$$\tau \propto \int dz\, x_e \sqrt{1+z} \propto \int d[(1+z)^{3/2}] x_e \,, \tag{7.2}$$

and in order to parameterize x_e as a function of $y \equiv (1+z)^{3/2}$, we have

$$x_e(y) = \frac{f}{2}\left[1 + \tanh\left(\frac{y(z_{re}) - y}{\Delta y}\right)\right], \tag{7.3}$$

where $y(z_{re}) = (1 + z_{re})^{3/2}$ is the value for which $x_e = f/2$, i.e. z_{re} measures where the reionization fraction is half of its maximum, while Δy is taken to be $1.5\sqrt{1 + z_{re}}\Delta z$.

The idea is to use a tomographic analysis of reionization [13] with error bars from Quasars that are comparable with CMB. We made the first steps in this project, and we modified the reionization part of the CAMB code [4] and added the reionization history from [13] which is more physical and in better agreement with the reionization process probed by Quasars. The electron fraction (for both the tanh model and the one inferred data) as a function of redshift/scale factor is illustrated in Figure 7.2. The TT, EE and TE CMB specta from tanh reionization history as well as the modified one are shown in Figure 7.3. Preliminary results obtained using the COSMOMC package [90] refer to the following cosmologies: ΛCDM, ΛCDM+Σ_{m_ν}, ΛCDM+$N_{\rm eff}$, ΛCDM+$\Sigma_{m_\nu} + N_{\rm eff}$. These results exploites the following data: high-ℓ Planck 2018 CMB data, low-ℓ Sroll-2 CMB likelihood, SPT-high-ℓ 2020 data, BAO dataset, Lyman-α- Quasar data and SNe data-set. To further investigate the effect of the optical depth and estimate the cosmological parameters in light of priors over reionization history. We aim at extending these results and explore the EDE scenario and the $N_{\rm eff}$ simultaneously [60]: an extremely light axion-like field, obtaining a mass through a periodic potential, decaying as fast or faster than radiation which in turn could solve the Hubble tension.

- More specifically along the lines of the works presented in this thesis, new EDE models [60] could be investigated with the same methodology developed in our first project. The latter is one of the proposals to solve the H_0 tension by decreasing the physical size of the sound horizon imprinted in the CMB and increasing the inferred H_0. Previous EDE analyses [60, 290, 291] have shown that this model can alleviate the Hubble tension, but the late-time amplitude of density fluctuation σ_8 increases as compared to ΛCDM, worsening the tension with LSS data [292, 293]. In Figure 7.4, we presented the behaviour of
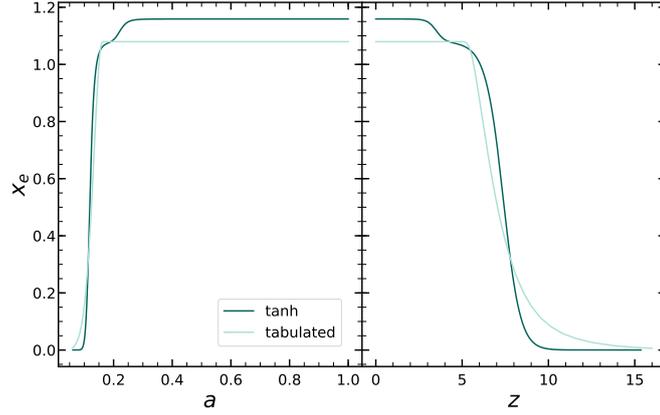
Figure 7.2: Number of free electrons per Hydrogen atom as a function of scale factor/redshift (left/right panel). Reionization histories with $\tau = 0.0532$, $z_{re} = 7.365$ and $\Delta z = 1$, the dark curve shows the tanh model while the light curve shows the result of the modified CAMB code [4] with implementation of the reionization data from [13]. The data from [13] are indeed the volume-weighted ionized hydrogen fraction (x_{HII}) in different redshifts from a tomographic analysis of reionization with error bars from Quasars that are comparable to CMB [14], which can be calculated as: $x_e = (x_{\mathrm{HII}} \times n_{\mathrm{H}} + x_{\mathrm{HII}} \times n_{\mathrm{He}})/n_{\mathrm{H}}$ where n_{H} and n_{He} are the average H and He number densities.
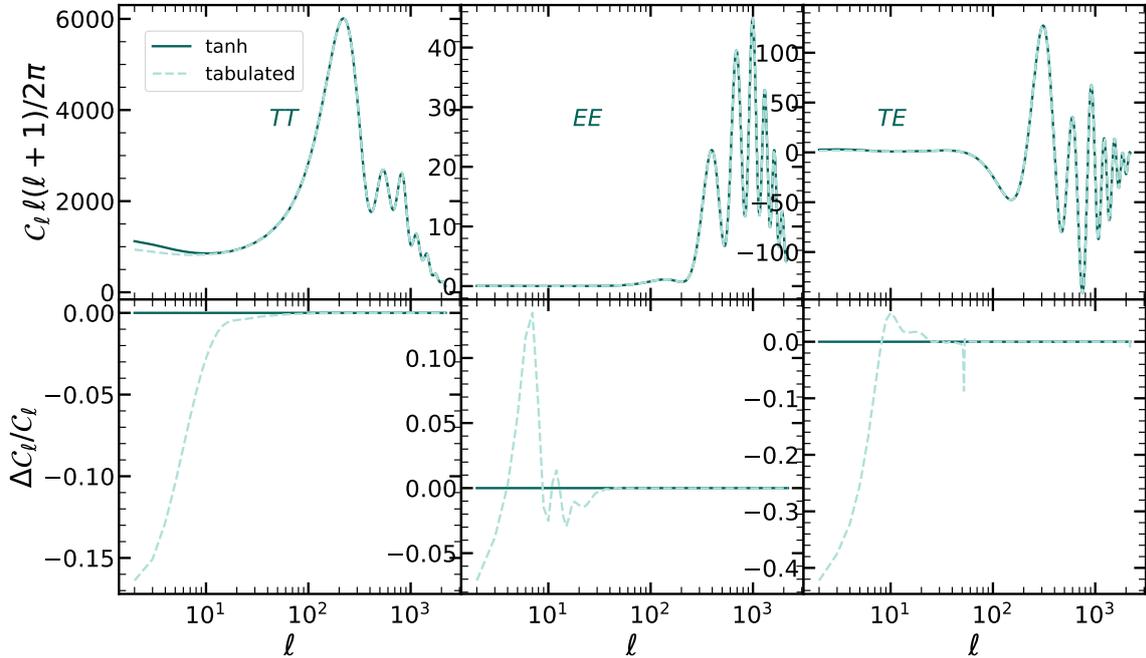
Figure 7.3: TT,TE and EE CMB spectra with the same $\tau = 0.0532$, $z_{re} = 7.365$ and $\Delta z = 1$, The light-dashed curve in each panel shows the spectra from the modified reionization by using the data from [13]. The dark-solid curve shows the spectra from the tanh parametrization. The z_{re} and Δz values in the modified reionization (light-dashed) have been chosen in such a way that get the same τ as the tanh reionization. The lower panel illustrate the $(\mathcal{C}_{\ell,\mathrm{tabulated}} - \mathcal{C}_{\ell,\mathrm{tanh}})/\mathcal{C}_{\ell,\mathrm{tanh}}$ in each case.

H_0's 1σ and 2σ confidence levels in EDE and ΛCDM models in the light of Planck 2018 [204] and SPTPol 2018 [294] CMB data sets. Here we used the EDE modified version of the CLASS code [93], which has been modified by [292] and the MONTE PYTHON MCMC package [92]. The idea is to exclude the Planck temperature power spectrum at high multipoles and combine the rest by polarization and lensing measurements of SPTPol likelihood [295], in order to remove the lensing tension in the Planck spectra on small scales in order to disentangle early and late time physics [296, 297].
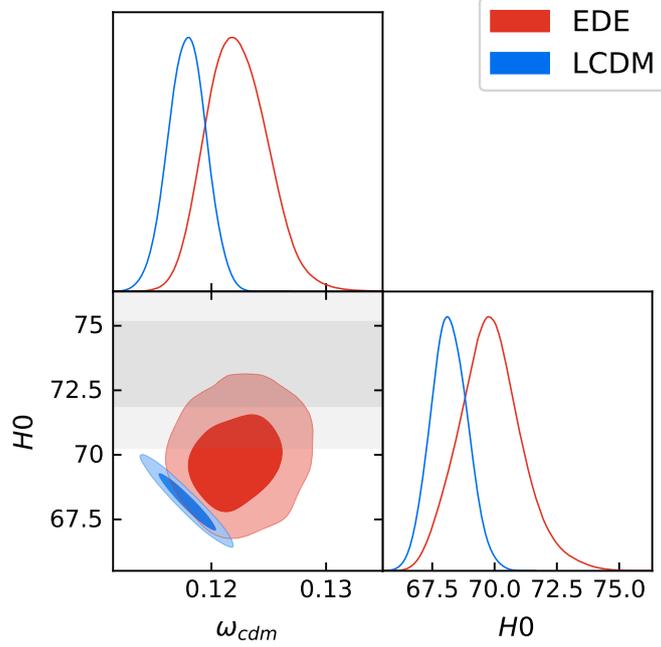
Figure 7.4: Comparing EDE and ΛCDM model using Planck and SPTpol CMB datasets. This figure illustrates the parameter space of H_0 and CDM density (ω_{cdm}). The blue contour shows the 1-σ and 2-σ confidence level for H_0 and CDM density in the ΛCDM model. The red contour demonstrate the same region for EDE. The grey shaded area shows the local measurements of H_0 which are in tension with ΛCDM (the blue contour). EDE (red contours) can weaken the same tension. Finally, in the upper/right panel, the posterior distribution for the CDM density/H_0 cosmological parameters is shown.

- Also, by combining the analyses and methodologies concerning the second and third projects, one interesting direction would be utilizing the 2-fluid simulations we performed in the second and third projects with a large box size $500\,\mathrm{h}^{-1}\,\mathrm{Mpc}$ on each side in order to measure the relative velocity biases ($b_{\theta_{bc}}$ and b_{v^2} terms in Eq. 7.4). In the second project in chapter 5 we measured the $b_{\delta_{bc}}$ and our future goal would be measuring the velocity terms using the 2-fluid simulations in the larger box size:

$$\delta_g(\mathbf{x}, z) = b_1(z)\delta_m(\mathbf{x}, z) + b_{\delta_{bc}}(z)\delta_{bc}(\mathbf{x}) + b_{\theta_{bc}}(z)\theta_{bc}(\mathbf{x}, z) + b_{v^2}^{bc}(z)[\mathbf{v}_{bc}^2 - \langle \mathbf{v}_{bc}^2 \rangle(z)]. \qquad (7.4)$$

However, we note that the velocity term is the subdominant term and would decay faster than the δ_{bc} term.

- Along this lines, it would be also interesting to study Compensated Isocurvature Perturbations (CIP)s, which are opposite spatial fluctuations in the baryon and CDM densities that can be generated during inflation [298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310]. A noteworthy aspect of CIPs is that they exhibit strong BAO features that are not completely in phase with those imprinted in the total matter fluctuations. This, in turn, can lead to a shift in the BAO scales imprinted in the galaxy distribution which can be measured by the oscillatory behaviour of $P_{m\,\delta_{bc}}$ in the Fourier space and potentially would lead to a shift in the 2-point correlation function in the configuration space. To assess CIPs we need to perform separate Universe simulations, with the ansatz stating that local structure formation inside a long-wavelength perturbation in some fiducial cosmology is equivalent to global structure formation in an appropriately modified cosmology (see [15]). For the case of CIPs, the modification to the cosmological model would include a change in the background density of Ω_b and

Ω_c, while keeping the total matter density the same, $\Omega_m = \Omega_b + \Omega_c$. Figure 7.5 represents the effect of the large-scale CIPs on the BAO in the correlation function in these hypotheses. In cosmologies with $\tilde{\Omega}_b = [1 + \Delta_b]$ and $\tilde{\Omega}_c = [1 - f_b \Delta_b]$, we can measure a possible shift in the BAO peak.
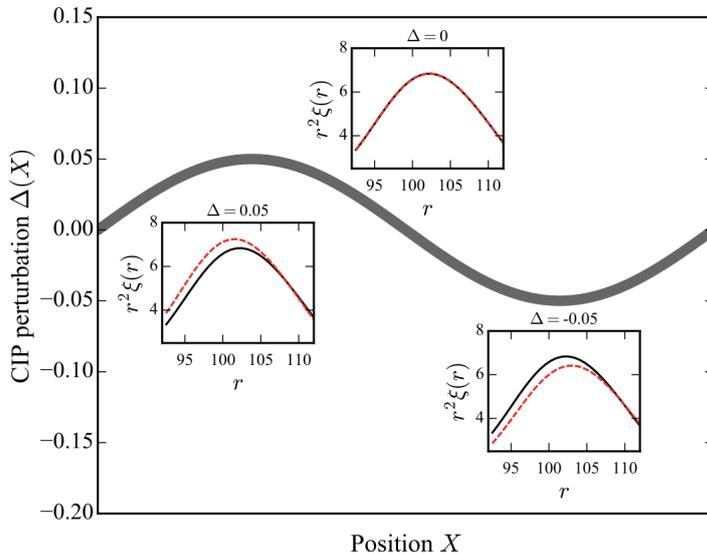
Figure 7.5: The illustration of the CIPs $\Delta(x)$, on the BAO peak regarding the separate Universe limit. The dashed-red curves show the behaviour including the CIPS and the solid-black curves are the results without having CIPs. We see that in the cosmologies with higher values of Δ, the BAO peak would shift to the left and in cosmologies with lower values we see the opposite effect. This effect can be measured directly using the separate Universe simulations. From [15].

- Indeed it would be also interesting to study the impact of CIPs generated during Inflation on the voids statistics, using the aforementioned separate Universe simulations described in [73].

The items above represent our main ideas concerning the improvements and new ideas for exploiting the results and methodologies presented in this thesis. We think that the lines we mentioned, both in terms of proposals for future observations and investigations following the works presented here, would need to be undertaken for comparison with forthcoming observations.

Bibliography

[1] Antonio J. Cuesta et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations in the correlation function of LOWZ and CMASS galaxies in Data Release 12. *Mon. Not. Roy. Astron. Soc.*, 457(2):1770–1785, 2016.

[2] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020.

[3] W. L. Freedman et al. Final results from the Hubble Space Telescope key project to measure the Hubble constant. *Astrophys. J.*, 553:47–72, 2001.

[4] A. Lewis, A. Challinor, and A. Lasenby. Efficient computation of CMB anisotropies in closed FRW models. *The Astrophysical Journal*, 538:473–476, 2000.

[5] S. Dodelson and F. Schmidt. *Modern Cosmology.* Elsevier Science, 2020.

[6] Florian Beutler, Hee-Jong Seo, Ashley J. Ross, Patrick McDonald, Shun Saito, Adam S. Bolton, Joel R. Brownstein, Chia-Hsun Chuang, Antonio J. Ĉuesta, Daniel J. Eisenstein, Andreu Font-Ribera, Jan Niklas Grieb, Nick Hand, Francisco-Shu Kitaura, Chirag Modi, Robert C. Nichol, Will J. Pẽrcival, Francisco Prada, Sergio Rodriguez-Torres, Natalie A. Roe, Nicholas P. Ross, Salvador Salazar-Albornoz, Ariel G. Sánchez, D̃onald P. Schneider, Anže Slosar, Jeremy Tinker, Rita Tojeiro, Mariana Vargas-Magaña, and Jose A. Vazquez. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Fourier space. *Monthly Notices of the Royal Astronomical Society*, 464(3):3409–3430, 09 2016.

[7] N. Y. Gnedin, S. C. O. Glover, R. S. Klessen, and V. Springel. *Star Formation in Galaxy Evolution: Connecting Numerical Models to Reality.* Springer, 2015.

[8] Darren J. Croton, Volker Springel, Simon D. M. White, G. De Lucia, C. S. Frenk, L. Gao, A. Jenkins, G. Kauffmann, J. F. Navarro, and N. Yoshida. The Many lives of AGN: Cooling flows, black holes and the luminosities and colours of galaxies. *Mon. Not. Roy. Astron. Soc.*, 365:11–28, 2006. [Erratum: Mon.Not.Roy.Astron.Soc. 367, 864 (2006)].

[9] Mark C. Neyrinck. ZOBOV: a parameter-free void-finding algorithm. *MNRAS*, 386(4):2101–2109, June 2008.

[10] Oliver Hahn, Cornelius Rampf, and Cora Uhlemann. Higher order initial conditions for mixed baryon–CDM simulations. *Mon. Not. Roy. Astron. Soc.*, 503(1):426–445, 2021.

[11] Fabian Schmidt. Effect of relative velocity and density perturbations between baryons and dark matter on the clustering of galaxies. *Phys. Rev. D*, 94(6):063508, 2016.

[12] Hasti Khoraminezhad, Titouan Lazeyras, Raul E. Angulo, Oliver Hahn, and Matteo Viel. Quantifying the impact of baryon-CDM perturbations on halo clustering and baryon fraction. *JCAP*, 03:023, 2021.

[13] Girish Kulkarni, Laura C. Keating, Martin G. Haehnelt, Sarah E.I. Bosman, Ewald Puchwein, Jonathan Chardin, and Dominique Aubert. Large Lyman-α opacity fluctuations and low CMB τ ... *MNRAS*, 485(1):L24–L28, 2019.

[14] Elisa Boera, George D. Becker, James S. Bolton, and Fahad Nasir. Revealing reionization with the thermal history of the intergalactic medium: New constraints from the lyα flux power spectrum. *The Astrophysical Journal*, 872(1):101, feb 2019.

[15] Chen Heinrich and Marcel Schmittfull. BAO Modulation as a Probe of Compensated Isocurvature Perturbations. *Phys. Rev. D*, 100(6):063503, 2019.

[16] E. P. Hubble. A spiral nebula as a stellar system, Messier 31. *Astrophysical Journal*, 69:103–158, 1929.

[17] Edwin Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the National Academy of Sciences*, 15(3):168–173, 1929.

[18] A. Einstein. The basics of general relativity theory. *Annalen der Physik*, 49:769–822, 1915.

[19] A. Friedman. Über die Krümmung des Raumes. *Zeitschrift für Physik*, 10:377–386, 1922.

[20] A. Friedman. Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes. *Zeitschrift für Physik*, 21:326–332, 1924.

[21] R. ALPHER and R. HERMAN. Evolution of the Universe. *Nature*, 1948.

[22] A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophysical Journal*, 1965.

[23] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, S. S. Mather, J. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson. Structure in the COBE Differential Microwave Radiometer First-Year Maps. *Astrophysical Journal Letters*, 396:L1, 1992.

[24] F. HOYLE and R. J. TAYLER. The Mystery of the Cosmic Helium Abundance. *Nature*, 203(4950):1108–1110, 1964.

[25] P. J. E. Peebles. Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations. *Astrophysical Journal*, 263:L1–L5, 1982.

[26] Douglas Clowe, Marusa Bradac, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, Christine Jones, and Dennis Zaritsky. A direct empirical proof of the existence of dark matter. *Astrophys. J. Lett.*, 648:L109–L113, 2006.

[27] Iu. P. Pskovskii. Light curves, color curves, and expansion velocity of type I supernovae as functions of the rate of brightness decline. *Soviet Astronomy*, 21:675, 1977.

[28] Adam G. Riess, Alexei V. Filippenko, Peter Challis, Alejandro Clocchiatti, Alan Diercks, Peter M. Garnavich, Ron L. Gilliland, Craig J. Hogan, Saurabh Jha, Robert P. Kirshner, B. Leibundgut, M. M. Phillips, David Reiss, Brian P. Schmidt, Robert A. Schommer, R. Chris Smith, J. Spyromilio, Christopher Stubbs, Nicholas B. Suntzeff, and John Tonry. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009–1038, sep 1998.

[29] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and The Supernova Cosmology Project. Measurements of ω and λ from 42 high-redshift supernovae. *The Astrophysical Journal*, 517(2):565–586, jun 1999.

[30] C. L. Bennett, R. S. Hill, G. Hinshaw, M. R. Nolta, N. Odegard, L. Page, D. N. Spergel, J. L. Weiland, E. L. Wright, M. Halpern, Ñ. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. Wollack. First-year wilkinson microwave anisotropy probe (WMAP) observations: Foreground emission. *The Astrophysical Journal Supplement Series*, 148(1):97–117, sep 2003.

[31] Matthew Colless et al. The 2dF Galaxy Redshift Survey: Final data release. 6 2003.

[32] Shadab Alam et al. Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological implications from two decades of spectroscopic surveys at the Apache Point Observatory. *Phys. Rev. D*, 103(8):083533, 2021.

[33] Ariel G. Sánchez, Jan Niklas Grieb, Salvador Salazar-Albornoz, Shadab Alam, Florian Beutler, Ashley J. Ross, Joel R. Brownstein, Chia-Hsun Chuang Antonio J. Cuesta, Daniel J. Eisenstein, Francisco-Shu Kitaura, Will J. Percival, Francisco Prada, Sergio Rodríguez-Torres, Hee-Jong Seo, Jeremy Tinker, Rita Tojeiro, Mariana Vargas-Magaña, Jose A. Vazquez, and Gong-Bo Zhao. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: combining correlated Gaussian posterior distributions. *Monthly Notices of the Royal Astronomical Society*, 464(2):1493–1501, 09 2016.

[34] Jan Niklas Grieb et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological implications of the Fourier space wedges of the final sample. *Mon. Not. Roy. Astron. Soc.*, 467(2):2085–2112, 2017.

[35] Florian Beutler et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Fourier space. *Mon. Not. Roy. Astron. Soc.*, 464(3):3409–3430, 2017.

[36] Vera C. Rubin and Jr. Ford, W. Kent. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *Astrophysical Journal*, 159:379, 1970.

[37] Virginia Trimble. Existence and nature of dark matter in the universe. *Annual Review of Astronomy and Astrophysics*, 25(1):425–472, 1987.

[38] Alan H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D*, 23:347–356, Jan 1981.

[39] Katsuhiko Sato. First-order phase transition of a vacuum and the expansion of the Universe. *Monthly Notices of the Royal Astronomical Society*, 195(3):467–479, 07 1981.

[40] A.D. Linde. A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B*, 108(6):389–393, 1982.

[41] A.A. Starobinsky. Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations. *Physics Letters B*, 117(3):175–178, 1982.

[42] Andreas Albrecht and Paul J. Steinhardt. Cosmology for grand unified theories with radiatively induced symmetry breaking. *Phys. Rev. Lett.*, 48:1220–1223, Apr 1982.

[43] David H. Weinberg, Michael J. Mortonson, Daniel J. Eisenstein, Christopher Hirata, Adam G. Riess, and Eduardo Rozo. Observational probes of cosmic acceleration. *Physics Reports*, 530(2):87–255, Sep 2013.

[44] L. Verde, T. Treu, and A. G. Riess. Tensions between the Early and the Late Universe. *Nature Astron.*, 3:891, 7 2019.

[45] Eleonora Di Valentino, Alessandro Melchiorri, and Joseph Silk. Planck evidence for a closed Universe and a possible crisis for cosmology. *Nature Astron.*, 4(2):196–203, 2019.

[46] Adam G. Riess, Stefano Casertano, Wenlong Yuan, J. Bradley Bowers, Lucas Macri, Joel C. Zinn, and Dan Scolnic. Cosmic Distances Calibrated to 1% Precision with Gaia EDR3 Parallaxes and Hubble Space Telescope Photometry of 75 Milky Way Cepheids Confirm Tension with ΛCDM. *Astrophys. J. Lett.*, 908(1):L6, 2021.

[47] Shahab Joudaki et al. KiDS-450: Testing extensions to the standard cosmological model. *Mon. Not. Roy. Astron. Soc.*, 471(2):1259–1279, 2017.

[48] H. Hildebrandt et al. KiDS-450: Cosmological parameter constraints from tomographic weak gravitational lensing. *Mon. Not. Roy. Astron. Soc.*, 465:1454, 2017.

[49] H. Hildebrandt et al. KiDS+VIKING-450: Cosmic shear tomography with optical and infrared data. *Astron. Astrophys.*, 633:A69, 2020.

[50] Marika Asgari et al. KiDS-1000 Cosmology: Cosmic shear constraints and comparison between two point statistics. *Astron. Astrophys.*, 645:A104, 2021.

[51] T. M. C. Abbott et al. Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing. *Phys. Rev. D*, 98(4):043526, 2018.

[52] S. Joudaki et al. KiDS+VIKING-450 and DES-Y1 combined: Cosmology with cosmic shear. *Astron. Astrophys.*, 638:L1, 2020.

[53] Eleonora Di Valentino et al. Cosmology intertwined iii: $f\sigma_8$ and s_8. *Astroparticle Physics*, 131:102604, 2021.

[54] Erminia Calabrese, Anze Slosar, Alessandro Melchiorri, George F. Smoot, and Oliver Zahn. Cosmic Microwave Weak lensing data as a test for the dark universe. *Phys. Rev. D*, 77:123531, 2008.

[55] C. Wetterich. Phenomenological parameterization of quintessence. *Physics Letters B*, 594(1-2):17–22, 2004.

[56] M. Doran and G. Robbers. Early dark energy cosmologies. *Journal of Cosmology and Astroparticle Physics*, 2006(06):026–026, 2006.

[57] Edmund J. Copeland, M. Sami, and Shinji Tsujikawa. Dynamics of dark energy. *Int. J. Mod. Phys.*, D15:1753–1936, 2006.

[58] P. J. E. Peebles and Bharat Ratra. The cosmological constant and dark energy. *Rev. Mod. Phys.*, 75:559, 2003.

[59] Robert R. Caldwell and Marc Kamionkowski. The Physics of Cosmic Acceleration. *Ann. Rev. Nucl. Part. Sci.*, 59:397–429, 2009.

[60] Vivian Poulin, Tristan L. Smith, Tanvi Karwal, and Marc Kamionkowski. Early Dark Energy Can Resolve The Hubble Tension. *Phys. Rev. Lett.*, 122(22):221301, 2019.

[61] Shi-Fan Chen, Emanuele Castorina, and Martin White. Biased Tracers of Two Fluids in the Lagrangian Picture. *JCAP*, 06:006, 2019.

[62] Naoki Yoshida, Naoshi Sugiyama, and Lars Hernquist. The evolution of baryon density fluctuations in multi-component cosmological simulations. *Mon. Not. Roy. Astron. Soc.*, 344:481–491, 2003.

[63] Ryan M. O'Leary and Matthew McQuinn. The formation of the first cosmic structures and the physics of the z 20 Universe. *Astrophys. J.*, 760:4, 2012.

[64] Raul E. Angulo, Oliver Hahn, and Tom Abel. How closely do baryons follow dark matter on large scales? *Mon. Not. Roy. Astron. Soc.*, 434:1756, 2013.

[65] Simeon Bird, Yu Feng, Christian Pedersen, and Andreu Font-Ribera. More accurate simulations with separate initial conditions for baryons and dark matter. *JCAP*, 06:002, 2020.

[66] Michaël Michaux, Oliver Hahn, Cornelius Rampf, and Raul E. Angulo. Accurate initial conditions for cosmological N-body simulations: Minimizing truncation and discreteness errors. *Mon. Not. Roy. Astron. Soc.*, 500(1):663–683, 2020.

[67] Cornelius Rampf, Cora Uhlemann, and Oliver Hahn. Cosmological perturbations for two cold fluids in ΛCDM. *Mon. Not. Roy. Astron. Soc.*, 503(1):406–425, 2021.

[68] Francesca Iannuzzi and Klaus Dolag. Adaptive gravitational softening in gadget. *Monthly Notices of the Royal Astronomical Society*, 417:2846–2859, 2011.

[69] Neal Dalal, Ue-Li Pen, and Uros Seljak. Large-scale BAO signatures of the smallest galaxies. *Journal of Cosmology and Astroparticle Physics*, 2010(11):007–007, nov 2010.

[70] Jaiyul Yoo and Uroš Seljak. Signatures of first stars in galaxy surveys: Multitracer analysis of the supersonic relative velocity effect and the constraints from the BOSS power spectrum measurements. *Phys. Rev. D*, 88(10):103520, 2013.

[71] Rennan Barkana and Abraham Loeb. Scale-Dependent Bias of Galaxies from Baryonic Acoustic Oscillations. *Mon. Not. Roy. Astron. Soc.*, 415:3113, 2011.

[72] Florian Beutler, Uros Seljak, and Zvonimir Vlah. Constraining the relative velocity effect using the Baryon Oscillation Spectroscopic Survey. *Mon. Not. Roy. Astron. Soc.*, 470(3):2723–2735, 2017.

[73] Alexandre Barreira, Giovanni Cabass, Dylan Nelson, and Fabian Schmidt. Baryon-CDM isocurvature galaxy bias with IllustrisTNG. *JCAP*, 02:005, 2020.

[74] S. Weinberg. *Cosmology*. Springer, 2009.

[75] Scott Dodelson. Modern cosmology. Academic Press, 2003.

[76] Sean M. Carroll. *Spacetime and Geometry*. Cambridge University Press, 7 2019.

[77] L. Amendola and S. Tsujikawa. *Dark Energy: Theory and Observations*. Cambridge University Press, 2010.

[78] Houjun Mo, Frank van den Bosch, and Simon White. *Galaxy Formation and Evolution*. Cambridge University Press, 2012.

[79] T. Padmanabhan. *Cosmology and Astrophysics through Problems*. Cambridge University Press, 1996.

[80] Wolfgang Rindler. *Relativity. Special, general, and cosmological*. Springer, 2007.

[81] G Gamow. The Origin of Elements and the Separation of Galaxies. *PHYSICAL REVIEW JOURNALS*, 1948.

[82] COBE collaboration. Measurement of the Cosmic Microwave Background Spectrum by the COBE FIRAS Instrument. *Astrophysical Journal*, 1994.

[83] T. M. C. Abbott et al. Dark Energy Survey Year 3 Results: Cosmological Constraints from Galaxy Clustering and Weak Lensing. 5 2021.

[84] G. Hinshaw et al. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, and Basic Results. *Astrophys. J. Suppl.*, 180:225–245, 2009.

[85] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 594:A13, 2016.

[86] C. J. Copi, D. N. Schramm, and M. S. Turner. Big-bang nucleosynthesis and the baryon density of the universe. *Science*, 267:192–199, 1995.

[87] R. K. Sachs and A. M. Wolf. Perturbations of a Cosmological Model and Angular Variations of the Microwave Background. *Astrophys. J.*, 147:73, 1967.

[88] Diego Blas, Mathias Garny, and Thomas Konstandin. Cosmological perturbation theory at three-loop order. *Journal of Cosmology and Astroparticle Physics*, 2014(01):010–010, jan 2014.

[89] Martín Crocce and Román Scoccimarro. Renormalized cosmological perturbation theory. *Phys. Rev. D*, 73:063519, Mar 2006.

[90] A. Lewis and S. Bridle. Cosmological parameters from CMB and other data: A Monte Carlo approach. *PHYSICAL REVIEW D*, 66:103511, 2002.

[91] R. J. Bouwens, G. D. Illingworth, P. A. Oesch, J. Caruana, B. Holwerda, R. Smit, and S. Wilkins. REIONIZATION AFTERPLANCK: THE DERIVED GROWTH OF THE COSMIC IONIZING EMISSIVITY NOW MATCHES THE GROWTH OF THE GALAXY UV LUMINOSITY DENSITY. *The Astrophysical Journal*, 811(2):140, sep 2015.

[92] Thejs Brinckmann and Julien Lesgourgues. MontePython 3: boosted MCMC sampler and other features. *Phys. Dark Univ.*, 24:100260, 2019.

[93] Diego Blas, Julien Lesgourgues, and Thomas Tram. The cosmic linear anisotropy solving system (class). part ii: Approximation schemes. *Journal of Cosmology and Astroparticle Physics*, 2011(07):034–034, Jul 2011.

[94] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. emcee: The mcmc hammer. *Publications of the Astronomical Society of the Pacific*, 125(925):306–312, Mar 2013.

[95] Stephen D. Landy and Alexander S. Szalay. Bias and Variance of Angular Correlation Functions. *Astrophys. J.*, 412:64, July 1993.

[96] Nick Kaiser. Clustering in real space and in redshift space. *Monthly Notices of the Royal Astronomical Society*, 227(1):1–21, 07 1987.

[97] Vincent Desjacques, Donghui Jeong, and Fabian Schmidt. Large-Scale Galaxy Bias. *Phys. Rept.*, 733:1–193, 2018.

[98] C. Alcock and B. Paczynski. An evolution free test for non-zero cosmological constant. *Nature*, 281:358,359, 10 1979.

[99] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro. Large scale structure of the universe and cosmological perturbation theory. *Phys. Rept.*, 367:1–248, 2002.

[100] Ya. B. Zeldovich. Gravitational Instability: An Approximate Theory for Large Density Perturbations. *A&A*, 500:13, 1970.

[101] Sverre J. Aarseth. *Gravitational N-Body Simulations*. Cambridge University Press, 2003.

[102] S. J. Aarseth and F. Hoyle. Dynamical Evolution of Clusters of Galaxies, I. *Monthly Notices of the Royal Astronomical Society*, 126(3):223–255, 06 1963.

[103] Adrian L. Melott, Randall J. Splinter, and Sergei F. Shandarin. Demonstrating discreteness and collision error in cosmological n body simulations of dark matter gravitational clustering. *Astrophys. J. Lett.*, 479:L79, 1997.

[104] Joshua Barnes and P. Hut. A hierarchical o(n log n) force-calculation algorithm. *Nature*, 324:446–449, 1986.

[105] R.W. Hockney and J.W. Eastwood. *Computer Simulation Using Particles*. CRC Press, 1988.

[106] W. H. Press and P. Schechter. Formation of galaxies and clusters of galaxies by self-similar gravitational condensation. *Astrophysical Journal*, 187:425–438, 1974.

[107] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. *Astrophysical Journal*, 292:371–394, 1985.

[108] W. H. Press and M. Davis. How to identify and weigh virialized clusters of galaxies in a complete redshift catalog. *Astrophysical Journal*, 259:449–473, 1982.

[109] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. *Astrophysical Journal*, 292:371–394, 1985.

[110] Cedric Lacey and Shanu Cole. Merger rates in hierarchical models of galaxy formation – II. Comparison with N-body simulations. *Monthly Notices of the Royal Astronomical Society*, 271(3):676–692, 12 1994.

[111] Greg L. Bryan and Michael L. Norman. Statistical properties of x-ray clusters: Analytic and numerical comparisons. *The Astrophysical Journal*, 495(1):80–99, mar 1998.

[112] James M. Gelb and Edmund Bertschinger. Cold dark matter. 1: The Formation of dark halos. *Astrophys. J.*, 436:467, 1994.

[113] D. W. Pfitzner and J. K. Salmon. Parallel halo finding in n-body cosmology simulations. *Astrophysical Journal*, 292:371–394, 1996.

[114] Anatoly Klypin and Jon Holtzman. Particle mesh code for cosmological simulations. 12 1997.

[115] Daniel J. Eisenstein and Piet Hut. HOP: A New group finding algorithm for N body simulations. *Astrophys. J.*, 498:137–142, 1998.

[116] Joachim Gerhard Stadel. Cosmological N-body simulations and their analysis. *Thesis (PhD). UNIVERSITY OF WASHINGTON*, 498:3657, 2001.

[117] James S. Bullock, Tsafrir S. Kolatt, Yair Sigad, Rachel S. Somerville, Andrey V. Kravtsov, Anatoly A. Klypin, Joel R. Primack, and Avishai Dekel. Profiles of dark haloes. Evolution, scatter, and environment. *Mon. Not. Roy. Astron. Soc.*, 321:559–575, 2001.

[118] Volker Springel, Simon D. M. White, Giuseppe Tormen, and Guinevere Kauffmann. Populating a cluster of galaxies. 1. Results at z = 0. *Mon. Not. Roy. Astron. Soc.*, 328:726, 2001.

[119] Stuart P.D. Gill, Alexander Knebe, and Brad K. Gibson. The Evolution substructure 1: A New identification method. *Mon.Not.Roy.Astron.Soc.*, 351:399, 2004.

[120] Dominique Aubert, Christophe Pichon, and Stephane Colombi. The Origin and implications of dark matter anisotropic cosmic infall on ~ L(*) halos. *Mon. Not. Roy. Astron. Soc.*, 352:376, 2004.

[121] Mark C. Neyrinck, Nickolay Y. Gnedin, and Andrew J. S. Hamilton. voboz: an almost-parameter-free halo-finding algorithm. *Monthly Notices of the Royal Astronomical Society*, 356(4):1222–1232, 02 2005.

[122] Juhan Kim and Changbom Park. A new halo-finding method forN-body simulations. *The Astrophysical Journal*, 639(2):600–616, mar 2006.

[123] Jurg Diemand, Michael Kuhlen, and Piero Madau. Early supersymmetric cold dark matter substructure. *Astrophys. J.*, 649:1–13, 2006.

[124] Laurie D. Shaw, Jochen Weller, Jeremiah P Ostriker, and Paul Bode. The bound mass of substructures in dark matter halos. *The Astrophysical Journal*, 659(2):1082–1095, apr 2007.

[125] Jeffrey P. Gardner, Andrew Connolly, and Cameron McBride. Enabling Rapid Development of Parallel Tree Search Applications. 9 2007.

[126] J. Gardner, A. Connolly, and C. McBride. Enabling knowledge discovery in a virtual universe. 2007.

[127] M. Maciejewski, S. Colombi, V. Springel, C. Alard, and F. R. Bouchet. Phase–space structures – II. Hierarchical Structure Finder. *Monthly Notices of the Royal Astronomical Society*, 396(3):1329–1348, 06 2009.

[128] S. Habib, A. Pope, Z. Luzic, D. Daniel, P. Fasel, N. Desai, K. Heitmann, C. Hsu, L. Ankeny, G. Mark, S. Bhattacharya, and J. Ahrens. Lanl finder. 2009.

[129] Steffen R. Knollmann and Alexander Knebe. Ahf: Amiga's Halo Finder. *Astrophys.J.Suppl.*, 182:608–624, 2009.

[130] Stephen Skory, Matthew J. Turk, Michael L. Norman, and Alison L. Coil. PARALLEL HOP: A SCALABLE HALO FINDER FOR MASSIVE COSMOLOGICAL DATA SETS. *The Astrophysical Journal Supplement Series*, 191(1):43–57, oct 2010.

[131] S. Planelles and V Quilis. Lanl finder. 2010.

[132] P. M. Sutter and P. M. Ricker. EXAMINING SUBGRID MODELS OF SUPERMASSIVE BLACK HOLES IN COSMOLOGICAL SIMULATION. *The Astrophysical Journal*, 723(2):1308–1318, oct 2010.

[133] Y. Rasera, J. Alimi, J. Courtin, F. Roy, P. Corasaniti, A. Fuzfa, and V. Boucher. 1241:1134–1139, 2010.

[134] Bridget L. Falck, Mark C. Neyrinck, Miguel A. Aragon-Calvo, Guilhem Lavaux, and Alexander S. Szalay. STRAIGHTENING THE DENSITY-DISPLACEMENT RELATION WITH a LOGARITHMIC TRANSFORM. *The Astrophysical Journal*, 745(1):17, dec 2011.

[135] Peter S. Behroozi, Risa H. Wechsler, and Hao-Yi Wu. THE ROCKSTAR PHASE-SPACE TEMPORAL HALO FINDER AND THE VELOCITY OFFSETS OF CLUSTER CORES. *The Astrophysical Journal*, 762(2):109, dec 2012.

[136] Marius Cautun, Rien van de Weygaert, Bernard J. T. Jones, and Carlos S. Frenk. Evolution of the cosmic web. *MNRAS*, 441(4):2923–2973, July 2014.

[137] B. Falck and M. C. Neyrinck. The persistent percolation of single-stream voids. *MNRAS*, 450(3):3239–3253, July 2015.

[138] M. A. Hausman, D. W. Olson, and B. D. Roth. The evolution of voids in the expanding universe. *The Astrophysical Journal*, 270:351–359, 1983.

[139] S. A. Gregory, L. A. Thompson, and W. G. Tifft. Observations of the Large Scale Distribution of Galaxies. *BAAS*, 10:622, 1978.

[140] R. P. Kirshner, Jr. Oemler, A., P. L. Schechter, and S. A. Shectman. A million cubic megaparsec void in Bootes ? *The Astrophysical Journal*, 248:L57–L60, 1981.

[141] Jorg M. Colberg, Ravi K. Sheth, Antonaldo Diaferio, Liang Gao, and Naoki Yoshida. Voids in a ΛCDM universe. *Mon. Not. Roy. Astron. Soc.*, 360:216–226, 2005.

[142] Nico Hamaus, P. M. Sutter, and Benjamin D. Wandelt. Universal density profile for cosmic voids. *Phys. Rev. Lett.*, 112:251302, Jun 2014.

[143] E. Ricciardelli, V. Quilis, and J. Varela. On the universality of void density profiles. *Monthly Notices of the Royal Astronomical Society*, 440(1):601–609, 03 2014.

[144] E. Ricciardelli, V. Quilis, and S. Planelles. The structure of cosmic voids in a ΛCDM Universe. *Monthly Notices of the Royal Astronomical Society*, 434(2):1192–1204, 07 2013.

[145] S. Nadathur, S. Hotchkiss, J. M. Diego, I. T. Iliev, S. Gottlöber, W. A. Watson, and G. Yepes. Universal void density profiles from simulation and SDSS. *IAU Symp.*, 308:542–545, 2014.

[146] C. Alcock and B. Paczynski. An evolution free test for non-zero cosmological constant. *Nature*, 281:358, 1979.

[147] Guilhem Lavaux and Benjamin D. Wandelt. PRECISION COSMOGRAPHY WITH STACKED VOIDS. *The Astrophysical Journal*, 754(2):109, jul 2012.

[148] P. M. Sutter, Guilhem Lavaux, Benjamin D. Wandelt, and David H. Weinberg. A first application of the alcock-paczynski test to stacked cosmic voids. *The Astrophysical Journal*, 761(2):187, Dec 2012.

[149] P. M. Sutter, Alice Pisani, Benjamin D. Wandelt, and David H. Weinberg. A measurement of the Alcock–Paczyński effect using cosmic voids in the SDSS. *Monthly Notices of the Royal Astronomical Society*, 443(4):2983–2990, 08 2014.

[150] Nico Hamaus, P. M. Sutter, Guilhem Lavaux, and Benjamin D. Wandelt. Probing cosmology and gravity with redshift-space distortions around voids. *JCAP*, 11:036, 2015.

[151] Nico Hamaus, Alice Pisani, Paul M. Sutter, Guilhem Lavaux, Stéphanie Escoffier, Benjamin D. Wandelt, and Jochen Weller. Constraints on Cosmology and Gravity from the Dynamics of Voids. *Phys. Rev. Lett.*, 117(9):091302, 2016.

[152] Qingqing Mao et al. A Cosmic Void Catalog of SDSS DR12 BOSS Galaxies. *Astrophys. J.*, 835(2):161, 2017.

[153] N. Hamaus, M. Aubert, A. Pisani, S. Contarini, G. Verza, M. C. Cousinou, S. Escoffier, A. Hawken, G. Lavaux, G. Pollina, B. D. Wandelt, J. Weller, M. Bonici, C. Carbone, L. Guzzo, A. Kovacs, F. Marulli, E. Massara, L. Moscardini, P. Ntelis, W. J. Percival, S. Radinović, M. Sahlén, Z. Sakr, A. G. Sánchez, H. A. Winther, N. Auricchio, S. Awan, R. Bender, C. Bodendorf, D. Bonino, E. Branchini, M. Brescia, J. Brinchmann, V. Capobianco, J. Carretero, F. J. Castander, M. Castellano, S. Cavuoti, A. Cimatti, R. Cledassou, G. Congedo, L. Conversi, Y. Copin, L. Corcione, M. Cropper, A. Da Silva, H. Degaudenzi, M. Douspis, F. Dubath, C. A. J. Duncan, X. Dupac, S. Dusini, A. Ealet, S. Ferriol, P. Fosalba, M. Frailis, E. Franceschi, P. Franzetti, M. Fumana, B. Garilli, B. Gillis, C. Giocoli, A. Grazian, F. Grupp, S. V. H. Haugan, W. Holmes, F. Hormuth, K. Jahnke, S. Kermiche, A. Kiessling, M. Kilbinger, T. Kitching, M. Kümmel, M. Kunz, H. Kurki-Suonio, S. Ligori, P. B. Lilje, I. Lloro, E. Maiorano, O. Marggraf, K. Markovic, R. Massey, S. Maurogordato, M. Melchior, M. Meneghetti, G. Meylan, M. Moresco, E. Munari, S. M. Niemi, C. Padilla, S. Paltani, F. Pasian, K. Pedersen, V. Pettorino, S. Pires, M. Poncet, L. Popa, L. Pozzetti, R. Rebolo, J. Rhodes, H. Rix, M. Roncarelli, E. Rossetti, R. Saglia, P. Schneider,

A. Secroun, G. Seidel, S. Serrano, C. Sirignano, G. Sirri, J. L. Starck, P. Tallada-Crespí, D. Tavagnacco, A. N. Taylor, I. Tereno, R. Toledo-Moreo, F. Torradeflot, E. A. Valentijn, L. Valenziano, Y. Wang, N. Welikala, G. Zamorani, J. Zoubian, S. Andreon, M. Baldi, S. Camera, S. Mei, C. Neissner, and E. Romelli. Euclid: Forecasts from redshift-space distortions and the alcock-paczynski test with cosmic voids, 2021.

[154] Alice Pisani, P. M. Sutter, Nico Hamaus, Esfandiar Alizadeh, Rahul Biswas, Benjamin D. Wandelt, and Christopher M. Hirata. Counting voids to probe dark energy. *Phys. Rev. D*, 92:083531, Oct 2015.

[155] Andrzej Odrzywołek. Holes in the static einstein universe and a model of the cosmological voids. *Phys. Rev. D*, 80:103515, Nov 2009.

[156] Guido D'Amico, Marcello Musso, Jorge Noreña, and Aseem Paranjape. Excursion sets and non-gaussian void statistics. *Phys. Rev. D*, 83:023521, Jan 2011.

[157] E. G. Patrick Bos, Rien van de Weygaert, Klaus Dolag, and Valeria Pettorino. The darkness that shaped the void: dark energy and cosmic voids. *Mon. Not. Roy. Astron. Soc.*, 426:440, 2012.

[158] G. W. Gibbons, M. C. Werner, N. Yoshida, and S. Chon. On de Sitter geometry in cosmic void statistics. *Monthly Notices of the Royal Astronomical Society*, 438(2):1603–1610, 12 2013.

[159] Guilhem Lavaux and Benjamin D. Wandelt. Precision cosmology with voids: definition, methods, dynamics. *Monthly Notices of the Royal Astronomical Society*, 403(3):1392–1408, Apr 2010.

[160] Joseph Clampitt, Yan-Chuan Cai, and Baojiu Li. Voids in modified gravity: excursion set predictions. *Monthly Notices of the Royal Astronomical Society*, 431(1):749–766, 03 2013.

[161] Baojiu Li. Voids in coupled scalar field cosmology. *Monthly Notices of the Royal Astronomical Society*, 411(4):2615–2627, 03 2011.

[162] Yan-Chuan Cai, Nelson Padilla, and Baojiu Li. Testing Gravity using Cosmic Voids. *Mon. Not. Roy. Astron. Soc.*, 451(1):1036–1055, 2015.

[163] Alexandre Barreira, Marius Cautun, Baojiu Li, Carlton Baugh, and Silvia Pascoli. Weak lensing by voids in modified lensing potentials. *JCAP*, 08:028, 2015.

[164] Paul Zivick, P. M. Sutter, Benjamin D. Wandelt, Baojiu Li, and Tsz Yan Lam. Using cosmic voids to distinguish f(R) gravity in future galaxy surveys. *Monthly Notices of the Royal Astronomical Society*, 451(4):4215–4222, 06 2015.

[165] Giorgia Pollina, Marco Baldi, Federico Marulli, and Lauro Moscardini. Cosmic voids in coupled dark energy cosmologies: the impact of halo bias. *Monthly Notices of the Royal Astronomical Society*, 455(3):3075–3085, 11 2015.

[166] Marco Baldi and Francisco Villaescusa-Navarro. Cosmic degeneracies – II. Structure formation in joint simulations of warm dark matter and $f(R)$ gravity. *Mon. Not. Roy. Astron. Soc.*, 473(3):3226–3240, 2018.

[167] Elena Massara, Francisco Villaescusa-Navarro, Matteo Viel, and P. M. Sutter. Voids in massive neutrino cosmologies. *JCAP*, 2015(11):018, November 2015.

[168] Christina D Kreisch, Alice Pisani, Carmelita Carbone, Jia Liu, Adam J Hawken, Elena Massara, David N Spergel, and Benjamin D Wandelt. Massive neutrinos leave fingerprints on cosmic voids. *Monthly Notices of the Royal Astronomical Society*, 488(3):4413–4426, 07 2019.

[169] Sofia Contarini, Federico Marulli, Lauro Moscardini, Alfonso Veropalumbo, Carlo Giocoli, and Marco Baldi. Cosmic voids in modified gravity models with massive neutrinos. *Monthly Notices of the Royal Astronomical Society*, 504(4):5021–5038, Apr 2021.

[170] C. Baccigalupi, L. Amendola, and F. Occhionero. Imprints of primordial voids on the cosmic microwave background. *Monthly Notices of the Royal Astronomical Society*, 288(2):387–396, 06 1997.

[171] Carlo Baccigalupi. Cosmic microwave background: Polarization and temperature anisotropies from symmetric structures. *Phys. Rev. D*, 59(12):123004, June 1999.

[172] Benjamin R. Granett, Mark C. Neyrinck, and István Szapudi. An Imprint of Superstructures on the Microwave Background due to the Integrated Sachs-Wolfe Effect. *Astrophys. J. Lett.*, 683(2):L99, August 2008.

[173] Yan-Chuan Cai, Mark C. Neyrinck, István Szapudi, Shaun Cole, and Carlos S. Frenk. A possible cold imprint of voids on the microwave background radiation. *The Astrophysical Journal*, 786(2):110, Apr 2014.

[174] S. Hotchkiss, S. Nadathur, S. Gottlöber, I. T. Iliev, A. Knebe, W. A. Watson, and G. Yepes. The Jubilee ISW Project - II. Observed and simulated imprints of voids and superclusters on the cosmic microwave background. *MNRAS*, 446(2):1321–1334, January 2015.

[175] P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, N. Bartolo, and et al. Planck2015 results. *Astronomy & Astrophysics*, 594:A21, Sep 2016.

[176] B. R. Granett, A. Kovács, and A. J. Hawken. The integrated Sachs-Wolfe signal from BOSS superstructures. *MNRAS*, 454(3):2804–2814, December 2015.

[177] Seshadri Nadathur and Robert Crittenden. A Detection of the Integrated Sachs-Wolfe Imprint of Cosmic Superstructures Using a Matched-filter Approach. *Astrophys. J. Lett.*, 830(1):L19, October 2016.

[178] A. Kovács, C. Sánchez, J. García-Bellido, S. Nadathur, R. Crittenden, D. Gruen, D. Huterer, D. Bacon, J. Clampitt, J. DeRose, S. Dodelson, E. Gaztañaga, B. Jain, D. Kirk, O. Lahav, R. Miquel, K. Naidoo, J. A. Peacock, B. Soergel, L. Whiteway, F. B. Abdalla, S. Allam, J. Annis, A. Benoit-Lévy, E. Bertin, D. Brooks, E. Buckley-Geer, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, C. E. Cunha, C. B. D'Andrea, L. N. da Costa, D. L. DePoy, S. Desai, T. F. Eifler, D. A. Finley, B. Flaugher, P. Fosalba, J. Frieman, T. Giannantonio, D. A. Goldstein, R. A. Gruendl, G. Gutierrez, D. J. James, K. Kuehn, N. Kuropatkin, J. L. Marshall, P. Melchior, F. Menanteau, B. Nord, R. Ogando, A. A. Plazas, A. K. Romer, E. Sanchez, V. Scarpine, I. Sevilla-Noarbe, F. Sobreira, E. Suchyta, M. Swanson, G. Tarle, D. Thomas, A. R. Walker, and DES Collaboration. Imprint of DES superstructures on the cosmic microwave background. *MNRAS*, 465(4):4166–4179, March 2017.

[179] A. Kovács, C. Sánchez, J. García-Bellido, J. Elvin-Poole, N. Hamaus, V. Miranda, S. Nadathur, T. Abbott, F. B. Abdalla, J. Annis, S. Avila, E. Bertin, D. Brooks, D. L. Burke, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, R. Cawthon, M. Crocce, C. Cunha, L. N. da Costa, C. Davis, J. De Vicente, D. DePoy, S. Desai, H. T. Diehl, P. Doel, E. Fernandez, B. Flaugher, P. Fosalba, J. Frieman, E. Gaztañaga, D. Gerdes, R. Gruendl, G. Gutierrez, W. Hartley, D. L. Hollowood, K. Honscheid, B. Hoyle, D. J. James, E. Krause, K. Kuehn, N. Kuropatkin, O. Lahav, M. Lima, M. Maia, M. March, J. Marshall, P. Melchior, F. Menanteau, C. J. Miller, R. Miquel, J. Mohr, A. A. Plazas, K. Romer, E. Rykoff, E. Sanchez, V. Scarpine, R. Schindler, M. Schubnell, I. Sevilla-Noarbe, M. Smith, R. C. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, M. Swanson, G. Tarle, D. Thomas, V. Vikram, J. Weller, and DES Collaboration. More out of less: an excess integrated Sachs-Wolfe signal from supervoids mapped out by the Dark Energy Survey. *MNRAS*, 484(4):5267–5277, April 2019.

[180] Qianjun Hang, Shadab Alam, Yan-Chuan Cai, and John Peacock. Stacked CMB lensing and ISW signals around superstructures in the DESI Legacy Survey. *arXiv e-prints*, page arXiv:2105.11936, May 2021.

[181] Yan-Chuan Cai, Mark Neyrinck, Qingqing Mao, John A. Peacock, Istvan Szapudi, and Andreas A. Berlind. The lensing and temperature imprints of voids on the cosmic microwave background. *MNRAS*, 466(3):3364–3375, April 2017.

[182] P. Vielzeuf, A. Kovács, U. Demirbozan, P. Fosalba, E. Baxter, N. Hamaus, D. Huterer, R. Miquel, S. Nadathur, G. Pollina, C. Sánchez, L. Whiteway, T. M. C. Abbott, S. Allam, J. Annis, S. Avila, D. Brooks, D. L. Burke, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, R. Cawthon, M. Costanzi, L. N. da Costa, J. De Vicente, S. Desai, H. T. Diehl, P. Doel, T. F. Eifler, S. Everett, B. Flaugher, J. Frieman, J. García-Bellido, E. Gaztanaga, D. W. Gerdes, D. Gruen, R. A. Gruendl, J. Gschwend, G. Gutierrez, W. G. Hartley, D. L. Hollowood, K. Honscheid, D. J. James, K. Kuehn, N. Kuropatkin, O. Lahav, M. Lima, M. A. G. Maia, M. March, J. L. Marshall, P. Melchior, F. Menanteau, A. Palmese, F. Paz-Chinchón, A. A. Plazas, E. Sanchez, V. Scarpine, S. Serrano, I. Sevilla-Noarbe, M. Smith, E. Suchyta, G. Tarle, D. Thomas, J. Weller, J. Zuntz, and DES Collaboration. Dark Energy Survey Year 1 results: the lensing imprint of cosmic voids on the cosmic microwave background. *MNRAS*, 500(1):464–480, January 2021.

[183] Srinivasan Raghunathan, Seshadri Nadathur, Blake D. Sherwin, and Nathan Whitehorn. The Gravitational Lensing Signatures of BOSS Voids in the Cosmic Microwave Background. *Astrophys. J. Lett.*, 890(2):168, February 2020.

[184] M. J. Rees, D. W. Sciama, and S. H. Stobbs. Metastable Helium in Interstellar and Intergalactic Space. *Astrophysical Letters*, 2:243, 1968.

[185] F. Finelli, J. García-Bellido, A. Kovács, F. Paci, and I. Szapudi. A Supervoid Explanation of the Cosmic Microwave Background Cold Spot. *Proceedings of the International Astronomical Union*, 10(S306):153–155, 2014.

[186] K. Kovač, S. J. Lilly, C. Knobel, T. J. Bschorr, Y. Peng, C. M. Carollo, T. Contini, O. Kneib, J.-P. an d Le Févre, V. Mainieri, A. Renzini, M. Scodeggio, G. Zamorani, S. Bardelli, M. Bolzonella, A. Bongiorno, K. Cãputi, O. Cucciati, S. de la Torre, L. de Ravel, P. Franzetti, B. Garilli, A. Iovino, P. Kampczyk, F. Lamareillẽ, J.-F. Le Borgne, V. Le Brun, C. Maier, M. Mignoli, P. Oesch, R. Pello, E. Perez Montero, J. Presotto, V. a nd Silverman, M. Tanaka, L. Tasca, L. Tresse, D. Vergani, E. Zucca, H. Aussel, A. M. Koekemoer, Ẽ. Le Floc'h, M. Moresco, and L. Pozzetti. zCOSMOS 20k: satellite galaxies are the main drivers of environmental effects in the galaxy population at least to $z \sim 0.7$. *Monthly Notices of the Royal Astronomical Society*, 438(1):717–738, 12 2013.

[187] Seshadri Nadathur, Mikko Lavinto, Shaun Hotchkiss, and Syksy Räsänen. Can a supervoid explain the cold spot? *Phys. Rev. D*, 90:103510, Nov 2014.

[188] Matteo Viel, Jörg M. Colberg, and T.-S. Kim. On the importance of high-redshift intergalactic voids. *Monthly Notices of the Royal Astronomical Society*, 386(3):1285–1293, 04 2008.

[189] Erwin Platen, Rien van de Weygaert, and Bernard J. T. Jones. A Cosmic Watershed: The WVF Void Detection Technique. *Mon. Not. Roy. Astron. Soc.*, 380:551–570, 2007.

[190] W. E. Schaap. PhD Thesis. *PhD Thesis - University of Groningen*, 2007.

[191] N Jeffrey, M Gatti, C Chang, L Whiteway, U Demirbozan, A Kovacs, G Pollina, D Bacon, N Hamaus, T Kacprzak, and et al. Dark energy survey year 3 results: Curved-sky weak lensing mass map reconstruction. *Monthly Notices of the Royal Astronomical Society*, 505(3):4626–4645, May 2021.

[192] Seshadri Nadathur et al. The completed SDSS-IV extended baryon oscillation spectroscopic survey: geometry and growth from the anisotropic void–galaxy correlation function in the luminous red galaxy sample. *Mon. Not. Roy. Astron. Soc.*, 499(3):4140–4157, 2020.

[193] R.R. Caldwell, Rahul Dave, and Paul J. Steinhardt. Cosmological imprint of an energy component with general equation of state. *Phys. Rev. Lett.*, 80:1582–1585, 1998.

[194] Shinji Tsujikawa. Quintessence: A Review. *Class. Quant. Grav.*, 30:214003, 2013.

[195] Valeria Pettorino, Luca Amendola, and Christof Wetterich. How early is early dark energy? *Phys. Rev. D*, 87:083009, 2013.

[196] Wayne Hu. Structure formation with generalized dark matter. *The Astrophysical Journal*, 506(2):485–494, oct 1998.

[197] R. R. Caldwell and Paul J. Steinhardt. The Imprint of gravitational waves in models dominated by a dynamical cosmic scalar field. *Phys. Rev. D*, 57:6057–6064, 1998.

[198] James M. Bardeen. Gauge-invariant cosmological perturbations. *Phys. Rev. D*, 22:1882–1905, Oct 1980.

[199] C.-P. Ma and E. Bertschinger. Cosmological Perturbation Theory in the Synchronous and Conformal Newtonian Gauges. *The Astronomical Journal*, 455(1):7, 1995.

[200] Hideo Kodama and Misao Sasaki. Cosmological Perturbation Theory. *Progress of Theoretical Physics Supplement*, 78:1–166, 01 1984.

[201] Valeria Pettorino and Carlo Baccigalupi. Coupled and Extended Quintessence: theoretical differences and structure formation. *Phys. Rev.*, D77:103003, 2008.

[202] E. Calabrese, R. de Putter, D. Huterer, E. V. Linder, and A. Melchiorri. Future CMB constraints on early, cold, or stressed dark energy. *PHYSICAL REVIEW D*, 83(2):3011, 2011.

[203] M. Archidiacono, L. Lopez-Honorez, and O. Mena. Current constraints on early and stressed dark energy models and future 21 cm perspectives. *PHYSICAL REVIEW D*, 90(12):3016, 2014.

[204] N. Aghanim et al. Planck 2018 results. V. CMB power spectra and likelihoods. 2019.

[205] P. A. R. Ade and others [Planck Collaboration]. Planck 2013 results. XV. CMB power spectra and likelihood. *Astronomy and Astrophysics*, 571:A15, 2014.

[206] N. Aghanim and others [Planck Collaboration]. Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters. *Astronomy and Astrophysics*, 594(A11):99, 2016.

[207] N. Aghanim et al. Planck intermediate results. XLVI. Reduction of large-scale systematic effects in HFI polarization maps and estimation of the reionization optical depth. *Astron. Astrophys.*, 596:A107, 2016.

[208] N. Aghanim et al. Planck 2018 results. III. High Frequency Instrument data processing and frequency maps. 2018.

[209] P. A. R. Ade et al. Planck 2015 results. XV. Gravitational lensing. *Astron. Astrophys.*, 594:A15, 2016.

[210] Florian Beutler, Chris Blake, Matthew Colless, D. Heath Jones, Lister Staveley-Smith, Lachlan Campbell, Quentin Parker, Will Saunders, and Fred Watson. The 6dF Galaxy Survey: baryon acoustic oscillations and the local Hubble constant. *Monthly Notices of the Royal Astronomical Society*, 416(4):3017–3032, 09 2011.

[211] A. J. Ross, L. Samushia, C. Howlett, W. J. Percival, A. Burden, and M. Manera. The clustering of the SDSS DR7 main Galaxy sample I: A 4 per cent distance measure at $z = 0.15$. *Monthly Notices of the Royal Astronomical Society*, 449(1):835–847, 2015.

[212] S. Alam et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. *Monthly Notices of the Royal Astronomical Society*, 470(3):2617–2652, 2017.

[213] J. E. Bautista et al. Measurement of baryon acoustic oscillation correlations at $z = 2.3$ with SDSS DR12 Lyα-Forests. *Astronomy and Astrophysics*, 603(A12):23, 2017.

[214] H. M. Bourboux et al. Baryon acoustic oscillations from the complete SDSS-III Lyα-quasar cross-correlation function at $z = 2.4$. *Astronomy and Astrophysics*, 608(A130):22, 2017.

[215] V. de Sainte Agathe et al. Baryon acoustic oscillations at z = 2.34 from the correlations of Lyα absorption in eBOSS DR14. 2019.

[216] M. Blomqvist et al. Baryon acoustic oscillations from the cross-correlation of Lyα absorption and quasars in eBOSS DR14. 2019.

[217] M. Betoule et al. Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples. *Astronomy and Astrophysics*, 568:A22, 2014.

[218] G. Risaliti and E. Lusso. A Hubble diagram for quasars. *The Astrophysical Journal*, 815(1):33, 2015.

[219] G. Risaliti and E. Lusso. Cosmological constraints from the Hubble diagram of quasars at high redshifts. *Nature Astronomy*, 3(3):272–277, 2019.

[220] Daniel J. Eisenstein. An Analytic expression for the growth function in a flat universe with a cosmological constant. *Submitted to: Astrophys. J.*, 1997.

[221] Daniel J. Eisenstein, Wayne Hu, Joseph Silk, and Alexander S. Szalay. Can baryonic features produce the observed 100 h^{-1} mpc clustering? *The Astrophysical Journal*, 494(1):L1–L4, feb 1998.

[222] Daniel J. Eisenstein, Wayne Hu, and Max Tegmark. Cosmic complementarity: h_0 and ω_m from combining cosmic microwave background experiments and redshift surveys. *The Astrophysical Journal*, 504(2):L57–L60, sep 1998.

[223] Daniel Eisenstein. Large scale structure and future surveys. In *Conference on Next Generation Wide-Field Multi-Object Spectroscopy Tuscon, Arizona, October 11-12, 2001*, 2003. [Submitted to: ASP Conf. Ser.(2003)].

[224] Daniel Eisenstein and Martin White. Theoretical uncertainty in baryon oscillations. *Phys. Rev. D*, 70:103523, 2004.

[225] Daniel J. Eisenstein et al. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *Astrophys. J.*, 633:560–574, 2005.

[226] Daniel J. Eisenstein, Hee-jong Seo, Edwin Sirko, and David Spergel. Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak. *Astrophys. J.*, 664:675–679, 2007.

[227] Daniel J. Eisenstein, Hee-Jong Seo, and Martin White. On the robustness of the acoustic scale in the low-redshift clustering of matter. *The Astrophysical Journal*, 664(2):660–674, aug 2007.

[228] Bruce A. Bassett and Renee Hlozek. Baryon Acoustic Oscillations. 2009.

[229] P. A. R. Ade and others [Planck Collaboration]. Planck 2015 results. XIV. Dark energy and modified gravity. *Astronomy and Astrophysics*, 594:A14, 2016.

[230] J. Xia and M. Viel. Early Dark Energy at High Redshifts: Status and Perspectives. *Journal of Cosmology and Astroparticle Physics*, 2009, 2009.

[231] D.M. Scolnic et al. The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample. *Astrophys. J.*, 859(2):101, 2018.

[232] A. G. Riess et al. A 3% Solution: Determination of the Hubble Constant with the Hubble Space Telescope and Wide Field Camera 3. *The Astrophysical Journal*, 730(2):119, 2011.

[233] G. Efstathiou. H0 Revisited. *Monthly Notices of the Royal Astronomical Society*, 440(2):1138–1152, 2014.

[234] E. M. L. Humphreys, Mark J. Reid, Jim M. Moran, Lincoln J. Greenhill, and Alice L. Argon. Toward a New Geometric Distance to the Active Galaxy NGC 4258. III. Final Results and the Hubble Constant. *The Astrophysical Journal*, 775:13, 2013.

[235] A. G. Riess, S. Casertano, W. Yuan, L. M. Macri, and D. Scolnic. Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond ΛCDM. *The Astrophysical Journal*, 876(1):85, 2019.

[236] R. Bean and O. Dore. Probing dark energy perturbations: The Dark energy equation of state and speed of sound as measured by WMAP. *PHYSICAL REVIEW D*, 69:083503, 2004.

[237] J. Weller and A. M. Lewis. Large scale cosmic microwave background anisotropies and dark energy. *Monthly Notices of the Royal Astronomical Society*, 346:987–993, 2003.

[238] J. T. A. de Jong and others [KIDS collaboration]. The Kilo-Degree Survey. 2012.

[239] Tao Yang, Aritra Banerjee, and Eoin Ó. Colgáin. On cosmography and flat ΛCDM tensions at high redshift. 11 2019.

[240] Hermano Velten and Syrios Gomes. Is the Hubble diagram of quasars in tension with concordance cosmology? *Phys. Rev. D*, 101(4):043502, 2020.

[241] A. Aghamousa and others [DESI collaboration]. The DESI Experiment Part I: Science,Targeting, and Survey Design. 2016.

[242] Alexandre Barreira, Giovanni Cabass, Dylan Nelson, and Fabian Schmidt. Baryon-CDM isocurvature galaxy bias with IllustrisTNG. *JCAP*, 02:005, 2020.

[243] Oliver Hahn and Tom Abel. Multi-scale initial conditions for cosmological simulations. *Monthly Notices of the Royal Astronomical Society*, 415(3):2101–2121, Jul 2011.

[244] Raul E. Angulo and Andrew Pontzen. Cosmological N-body simulations with suppressed variance. *Mon. Not. Roy. Astron. Soc.*, 462(1):L1–L5, 2016.

[245] Volker Springel. The Cosmological simulation code GADGET-2. *Mon. Not. Roy. Astron. Soc.*, 364:1105–1134, 2005.

[246] Jeremy L. Tinker, Andrey V. Kravtsov, Anatoly Klypin, Kevork Abazajian, Michael S. Warren, Gustavo Yepes, Stefan Gottlober, and Daniel E. Holz. Toward a halo mass function for precision cosmology: The Limits of universality. *Astrophys. J.*, 688:709–728, 2008.

[247] Ping He, Long-Long Feng, and Li-Zhi Fang. Distributions of baryon fraction on large scales in the Universe. *Astrophys. J.*, 623:601–611, 2005.

[248] Stefano Ettori, K. Dolag, S. Borgani, and G. Murante. The baryon fraction in hydrodynamical simulations of galaxy clusters. *Mon. Not. Roy. Astron. Soc.*, 365:1021–1030, 2006.

[249] Andrey V. Kravtsov, Daisuke Nagai, and Alexey A. Vikhlinin. Effects of cooling and star formation on the baryon fractions in clusters. *Astrophys. J.*, 625:588–598, 2005.

[250] Robert A. Crain, Vincent R. Eke, Carlos S. Frenk, Adrian Jenkins, Ian G. McCarthy, Julio F. Navarro, and Frazer R. Pearce. The baryon fraction of Lambda-CDM haloes. *Mon. Not. Roy. Astron. Soc.*, 377:41–49, 2007.

[251] Stefan Gottloeber and Gustavo Yepes. Shape, spin and baryon fraction of clusters in the MareNostrum Universe. *Astrophys. J.*, 664:117–122, 2007.

[252] Francisco Villaescusa-Navarro et al. The CAMELS project: Cosmology and Astrophysics with MachinE Learning Simulations. 10 2020.

[253] Titouan Lazeyras, Christian Wagner, Tobias Baldauf, and Fabian Schmidt. Precision measurement of the local bias of dark matter halos. *JCAP*, 02:018, 2016.

[254] P.F. De Salas, S. Gariazzo, O. Mena, C.A. Ternes, and M. Tórtola. Neutrino Mass Ordering from Oscillations and Beyond: 2018 Status and Future Prospects. *Front. Astron. Space Sci.*, 5:36, 2018.

[255] Chi-Ting Chiang, Marilena LoVerde, and Francisco Villaescusa-Navarro. First detection of scale-dependent linear halo bias in N-body simulations with massive neutrinos. *Phys. Rev. Lett.*, 122(4):041302, 2019.

[256] Julian B. Muñoz and Cora Dvorkin. Efficient Computation of Galaxy Bias with Neutrinos and Other Relics. *Phys. Rev. D*, 98(4):043503, 2018.

[257] Seshadri Nadathur and Shaun Hotchkiss. The nature of voids – I. Watershed void finders and their connection with theoretical models. *Mon. Not. Roy. Astron. Soc.*, 454(2):2228–2241, 2015.

[258] Ravi K. Sheth and Rien van de Weygaert. A hierarchy of voids: much ado about nothing. *MNRAS*, 350(2):517–538, May 2004.

[259] Steven R. Furlanetto and Tsvi Piran. The evidence of absence: galaxy voids in the excursion set formalism. *MNRAS*, 366(2):467–479, February 2006.

[260] Alice Pisani, P. M. Sutter, Nico Hamaus, Esfandiar Alizadeh, Rahul Biswas, Benjamin D. Wandelt, and Christopher M. Hirata. Counting voids to probe dark energy. *Physical Review D*, 92(8):083531, October 2015.

[261] Giovanni Verza, Alice Pisani, Carmelita Carbone, Nico Hamaus, and Luigi Guzzo. The void size function in dynamical dark energy cosmologies. *JCAP*, 2019(12):040, December 2019.

[262] Sofia Contarini, Tommaso Ronconi, Federico Marulli, Lauro Moscardini, Alfonso Veropalumbo, and Marco Baldi. Cosmological exploitation of the size function of cosmic voids identified in the distribution of biased tracers. *MNRAS*, 488(3):3526–3540, September 2019.

[263] P. M. Sutter, Guilhem Lavaux, Nico Hamaus, Benjamin D. Wandelt, David H. Weinberg, and Michael S. Warren. Sparse sampling, galaxy bias, and voids. *MNRAS*, 442(1):462–471, July 2014.

[264] P. J. E. Peebles and M. G. Hauser. Statistical Analysis of Catalogs of Extragalactic Objects. III. The Shane-Wirtanen and Zwicky Catalogs. *Astrophysical Journal*, 28:19, 1974.

[265] M. Davis and P. J. E. Peebles. A survey of galaxy redshifts. V. The two-point position and velocity correlations. *Astrophysical Journal*, 267:465–482, 1983.

[266] P. C. Hewett. The estimation of galaxy angular correlation functions. *Astrophysical Journal*, 201:867–883, 1982.

[267] A. J. S. Hamilton. Toward Better Ways to Measure the Galaxy Correlation Function. *Astrophysical Journal*, 417:19, 1993.

[268] Nick Hand, Yu Feng, Florian Beutler, Yin Li, Chirag Modi, Uroš Seljak, and Zachary Slepian. nbodykit: An open-source, massively parallel toolkit for large-scale structure. *The Astronomical Journal*, 156(4):160, Sep 2018.

[269] Nico Hamaus, Benjamin D. Wandelt, P. M. Sutter, Guilhem Lavaux, and Michael S. Warren. Cosmology with Void-Galaxy Correlations. *Phys. Rev. Lett.*, 112(4):041304, 2014.

[270] Kwan Chuen Chan, Nico Hamaus, and Vincent Desjacques. Large-Scale Clustering of Cosmic Voids. *Phys. Rev. D*, 90(10):103521, 2014.

[271] Erwin Platen, Rien van de Weygaert, and Bernard J. T. Jones. Alignments of Voids in the Cosmic Web. *Mon. Not. Roy. Astron. Soc.*, 387:128, 2008.

[272] Giorgia Pollina, Nico Hamaus, Klaus Dolag, Jochen Weller, Marco Baldi, and Lauro Moscardini. On the linearity of tracer bias around voids. *Mon. Not. Roy. Astron. Soc.*, 469(1):787–799, 2017.

[273] L. Ceccarelli, D. Paz, M. Lares, N. Padilla, and D. García Lambas. Clues on void evolution I: Large scale galaxy distributions around voids. *Mon. Not. Roy. Astron. Soc.*, 434:1435, 2013.

[274] Joseph Clampitt, Bhuvnesh Jain, and Carles Sánchez. Clustering and Bias Measurements of SDSS Voids. *Mon. Not. Roy. Astron. Soc.*, 456(4):4425–4431, 2016.

[275] Fabian Schmidt. Sigma-Eight at the Percent Level: The EFT Likelihood in Real Space. *JCAP*, 04:032, 2021.

[276] Andrea Pezzotta, Martin Crocce, Alexander Eggemeier, Ariel G. Sánchez, and Román Scoccimarro. Testing one-loop galaxy bias: Cosmological constraints from the power spectrum. *Phys. Rev. D*, 104(4):043531, 2021.

[277] Nico Schuster, Nico Hamaus, Alice Pisani, Carmelita Carbone, Christina D. Kreisch, Giorgia Pollina, and Jochen Weller. The bias of cosmic voids in the presence of massive neutrinos. *JCAP*, 2019(12):055, December 2019.

[278] Kwan Chuen Chan, Nico Hamaus, and Matteo Biagetti. Constraint of void bias on primordial non-Gaussianity. *Phys. Rev. D*, 99(12):121304, June 2019.

[279] Kwan Chuen Chan, Yin Li, Matteo Biagetti, and Nico Hamaus. Measurement of Void Bias Using Separate Universe Simulations. *Astrophys. J.*, 889(2):89, February 2020.

[280] Jeremy L. Tinker, Brant E. Robertson, Andrey V. Kravtsov, Anatoly Klypin, Michael S. Warren, Gustavo Yepes, and Stefan Gottlober. The Large Scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests. *Astrophys. J.*, 724:878–886, 2010.

[281] Marco Peloso, Massimo Pietroni, Matteo Viel, and Francisco Villaescusa-Navarro. The effect of massive neutrinos on the BAO peak. *JCAP*, 07:001, 2015.

[282] Francisco Villaescusa-Navarro, David Alonso, and Matteo Viel. Baryonic acoustic oscillations from 21 cm intensity mapping: the Square Kilometre Array case. *Mon. Not. Roy. Astron. Soc.*, 466(3):2736–2751, 2017.

[283] Andrej Obuljen, Francisco Villaescusa-Navarro, Emanuele Castorina, and Matteo Viel. Baryon Acoustic Oscillations reconstruction with pixels. *JCAP*, 09:012, 2017.

[284] Atsushi Taruya, Takahiro Nishimichi, Shun Saito, and Takashi Hiramatsu. Nonlinear evolution of baryon acoustic oscillations from improved perturbation theory in real and redshift spaces. *Physical Review D*, 80, 2009.

[285] Stefano Anselmi, Pier-Stefano Corasaniti, Glenn D. Starkman, Ravi K. Sheth, and Idit Zehavi. Linear point standard ruler for galaxy survey data: Validation with mock catalogs. *Phys. Rev. D*, 98(2):023527, 2018.

[286] Alexandre Barreira, Giovanni Cabass, Kaloian D. Lozanov, and Fabian Schmidt. Compensated Isocurvature Perturbations in the Galaxy Power Spectrum. *JCAP*, 07:049, 2020.

[287] T. M. C. Abbott et al. Dark Energy Survey Year 1 Results: A Precise H0 Estimate from DES Y1, BAO, and D/H Data. *Mon. Not. Roy. Astron. Soc.*, 480(3):3879–3888, 2018.

[288] G. J. Hill et al. The Hobby-Eberly Telescope Dark Energy Experiment (HETDEX): Description and Early Pilot Survey Results. *ASP Conf. Ser.*, 399:115–118, 2008.

[289] G. Hill, K. Gebhardt, Eiichiro Komatsu, N. Drory, P. MacQueen, Jonathan Adams, R. Blanc, G. andKoehler, M. Rafal, Martin Roth, A. Kelz, C. Gronwall, R. Ciardullo, and D. Schneider. The hobby-eberly telescope dark energy experiment (hetdex): Description and early pilot survey results. 06 2008.

[290] Prateek Agrawal, Francis-Yan Cyr-Racine, David Pinner, and Lisa Randall. Rock 'n' Roll Solutions to the Hubble Tension. 4 2019.

[291] Tristan L. Smith, Vivian Poulin, and Mustafa A. Amin. Oscillating scalar fields and the Hubble tension: a resolution with novel signatures. *Phys. Rev. D*, 101(6):063523, 2020.

[292] J. Colin Hill, Evan McDonough, Michael W. Toomey, and Stephon Alexander. Early dark energy does not restore cosmological concordance. *Phys. Rev. D*, 102(4):043507, 2020.

[293] Mikhail M. Ivanov, Evan McDonough, J. Colin Hill, Marko Simonović, Michael W. Toomey, Stephon Alexander, and Matias Zaldarriaga. Constraining Early Dark Energy with Large-Scale Structure. *Phys. Rev. D*, 102(10):103502, 2020.

[294] J. W. Henning et al. Measurements of the Temperature and E-Mode Polarization of the CMB from 500 Square Degrees of SPTpol Data. *Astrophys. J.*, 852(2):97, 2018.

[295] Anton Chudaykin, Dmitry Gorbunov, and Nikita Nedelko. Combined analysis of Planck and SPTPol data favors the early dark energy models. *JCAP*, 08:013, 2020.

[296] Balakrishna S. Haridasu, Matteo Viel, and Nicola Vittorio. Sources of H_0-tension in dark energy scenarios. *Phys. Rev. D*, 103(6):063539, 2021.

[297] Licia Verde, Emilio Bellini, Cassio Pigozzo, Alan F. Heavens, and Raul Jimenez. Early Cosmology Constrained. *JCAP*, 04:023, 2017.

[298] David Polarski and Alexei A. Starobinsky. Isocurvature perturbations in multiple inflationary models. *Phys. Rev. D*, 50:6123–6129, 1994.

[299] Andrei D. Linde and Viatcheslav F. Mukhanov. Nongaussian isocurvature perturbations from inflation. *Phys. Rev. D*, 56:535–539, 1997.

[300] Andrew R Liddle and Anupam Mazumdar. Perturbation amplitude in isocurvature inflation scenarios. *Phys. Rev. D*, 61:123507, 2000.

[301] David Langlois and Alain Riazuelo. Correlated mixtures of adiabatic and isocurvature cosmological perturbations. *Phys. Rev. D*, 62:043504, 2000.

[302] A. Notari and A. Riotto. Isocurvature perturbations in the ekpyrotic universe. *Nucl. Phys. B*, 644:371–382, 2002.

[303] David H. Lyth, Carlo Ungarelli, and David Wands. The Primordial density perturbation in the curvaton scenario. *Phys. Rev. D*, 67:023503, 2003.

[304] Francesc Ferrer, Syksy Rasanen, and Jussi Valiviita. Correlated isocurvature perturbations from mixed inflaton-curvaton decay. *JCAP*, 10:010, 2004.

[305] Miao Li, Chunshan Lin, Tower Wang, and Yi Wang. Non-Gaussianity, Isocurvature Perturbation, Gravitational Waves and a No-Go Theorem for Isocurvaton. *Phys. Rev. D*, 79:063526, 2009.

[306] Daniel Grin, Olivier Dore, and Marc Kamionkowski. Compensated Isocurvature Perturbations and the Cosmic Microwave Background. *Phys. Rev. D*, 84:123003, 2011.

[307] Jussi Valiviita, Matti Savelainen, Marianne Talvitie, Hannu Kurki-Suonio, and Stanislav Rusak. Constraints on scalar and tensor perturbations in phenomenological and two-field inflation models: Bayesian evidences for primordial isocurvature and tensor modes. *Astrophys. J.*, 753:151, 2012.

[308] Ian Huston and Adam J. Christopherson. Isocurvature Perturbations and Reheating in Multi-Field Inflation. 2 2013.

[309] Adam J. Christopherson. Cosmological Perturbations: Vorticity, Isocurvature and Magnetic Fields. *Int. J. Mod. Phys. D*, 23(11):1430024, 2014.

[310] Chen He, Daniel Grin, and Wayne Hu. Compensated isocurvature perturbations in the curvaton model. *Phys. Rev. D*, 92(6):063018, 2015.

[311] Chung-Pei Ma and Edmund Bertschinger. Cosmological perturbation theory in the synchronous and conformal Newtonian gauges. *Astrophys. J.*, 455:7–25, 1995.

[312] B. P. Schmidt et al. The High-Z Supernova Search: Measuring Cosmic Deceleration and Global Curvature of the Universe Using Type Ia Supernovae. *The Astronomical Journal*, 507(1):46–63, 1998.

[313] Hasti Khoraminezhad, Matteo Viel, Carlo Baccigalupi, and Maria Archidiacono. Constraints on the Spacetime Dynamics of an Early Dark Energy Component. *JCAP*, 07:039, 2020.

[314] Michel Chevallier and David Polarski. Accelerating universes with scaling dark matter. *Int. J. Mod. Phys.*, D10:213–224, 2001.

[315] Eric V. Linder. Exploring the expansion history of the universe. *Phys. Rev. Lett.*, 90:091301, 2003.

[316] Francesca Perrotta, Carlo Baccigalupi, and Sabino Matarrese. Extended quintessence. *Phys. Rev.*, D61:023507, 1999.

[317] Wayne Hu, Daniel J. Eisenstein, Max Tegmark, and Martin J. White. Observationally determining the properties of dark matter. *Phys. Rev.*, D59:023512, 1999.

[318] Tobias Basse, Ole Eggers Bjaelde, Steen Hannestad, and Yvonne Y. Y. Wong. Confronting the sound speed of dark energy with future cluster surveys. 2012.

[319] Roland de Putter, Dragan Huterer, and Eric V. Linder. Measuring the Speed of Dark: Detecting Dark Energy Perturbations. *Phys. Rev. D*, 81:103513, 2010.

[320] Domenico Sapone and Martin Kunz. Fingerprinting dark energy. *Phys. Rev. D*, 80:083519, 2009.

[321] R. Laureijs and others [Euclid collaboration]. Euclid Definition Study Report. 2011.

[322] P. A. Abell and others [LSST collaboration]. LSST Science Book, Version 2.0. 2009.

[323] K. N. Abazajian and others [CMB-S4 collaboration]. CMB-S4 Science Book, First Edition. 2016.

[324] Kyungjin Ahn. How the Density Environment Changes the Influence of the Dark Matter-baryon Streaming Velocity on Cosmological Structure Formation. *Astrophys. J.*, 830(2):68, 2016.

[325] Rennan Barkana and Abraham Loeb. Scale-dependent bias of galaxies from baryonic acoustic oscillations. *mnras*, 415(4):3113–3118, August 2011.

[326] Selim C. Hotinli, James B. Mertens, Matthew C. Johnson, and Marc Kamionkowski. Probing correlated compensated isocurvature perturbations using scale-dependent galaxy bias. *Phys. Rev. D*, 100(10):103528, 2019.

[327] Patrick McDonald. Toward a measurement of the cosmological geometry at Z 2: predicting lyman-alpha forest correlation in three dimensions, and the potential of future data sets. *Astrophys. J.*, 585:34–51, 2003.

[328] Christian Wagner, Fabian Schmidt, Chi-Ting Chiang, and Eiichiro Komatsu. Separate Universe Simulations. *Mon. Not. Roy. Astron. Soc.*, 448(1):L11–L15, 2015.

[329] Lori M. Lubin, Renyue Cen, Neta A. Bahcall, and Jeremiah P. Ostriker. The Baryon fraction and velocity temperature relation in galaxy clusters: Models versus observations. *Astrophys. J.*, 460:10, 1996.

[330] R. Sadat and A. Blanchard. New light on the baryon fraction in galaxy clusters. *Astron. Astrophys.*, 371:19, 2001.

[331] Umberto Maio. Gas distribution, metal enrichment, and baryon fraction in Gaussian and non-Gaussian universes. *Class. Quant. Grav.*, 28:225015, 2011.

[332] Stacy McGaugh and Erwin de Blok. The baryon fraction distribution and the tully-fisher relation. *ASP Conf. Ser.*, 136:210, 1998.

[333] Stéphane Ilic, Marian Douspis, and Alain Blanchard. The baryon fraction in clusters and scaling relations in LCDM from X-ray and Planck data. *PoS*, FFP14:083, 2016.

[334] Emmanouil Papastergis, Andrea Cattaneo, Shan Huang, Riccardo Giovanelli, and Martha P. Haynes. A direct measurement of the baryonic mass function of galaxies & implications for the galactic baryon fraction. *Astrophys. J.*, 759:138, 2012.

[335] Shant Baghram. Measuring the baryon fraction in cluster of galaxies with Kinematic Sunyaev Zeldovich and a Standard Candle. *JCAP*, 02:015, 2019.

[336] Gilbert P. Holder, Kenneth M. Nollett, and Alexander van Engelen. On Possible Variation in the Cosmological Baryon Fraction. *Astrophys. J.*, 716:907–913, 2010.

[337] J. Weller, J. P. Ostriker, P. Bode, and L. Shaw. Fast identification of bound structures in large N-body simulations. *Monthly Notices of the Royal Astronomical Society*, 364(3):823–832, 12 2005.

[338] T. Padmanabhan. *Theoretical Astrophysics*. Cambridge University Press, 2000.

[339] E. Lifshitz. Republication of: On the gravitational stability of the expanding universe. *J. Phys. (USSR)*, 10(2):116, 1946.

[340] Eleonora Di Valentino, Olga Mena, Supriya Pan, Luca Visinelli, Weiqiang Yang, Alessandro Melchiorri, David F. Mota, Adam G. Riess, and Joseph Silk. In the Realm of the Hubble tension − a Review of Solutions. 3 2021.

[341] Will Handley. Curvature tension: evidence for a closed universe. *Phys. Rev. D*, 103(4):L041301, 2021.

[342] Seshadri Nadathur, Paul Carter, and Will Percival. A Zeldovich reconstruction method for measuring redshift space distortions using cosmic voids. *Mon. Not. Roy. Astron. Soc.*, 482(2):2459–2470, 2019.

[343] Seshadri Nadathur, Paul M. Carter, Will J. Percival, Hans A. Winther, and Julian Bautista. Beyond BAO: Improving cosmological constraints from BOSS data with measurement of the void-galaxy cross-correlation. *Phys. Rev. D*, 100(2):023504, 2019.

[344] Stefano Anselmi, Glenn D. Starkman, and Ravi K. Sheth. Beating non-linearities: improving the Baryon Acoustic Oscillations with the linear point. *Mon. Not. Roy. Astron. Soc.*, 455(3):2474–2483, 2016.

[345] Stefano Anselmi, Pier-Stefano Corasaniti, Ariel G. Sanchez, Glenn D. Starkman, Ravi K. Sheth, and Idit Zehavi. Cosmic distance inference from purely geometric BAO methods: Linear Point standard ruler and Correlation Function Model Fitting. *Phys. Rev. D*, 99(12):123515, 2019.

[346] Stefano Anselmi, Glenn D. Starkman, Pier-Stefano Corasaniti, Ravi K. Sheth, and Idit Zehavi. Galaxy Correlation Functions Provide a More Robust Cosmological Standard Ruler. *Phys. Rev. Lett.*, 121(2):021302, 2018.

[347] Márcio O'Dwyer, Stefano Anselmi, Glenn D. Starkman, Pier-Stefano Corasaniti, Ravi K. Sheth, and Idit Zehavi. Linear point and sound horizon as purely geometric standard rulers. *Phys. Rev. D*, 101:083517, Apr 2020.

[348] G. Parimbelli, S. Anselmi, M. Viel, C. Carbone, F. Villaescusa-Navarro, P. S. Corasaniti, Y. Rasera, R. Sheth, G. D. Starkman, and I. Zehavi. The effects of massive neutrinos on the linear point of the correlation function. *JCAP*, 01:009, 2021.

[349] Farnik Nikakhtar, Ravi K. Sheth, and Idit Zehavi. Laguerre reconstruction of the correlation function on Baryon Acoustic Oscillation scales. 1 2021.

[350] Dmitriy Tseliakhovich and Christopher Hirata. Relative velocity of dark matter and baryonic fluids and the formation of the first structures. *Physical Review D*, 82(8), Oct 2010.

[351] Will J. Percival, Carlton M. Baugh, Joss Bland-Hawthorn, Terry Bridges, Russell Cannon, Shaun Cole, Matthew Colless, Chris Collins, Warrick Couch, Gavin Dalton, Roberto De Propris, Simon P. Driver, George Efstathiou, Richard S. Ellis, Carlos S. Frenk, Karl Glazebrook, Carole Jackson, Ofer Lahav, Ian Lewis, Stuart Lumsden, Steve Maddox, Stephen Moody, Peder Norberg, John A. Peacock, Bruce A. Peterson, Will Sutherland, and Keith Taylor. The 2dF Galaxy Redshift Survey: the power spectrum and the matter content of the Universe. *Monthly Notices of the Royal Astronomical Society*, 327(4):1297–1306, 11 2001.

[352] Chris Blake and Karl Glazebrook. Probing dark energy using baryonic oscillations in the galaxy power spectrum as a cosmological ruler. *Astrophys. J.*, 594:665–673, 2003.

[353] Wayne Hu and Zoltan Haiman. Redshifting rings of power. *Phys. Rev. D*, 68:063004, Sep 2003.

[354] Hee-Jong Seo and Daniel J. Eisenstein. Probing dark energy with baryonic acoustic oscillations from future large galaxy redshift surveys. *Astrophys. J.*, 598:720–740, 2003.

[355] Eric V. Linder. Baryon oscillations as a cosmological probe. *Phys. Rev. D*, 68:083504, Oct 2003.

[356] Daniel J. Eisenstein et al. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *Astrophys. J.*, 633:560–574, 2005.

[357] Shaun Cole, Will J. Percival, John A. Peacock, Peder Norberg, Carlton M. Baugh, Carlos S. Frenk, Ivan Baldry, Joss Bland-Hawthorn, Terry Bridges, Russell Cannon, Matthew Colless, Chris Collins, Warrick Couch, Ni cholas J. G. Cross, Gavin Dalton, Vincent R. Eke, Roberto de Propris, Simon P. Driver, George Efstathiou, Karl Ellis, Richard S. a nd Glazebrook, Carole Jackson, Adrian Jenkins, Ofer Lahav, Ian Lewis, Stuart Lumsden, Steve Maddox, Darren Madgwick, Bruce A. Peterson, Will Sutherland, Keith Taylor, and The 2dFGRS Team. The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications. *Monthly Notices of the Royal Astronomical Society*, 362(2):505–534, 09 2005.

[358] Chris Blake, Eyal A. Kazin, Florian Beutler, Tamara M. Davis, David Parkinson, Sarah Brough, Matthew Colless, Carlos Contreras, Warrick Couch, Scott Croom, Darren Croton, Michael J. Drinkwater, Karl Forster, David Gilbank, Mike Gladders, Karl Glazebrook, Ben Jelliffe, Russell J. Jurek, I-hui Li, Barry Madore, D. Christopher Martin, Kevin Pimbblet, Gregory B. Poole, Michael Pracy, Rob Sharp, Emily Wisnioski, David Woods, Ted K. Wyder, and H. K. C. Yee. The WiggleZ Dark Energy Survey: mapping the distance–redshift relation with baryon acoustic oscillations. *Monthly Notices of the Royal Astronomical Society*, 418(3):1707–1724, 12 2011.

[359] Shadab Alam et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. *Mon. Not. Roy. Astron. Soc.*, 470(3):2617–2652, 2017.

[360] Jonathan Blazek, Joseph E. McEwen, and Christopher M. Hirata. Streaming velocities and the baryon-acoustic oscillation scale. *Phys. Rev. Lett.*, 116(12):121303, 2016.

[361] Zachary Slepian and Daniel Eisenstein. On the signature of the baryon–dark matter relative velocity in the two- and three-point galaxy correlation functions. *Mon. Not. Roy. Astron. Soc.*, 448(1):9–26, 2015.

[362] Zachary Slepian et al. Constraining the baryon–dark matter relative velocity with the large-scale three-point correlation function of the SDSS BOSS DR12 CMASS galaxies. *Mon. Not. Roy. Astron. Soc.*, 474(2):2109–2115, 2018.

[363] Jaiyul Yoo, Neal Dalal, and Uroš Seljak. Supersonic relative velocity effect on the baryonic acoustic oscillation measurements. *Journal of Cosmology and Astroparticle Physics*, 2011(07):018–018, Jul 2011.

[364] W. L. K. Wu et al. A Measurement of the Cosmic Microwave Background Lensing Potential and Power Spectrum from 500 deg^2 of SPTpol Temperature and Polarization Data. *Astrophys. J.*, 884:70, 2019.