Scuola Internazionale Superiore di Studi Avanzati - Trieste

Doctoral Thesis

Relativistic Cosmology

from the linear to the non-linear regime

Author:
Francesca Lepori

Supervisors:
Prof. Matteo Viel
Prof. Carlo Baccigalupi

Thesis submitted in fulfillment of the requirements
for the degree of Doctor Philosophiæ
in Astrophysics

September 2018

SISSA - Via Bonomea 265 - 34136 TRIESTE - ITALY

This thesis is dedicated to
Giulia and Andrea,
for you to know
why I was not there most of the time.

ABSTRACT

Cosmology is a data-driven science and in the past decades we have seen the Λ-Cold-Dark-Matter (ΛCDM) model establishing itself as the standard model for cosmology. However, the standard model still presents several open problems; among them, explaining the mechanism of the current cosmic acceleration. In the near future, Large Scale Structure (LSS) surveys are expected to play a pivotal role in shedding light on this subject. Galaxy surveys will be able to map the galaxies distribution on very large volumes with greatly improved statistics, while new techniques such as the intensity mapping applied to the 21 cm neutral hydrogen emission promise to explore the evolution of the Universe until the epoch of reionization. Therefore, it is of crucial importance to combine this huge effort from the observational side with a correct model for the observable quantities.

During my PhD I worked on several topics in the field of Large Scale Structure (LSS), which is considered here as a laboratory to test fundamental aspects of gravity and cosmology. The thesis is divided in three parts. In part I we will introduce the basics of modern cosmology. This introduction does not aim to be all-encompassing, but it will specifically address the topics that are relevant in understanding the main body of the thesis, i.e. part II and part III.

Part II of the thesis concerns the largest scales that we will be able to test with the future generation of galaxy surveys. These scales can be studied within the framework of linear perturbation theory. However, a full relativistic treatment is necessary in order to take into account all the horizon effects that are predicted by General Relativity and in general by any metric theory of gravity. The fact that photons travel from the source to the observer in an inhomogeneous universe introduces volumes and redshift correction to the observed quantities. In the linear regime they include the standard redshift-space distortions (RSD) also known as the Kaiser effect, the gravitational lensing magnification, the Doppler effect and gravitational redshift, Shapiro time-delay, Sachs-Wolfe and integrated Sachs-Wolfe effects. The standard LSS analysis includes the Kaiser effect, while the other relativistic effects are often neglected. In part II we investigated the relevance of

the relativistic effects for future planned galaxy survey, focusing on their impact on cosmological tests and parameter estimation (chapter 4 and chapter 5) and their detectability (chapter 6).

Part III of the thesis focuses on the highly non-linear regime of the LSS. The vector degrees of freedom are often neglected in modelling the cosmic velocities. This approximation works fairly well on large scales, i.e. where linear theory applies. However, on non-linear scales it is well-known that vorticity is generated. In chapter 7 we investigated numerically the generation of vorticity with the recently-born relativistic N-body code *gevolution*. Even if the generation of vorticity is a purely newtonian effect, a relativistic treatment extends the newtonian approach: for example, it allows to investigate the interplay between the vorticity and the vector degrees of freedom in the metric.

Finally, in chapter 8 we summarize the results and draw possible extensions of the work presented in this manuscript.

LIST OF PUBLICATIONS

The results presented in this thesis are based on the publications listed below.

- F. Lepori, E. Di Dio, M. Viel, C. Baccigalupi and R. Durrer,
  The Alcock Paczyński test with Baryon Acoustic Oscillations: systematic effects for future surveys,
  J. Cosmol. Astropart. Phys. **2017** 02 020 [1606.03114].

- E. Villa, E. Di Dio and F. Lepori,
  Lensing convergence in galaxy clustering in LambdaCDM and beyond,
  J. Cosmol. Astropart. Phys. **2018** 04 033 [1711.07466].

- F. Lepori, E. Di Dio, E. Villa and M. Viel,
  Optimal galaxy survey for detecting the dipole in the cross-correlation with 21 cm Intensity Mapping,
  J. Cosmol. Astropart. Phys. **2018** 05 043 [1709.03523].

- G. Jelic-Cizmek, F. Lepori, J. Adamek and R. Durrer,
  The generation of vorticity in cosmological N-body simulations,
  J. Cosmol. Astropart. Phys. **2018** 09 006 [1806.05146].

ACKNOWLEDGEMENTS

The moment to fill the last blank page of this thesis has eventually arrived and I would like to sincerely say "thank you" to every single one who, either directly or indirectly, contributed to the work here contained and supported me throughout the past four years.

Firstly, I want to express my gratitude to Martin Kunz and Giovanni Marozzi for accepting being examiners of my PhD defense and for spending time reading this thesis. I want to thank especially Giovanni for sending me his constructive comments that improved the quality of this manuscript and Martin for coming to Trieste and spending his time here paying attention to my defense, rather than chilling in Barcola.

I want to thank Matteo for having guided me through the PhD forest, from the end of the first year until now. I want to warmly thank you for understanding which topic was most interesting and suitable for me, for giving me the freedom I needed to work at my best and overall for contrasting my cosmic pessimism with your always positive and constructive attitude.

A large part of what is written in this thesis would have never seen light without the contribution of Enea. After every discussion or email exchange, I still cannot help thinking of how much I learned from your influence and how much is still left to be learned. I will always be deeply in debt with you for your guidance and patience.

I would like to thank the people in SISSA I had the pleasure and honor to work with during my PhD, who also heavily contributed to this thesis. I want to thank Carlo for insightful discussions and for his invaluable contribution to my scientific growth. I need to reserve particular thanks to Eleonora, for sharing with me her deep knowledge of Cosmology and General Relativity and for her wise and energetic advice concerning *Life, the Universe and Everything*.

I wish to thank the kind hospitality of the University of Geneva, where a part of the work reported in this thesis was carried out. I would like to express all my gratitude to Ruth for welcoming me in this rich environment, for offering me an

interesting project to work on and for deeply contributing to my education. During the 6 months I spent in Geneva I had the chance to work alongside Goran and Julian and to learn a lot from both of them. For these reasons I owe a great deal to you as well. In more general terms, I want to express my appreciation to all the members of the Geneva Cosmology Group for making me feel as part of the group, for insightful discussions... and for great fondue-based dinners!

When I started my PhD program in SISSA I met a bunch of people that throughout these years became my office mates and my closest friends. Your contribution to my PhD and to my personal development cannot be summarized in a couple of sentences (maybe a movie would not be enough!) and, for this reason, I will not even try. Anyway, I want to thank Andrej, Costantino, Elias, Gor, Lorenzo and Riki for their awesomeness and craziness and for making these years legen-dary to me.

My family deserves a special place in my acknowledgements. Even when thousands of kilometers are between me and you I know you will always be present in my life. I wish I could thank enough my parents, Mariangela and Antonio, for their unconditional love and for being always supportive, even when they do not really understand what I am doing with my life. I want to thank my sister Barbara, for being to me the go-to person whenever I have any trouble, and my sister Laura, for making me the ant of two awesome and sweet kids. Well, Giulia is great...As for Andrea, we will see!

Last, I need to dedicate a special thought to Federica, for supporting me daily against my random self-doubts and for turning any blueish day of my life into a rainbow one. For these reasons and many more the last "thank you" of this section is for you.

NOTATION AND CONVENTIONS

Unless otherwise stated, the following notation and conventions are adopted throughout the present thesis:

- Natural units are adopted, in particular the speed of light is set to $c = 1$.
- The metric signature is $(-, +, +, +)$.
- Derivatives with respect to the cosmic time are denoted with a prime, derivatives with respect to the conformal time are denoted by a dot.
- Latin indices run over spatial coordinates, while greek indices run over 4 coordinates (1 time + 3 spatial coordinates).
- For a generic function f, the Fourier transform convention is the following:

$$f(\mathbf{k}) = \int d^3x e^{i\mathbf{k}\cdot\mathbf{x}} f(\mathbf{x}),$$
$$f(\mathbf{x}) = \frac{1}{(2\pi)^3} \int d^3k e^{-i\mathbf{k}\cdot\mathbf{x}} f(\mathbf{k}).$$

- In the relativistic number count, the direction of propagation of photons is denoted with $\mathbf{n}$, while $-\mathbf{n}$ is the line-of-sight direction.

Contents

I Basic concepts in Modern Cosmology **1**

1 The smooth background **6**
1.1 The Friedmann equations 6
1.2 Distances in cosmology 9

2 Theory of cosmological perturbations **14**
2.1 Relativistic perturbation theory 15
2.2 The Vlasov equation 21
2.3 Beyond linear perturbation theory: the N-body method 24

3 Large Scale Structure observables **29**
3.1 Statistics of random fields 29
3.2 Primordial power spectrum and matter power spectrum 30
3.2.1 Baryon Acoustic Oscillations 33
3.3 Bias and redshift space distortions 34

II Relativistic effects in the Large Scale Structure **39**

4 The Alcock Paczyński test with Baryon Acoustic Oscillations: systematic effects for future surveys **44**
4.1 Introduction 44
4.2 Relativistic formalism for galaxy correlations 45
4.3 The Alcock Paczyński test 48
4.4 Methodology 49
4.5 Results 50
4.5.1 Relativistic effects on the correlation function 51

4.5.2 Galaxy bias . . . 55
4.5.3 The radial window function . . . 56
4.5.4 Impact of shot-noise on the AP test . . . 62
4.6 Summary . . . 66

5 Lensing convergence in galaxy clustering in ΛCDM and beyond 68
5.1 Introduction . . . 68
5.2 Galaxy Number Counts . . . 71
5.3 Methodology . . . 74
5.3.1 Bias on the parameter constraints . . . 76
5.3.2 Bias on the parameter estimation . . . 77
5.3.3 Surveys specifications . . . 79
5.4 Results . . . 81
5.4.1 The effect of lensing magnification: constraints on the cosmological parameters . . . 83
5.4.2 The effect of lensing magnification: bias on the cosmological parameters estimation . . . 85
5.5 Summary . . . 89

6 Optimal galaxy survey for detecting the dipole in the cross-correlation with 21 cm Intensity Mapping 92
6.1 Introduction . . . 92
6.2 Cross-correlation odd multipoles . . . 95
6.2.1 Galaxy number counts . . . 95
6.2.2 21 cm brightness temperature fluctuation . . . 96
6.2.3 21 cm - Galaxies cross-correlation . . . 97
6.2.4 Contaminations to the relativistic dipole . . . 101
6.2.5 Covariance for the cross-correlation dipole . . . 104
6.3 Halo Model approach . . . 107
6.3.1 HI model . . . 107
6.3.2 Galaxy model . . . 110
6.4 Signal-to-noise analysis . . . 112
6.5 HOD approach for modelling Luminosity-threshold galaxy catalogues . . . 116
6.6 Summary . . . 119

III LSS in the non-linear regime 122

7 The Generation of Vorticity in Cosmological Large Scale Structure 123
7.1 Introduction and motivation . . . 123
7.2 Method . . . 126
7.2.1 Computation of the velocity field . . . 127
7.2.2 Decomposition of the velocity field . . . 129
7.2.3 Method comparison . . . 131
7.3 Results . . . 133
7.3.1 Snapshots . . . 134
7.3.2 Large-scale behavior . . . 135
7.3.3 Resolution and redshift dependence . . . 138
7.3.4 Cross-spectra . . . 142
7.4 Summary . . . 146

8 Conclusions and outlook 148

Appendices

A The Alcock Paczyński test with Baryon Acoustic Oscillations: systematic effects for future surveys 154
A.1 Methodology tests . . . 154
A.2 Surveys specifications . . . 156
A.2.1 Euclid . . . 157
A.2.2 SKA . . . 157
A.2.3 DESI . . . 158

B Optimal galaxy survey for detecting the dipole in the cross-correlation with 21 cm Intensity Mapping 159
B.1 Flat-sky versus full-sky dipole . . . 159
B.2 Contribution to the wide-angle correction to the covariance . . . 163

C The Generation of Vorticity in Cosmological Large Scale Structure 166
C.1 Finite-box effect . . . 166
C.2 Comparison with the Delaunay Tessellation Field Estimator method 166

Bibliography 170

Part I

Basic concepts in Modern Cosmology

Modern cosmology pursues an extremely ambitious goal: to study the evolution of the Universe as a whole. The recent success of the discipline is even more remarkable if we think that, unlike many other branches of natural sciences, we do not have a full control of our investigation. First, the target of our experiments in unique: we have access to only one Universe for our investigation and cannot fine-tune the initial conditions of our system at will. Second, we live in a circumscribed region of our target and, therefore, we cannot observe the *whole* Universe but only that part of the Universe containing matter or radiation accessible to our astronomical observations. Despite these crucial limitations, in the past 100 years cosmology evolved from a purely mathematical/philosophical discipline into a quantitative science that is able to formulate predictions and test these predictions against observations.

Its exponential development has both theoretical and observational roots. On the theoretical side, the formulation of General Relativity by Einstein and its application to the entire universe by Einstein, Freedman, Lemaître, Robertson and Walker in the 1920s-1930s provided for the first time a scientific framework to model the global evolution of the Universe (Friedmann 1922; Lemaître 1931; Robertson 1935; Walker 1937). On the other hand, the observational guidance played a crucial role in discerning between the rich landscapes of models predicted from the theory. In 1929, Hubble first found that galaxies are receding from us and that their recessional velocities are proportional to their distance. The result of this study, the so-called Hubble diagram, provided the first clear evidence that our Universe is expanding (Hubble 1929). At this point two schools of thought emerged in the scientific community, providing different interpretations to Hubble's observations: the supporters of the steady-state theory (lead by Bondi, Gold, and Hoyle) and the proponents of the Big Bang theory [1] (among them Gamow and his student Alpher). According to the steady-state cosmology, the Universe has always been the same and galaxies and cosmic structures are constantly created in order to compensate for the expansion. In this model the average density in an expanding Universe is constant and, therefore, the total mass is not conserved. The Big Bang paradigm instead preserves the principle of mass conservation, implying that an expanding Universe must have gone through a very hot state, which characterizes the early stage of its evolution, and then cooled down. In 1948 Alpher predicted that, according to the Big Bang paradigm, the Universe should have been dominated in the early stages of its life by radiation and that the residual of this radiation should be observable today as a black body spectrum at the temperature of around

[1] Curiously, the word Big-Bang was coined by Fred Hoyle himself, maybe with the purpose of painting the theory in a negative light.

5 Kelvin (A. Alpher & C. Herman 1950). The crucial and definitive confirmation of this prediction arrived in the 60's from the discovery of the Cosmic Microwave Background (CMB) by Penzias and Wilson, who inadvertently detected the cold leftover radiation from the primordial stages of the Universe at a temperature of around 2.7 Kelvin (Penzias & Wilson 1965). The discovery of the CMB put an end to the controversial debate between the steady-state and the Big Bang theory in favor of the latter. The picture outlined above is perfectly consistent with a Universe described by the Friedmann-Lemaître-Robertson-Walker (FLRW) framework filled with matter and radiation. However, there are two missing pieces in this story that nowadays represent the two major subjects of investigation. The first piece is strictly related to the issue raised in the 1930s by Zwicky, who realized that there is a mismatch between the mass of the Coma cluster inferred from the virial theorem and the total mass measured from the observed galaxies (Zwicky 1937). Zwicky concluded that the virial mass of the cluster is dominated by an invisible form of matter. This incongruence on the Coma cluster was confirmed on galactic scales in the 1970s by Rubin's study on the galaxies rotational properties (Rubin et al. 1980). The invisible mass is now called *dark matter*. The second missing piece was discovered only 20 years ago. In 1998 two independent groups published the results of their observations of Type IA supernovae, which pointed out that, within the FLRW paradigm, the Universe is presently undergoing a phase of accelerated expansion (Perlmutter et al. 1998; Riess et al. 1998). These results were unexpected, in fact if the Universe was filled only with 'regular' matter (whose density decreases with the expansion of the Universe) the expansion rate should be decreasing. The exotic energy component driving the accelerated expansion of the Universe is denoted as *dark energy*. Nowadays, there are more evidence that points toward the existence of dark matter and dark energy components. Relevant to mention is the CMB, which in the past 20 years played a pivotal role in laying the foundations for the precision cosmology that we know. In fact, it turned out that most of the information that we can gain from the CMB does not come from its black-body spectrum, but from its tiny anisotropies that must be present in order to explain how the astrophysical sources we observe today could form. The CMB anisotropies have first been detected in 1992 by the Cosmic Background Observer (COBE) (Fixsen et al. 1996), then mapped with increasing precision first by the Wilkinson Microwave Anisotropy Probe (WMAP) (Hinshaw et al. 2013) and later by the Planck satellite (Ade et al. 2016a; Aghanim et al. 2018a). The picture that emerges from the analysis of the CMB anisotropies is a flat (or close to flat) Universe, filled with radiation, which today is a very small fraction of the

total energy budget, baryonic matter, which is the ordinary luminous matter and can be described within the standard model of particle physics, plus the already mentioned *dark matter* and *dark energy*, which make together the 95% of the total energy budget. Even if the nature of these dark components is still unknown, from the current observations a standard model for cosmology has been drawn, according to which the dark matter component is cold, i.e. pressureless, while the dark energy is provided by a cosmological constant. The standard cosmological model is labelled as the ΛCDM model. Currently, this minimal model is able to exquisitely reproduce all cosmological observations, including the already mentioned CMB (Akrami et al. 2018a; Aghanim et al. 2018a,b; Akrami et al. 2018b), the large scale structure galaxy clustering (see for example (Sanchez et al. 2014)), the cosmological distance measurements from supernovae (Riess et al. 2016) and the statistics of weak gravitational lensing (Heymans et al. 2013), with only six free parameters. The ΛCDM model provided, in the past decades, important predictions, such as the baryon acoustic oscillation feature, first detected in 2005 (Eisenstein et al. 2005), and the CMB polarization, first detected in 2002 (Kovac et al. 2002). Some mild evidence of discordance between different data sets have been recently pointed out (Addison et al. 2016). The strongest tension that has been highlighted so far concerns the present Hubble rate: measurements from the Planck collaboration (Aghanim et al. 2018a), in agreement with galaxy clustering analysis (Zhang et al. 2018), present a 3-σ tension with the local measurements carried out with supernovae (Riess et al. 2016), which prefer larger values for the present expansion rate. This discrepancy has been largely discussed in the literature (see for example (Bernal et al. 2016)) and modifications to the standard model, such as dynamical dark energy or modifications to the early Universe physics, have been advocated to alleviate it. Also, a minor internal anomaly has been found within the Planck data sets: the amplitude of the lensing contribution to the temperature power spectrum presents a 2-σ tension with the ΛCDM expectation, while the direct estimation of the lensing potential from the 4-point function does not present any tension (Calabrese et al. 2008; Motloch & Hu 2018). Despite these small anomalies that will be investigated in-depth in the near future, modern cosmology provides a simple and highly predictive standard model based on a theory of gravity that currently passed all the experimental tests (see (GRAVITY Collaboration et al. 2018) for a recent test of General Relativity near a supermassive black hole) and has, as its main targets, the following:

- Unveiling the nature of its dark components, for which a fully satisfying physical theory currently does not exist.

- Extending the current tests of the gravity theory to scales that were not available for our observation in the past.

This introductory part of the thesis will be devoted to the description of the ΛCDM model. The introduction is further divided into three chapters. In chapter 1 we briefly review the homogeneous and isotropic solutions of the Einstein equations and the different concepts of distance that are relevant for the observations in an expanding Universe. Chapter 2 is devoted to the theory of cosmological perturbations, with a particular attention to the dark matter perturbation theory, which is the most relevant component for the formation of the large scale structure. Finally, chapter 3 introduces the relevant statistical quantities that are needed in order to extract cosmological information from LSS observables.

Chapter 1

THE SMOOTH BACKGROUND

1.1 The Friedmann equations

The standard cosmological model is based on the following key assumptions:

- General Relativity (GR) describes the gravitational interactions on cosmological scales. According to GR, space-time is a 4-dimensional manifold $\mathcal{M}$, endowed with a Lorentzian metric g. Particles and photons move along geodesic trajectories and the dynamics of the metric is determined by the Einstein equations:

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{1.1}$$

  where G is the gravitational constant, $G_{\mu\nu}$ is the Einstein tensor which encodes the geometrical properties of our spacetime, while $T_{\mu\nu}$ is the stress energy tensor, describing its energy and matter content.

- The universe is statistically homogeneous and isotropic at *large* scales.

- The energy content of the universe can be modelled as a mixture of barotropic fluids, i.e. the pressure of each fluid can be expressed as a function of its density alone.

In this chapter, we will assume that the universe is perfectly homogeneous and isotropic, while the theory of cosmological perturbations on top of the herein described background will be summarized in the next chapter. Under these assumptions, it is possible to prove that the 4D space-time must admit a submanifold which is maximally symmetric (Weinberg 1972). Therefore, the full metric $g_{\mu\nu}$ can be expressed in the following form:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a^2(t)\gamma_{ij}dx^i dx^j, \tag{1.2}$$

where t denotes the cosmic time, a is the scale factor and γ_{ij} is the metric of the 3D maximally symmetric submanifold,

$$\gamma_{ij}dx^i dx^j = \frac{dr^2}{1-Kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{1.3}$$

The spatial curvature K is a constant which determines the topology of the spacetime. In this thesis, the discussion will always be restricted to a spatially-flat background, consistent with the current observations (Ade et al. 2016a; Aghanim et al. 2018a).

For $K = 0$, equation (1.3) simplifies to

$$\gamma_{ij}dx^i dx^j = dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{1.4}$$

The metric is often expressed in terms of the conformal time η

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = a^2(\eta)(-d\eta^2 + \gamma_{ij}dx^i dx^j), \tag{1.5}$$

where

$$\eta(t) = \int_0^t \frac{dt'}{a(t')}. \tag{1.6}$$

Scale factor, physical and conformal time are directly related to a primary observable is cosmology: the gravitational redshift. The gravitational redshift z measures the shift in frequency that a light ray moving toward the observer experiences because of the time-evolution of the scale factor: $1 + z = a_0/a$. Here a_0 and a denote the scale factor at the present and emission time, respectively. Since the scale factor is generally normalized to 1 at the present time, the relation between redshift and scale factor can be simply written as

$$1 + z = \frac{1}{a}. \tag{1.7}$$

The evolution of the background metric is fully encoded in the scale factor and its derivatives. In particular, a fundamental quantity in cosmology is the Hubble rate H:

$$H(t) = \frac{a'}{a} = \frac{\dot{a}}{a^2} = \mathcal{H}a^{-1}, \tag{1.8}$$

where a prime and a dot denote the derivatives with respect to the cosmic and conformal time, respectively, and $\mathcal{H}$ is the conformal Hubble parameter.

The expansion history of the Universe depends on the matter and energy content of the Universe itself. The symmetry of the metric implies that the stress-energy tensor for each fluid component X necessarily takes the same form as for an isotropic perfect fluid (Weinberg 1972)

$$(T_{\mu\nu})_X = (\rho_X + P_X)u_\mu u_\nu + P_X g_{\mu\nu}, \tag{1.9}$$

where u_μ is the 4-velocity of the fluid component, while the density ρ_X and the pressure P_X are the eigenvalues of the stress-energy tensor in equation (1.9).

The Einstein equations for the metric in equation (1.2) and the sources in (1.9) reduce to the following set of equations

$$H^2 = \frac{8\pi G}{3} \sum_X \rho_X, \tag{1.10}$$

$$\frac{a''}{a} = -\frac{4\pi G}{3} \sum_X (\rho_X + 3P_X), \tag{1.11}$$

while the conservation of the stress-energy tensor provides the following conservation law for the fluid components individually, i.e.

$$\nabla_\mu T_{\mu\nu} = 0 \implies \rho'_X = -3(\rho_X + P_X)H. \tag{1.12}$$

The equations (1.10) and (1.11) are known as *Friedmann* equations. Since the conservation law is not an independent relation between the unknown variables, it is necessary to provide a closure relation for solving the system, i.e. an equation of state (EOS) which relates for each component the pressure of the barotropic fluid to its density. The most common hypothesis is to assume a linear relation [1]

$$P = w\rho, \tag{1.13}$$

where w is a constant which depends on the nature of the fluid. For a generic fluid with $w =$ constant, the Friedmann equations lead to the following density evolution

$$\rho \propto a^{-3(1+w)}. \tag{1.14}$$

In the standard cosmological model the energy content of the Universe is made of several components, which coexist and dominate its evolution during the different phases of the Universe:

- *Radiation*, including photons and neutrinos, which is characterized by the equation of state $w = 1/3$.
- Dust, or *cold matter*, which is pressureless matter and, consequently, its EOS is $w = 0$.
- Vacuum energy, or *cosmological constant*, whose density is constant in time and thus requires $w = -1$.

[1] Here we omit the underscript X to denote that pressure and density refer to each fluid component.

It is common to divide the evolution of the Universe in three phases: the *radiation era*, which is its primordial phase and whereby radiation is the main component, the *matter dominated era*, which is the stage at which cosmic structure forms and the most recently started phase in which the dominant component is the cosmological constant. The transition between the radiation and the matter dominated era, according to the current measurements, happened at $z_{\rm eq} \approx 3400$, while the cosmological constant density overcame the matter density at $z_\Lambda \approx 0.3$.

The Friedmann equations are often conveniently expressed in terms of the critical density parameters Ω_X :

$$\Omega_X = \frac{\rho_X}{\rho_{\rm c}}(z=0), \tag{1.15}$$

where $\rho_{\rm c}$ is the critical density $\rho_{\rm c} = 3H^2/8\pi G$ and the underscript $X = {\rm r, m}, \Lambda$ denotes the radiation, matter and cosmological constant, respectively. In terms of the critical parameters, the Hubble rate can be written as

$$H = H_0\sqrt{\Omega_{\rm r}a^{-4} + \Omega_{\rm m}a^{-3} + \Omega_\Lambda}, \tag{1.16}$$

where $\Omega_{\rm X}(z) = \frac{\rho_X}{\rho_{\rm c}}(z)$ [2].

The present Hubble rate H_0 is called *Hubble constant* and can be parametrized in terms of the dimensionless parameter h:

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

The Hubble constant has been measured independently from the CMB (Ade et al. 2016a; Aghanim et al. 2018a) and type Ia Supernovae (SNe) in the local Universe (Riess et al. 2016). The two measurements are in tension of $> 3\sigma$ (local measurements prefer larger values of H_0) and it is still unclear if this is a hint of new physics or if any systematic effect was not properly taken into account in one of the two analysis.

In the next section we will see how the Hubble constant can be inferred from measurements of the cosmological distances.

1.2 Distances in cosmology

The expansion of the Universe has important consequences for the measurements of cosmological distances (see, for example, (Hogg 1999) for a concise review). In fact, we do not directly observe the proper distance or the proper size of the astronomical objects, but we infer these quantities from measurements of their redshift, their

[2] In this thesis, unless the redshift dependence is made explicit, Ω_X will denote the critical parameter for the X component at redshift 0.

luminosities or their angular size. In an Euclidean space all these conceptually different definitions of distances coincide. However, this is not true in an expanding Universe, where they depend on the geometrical properties of the Universe itself.

The first natural distance that can be defined from the FLRW metric is the distance travelled by a photon from a source at cosmic time $t_{\rm E}$ and detected by an observer at time $t_{\rm o}$. Since photons travel along null geodesic, we have $ds^2 = 0$, i.e.

$$r_{\rm com}(z) \equiv \int_{t_{\rm E}}^{t_{\rm o}} \frac{dt}{a(t)} = \int_0^z \frac{d\tilde{z}}{H(\tilde{z})}, \tag{1.18}$$

where z is the redshift of the source, while the observer is at $z = 0$. The quantity in (1.18) is the **comoving radial distance**. For a spatially curved Universe, it is relevant to introduce a *transverse* comoving distance, which is the comoving distance between two objects at a given angular separation. However, since we restrict this discussion to a spatially flat Universe, the *transverse* comoving distance coincide with the comoving distance in (1.18) and, therefore, there is no need to distinguish the two concepts.

The **proper physical distance** is the distance travelled by a photon as it would be measured by a comoving observer, i.e the distance that can be inferred by its proper time:

$$r_{\rm phys}(z) \equiv \int_{t_{\rm E}}^{t_{\rm o}} dt = a(z) r_{\rm com}(z). \tag{1.19}$$

As we mentioned above, comoving distance and proper physical distance cannot be directly observed. Therefore, it is important to relates these quantities to the observed distance.

The angular diameter distance for an object of proper transverse size $L_{\rm phys}$ and angular extension $\theta << 1$ is

$$D_{\rm A} = \frac{L_{\rm phys}}{\theta}, \tag{1.20}$$

and it is related to the comoving distance by

$$D_{\rm A}(z) = \frac{r_{\rm com}(z)}{1+z}. \tag{1.21}$$

Therefore, the angular distance of an object today is larger than it was in the past, due to the expansion of the Universe.

The luminosity distance is related to another crucial observable, the flux received from a source

$$D_{\rm L} = \sqrt{\frac{L}{4\pi F}}, \tag{1.22}$$

where F is the observed flux and L is the bolometric luminosity. Unlike the angular diameter distance, the luminosity distance at present time is lower than it was in the past. This is a consequence of the expansion of the Universe, in fact the number of photons emitted per unit area is diluted by the expansion.

The relation between luminosity distance and comoving distance is the following.

$$D_{\rm L}(z) = (1+z) r_{\rm com}(z). \tag{1.23}$$

Equations (1.18) and (1.23) provide the so called distance-to-redshift relation. In the local Universe, it can be used to directly measure the Hubble constant. In fact, the equations (1.18) can be approximated to

$$z = H_0 r_{\rm com}, \tag{1.24}$$

where r is the distance to the source and the redshift z coincides with its recessional velocity. This relation was observed for the first time by Hubble, who studied a sample of "extragalactic nebulae", i.e. galaxies, at $z < 0.003$ and presented the relations (1.24) in its seminal paper in 1929 (Hubble 1929). This was the first evidence of the expansion of the Universe.

At higher redshift, equation (1.24) is not accurate. However, the distance-to-redshift relation can be used to test the expansion rate of the Universe. What is strictly necessary for this task is to have a sample of standard candles, i.e. astronomical objects that emit with the same intrinsic luminosity L and whose redshift can be measured. An example of standard candles are type Ia Supernovae, i.e. Supernovae explosion caused by the merging of a white dwarf with another star in a binary system. The measurements of the luminosity distance from type Ia Supernovae shook the scientific community at the end of the last century (Perlmutter et al. 1998; Riess et al. 1998). Comparing the measurements of the luminosity distance for around 50 objects and the theoretical expectation for a homogeneous and isotropic Universe, it was inferred that the current expansion of the Universe is accelerating. Today, apart from galaxies and type Ia Supernovae, other astrophysical objects are used or have been proposed to be used to test the distance-to-redshift relation, for example the Gamma-Ray bursts (Amati & Della Valle 2013) or Quasars (Risaliti & Lusso 2015). Furthermore, the recent detection of the gravitational waves (GWs) signal from a binary neutron star inspiral (Abbott et al. 2017a) opened the door to the possibility of using the GWs as standard sirens, i.e. the analogue of standard candles for GWs. In fact, the distance to the source can be inferred from the GWs signal, while the electromagnetic counterpart can be used to estimate the redshift. A first measurement of the Hubble constant from a single GWs event

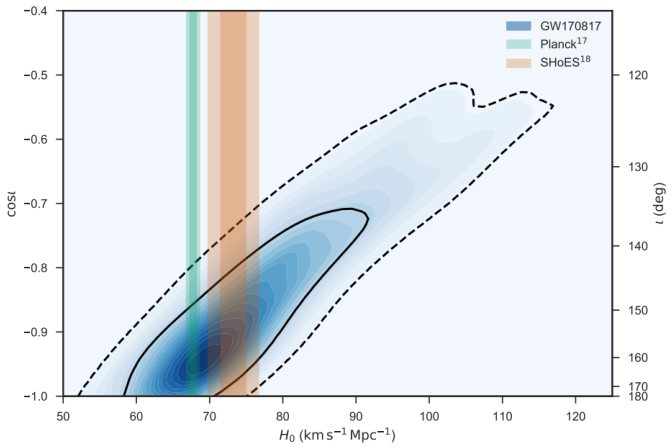


Figure 1.1: A plot from (Abbott et al. 2017b) which shows the constraints of the Hubble constant from different measurements. The green and orange bands show the 1-σ (dark shaded region) and $2-\sigma$ (light shaded region) constraints from CMB (Ade et al. 2016a) and SNe measurements (Riess et al. 2016), respectively. The blue contours represent the constraints from the GWs emission from the merger of the binary neutron star system and its electromagnetic counterpart. Continuous and dashed lines represent the 1-σ and 2-σ contour, respectively. In the GWs measurements, the Hubble constant is degenerate with the cosine of the inclination angle of the system (the angle between the line of sight and the angular momentum of the binary system). The cosine of the inclination angle is shown on the y-axis.

has been recently presented (Abbott et al. 2017b). Since it is based on a single event, the resulting measurement is not very accurate and it is consistent with both the CMB measurement and the local measurements from SNe (see figure 1.1). However, with the detection of a large number of events the statistics is expected to improve and to help to shed light on the H_0 tension.

Within the ΛCDM model, the current accelerated expansion of the Universe is interpreted as the effect of cosmological constant in the Einstein equations. However, the cosmological constant leads to some serious theoretical issues: a fine-tuning problem and a coincidence problem (Weinberg 1989; Frieman et al. 2008). Regarding the fine-tuning problem, the cosmological constant enters in the theory of gravity in the same way as the quantum vacuum energy does. Therefore, the cosmological constant that we estimate from observations will be a sum of a vacuum energy contribution $\Lambda_{\rm vac}$ plus the extra degree of freedom in our theory Λ. i.e.

$$\Lambda_{\rm obs} = \Lambda_{\rm vac} + \Lambda. \tag{1.25}$$

The contribution to the observed cosmological constant from the quantum vacuum can be computed in Quantum Field Theory (QFT) and it depends on the ultraviolet

cutoff of the theory. If we choose the Planck scale to be the cutoff energy, the calculation leads to a value of $\Lambda_{\rm vac}$ which is around 120 orders of magnitude larger than the measured $\Lambda_{\rm obs}$. Therefore, we need to fine-tune the value of Λ in order to compensate for the effect of the vacuum energy. Furthermore, the dark energy density predicted for the cosmological constant is comparable today to the matter density, despite the fact that matter and dark energy exhibit a different redshift evolution (the equality time, according to the current observations, happened at $z_\Lambda \approx 0.3$). This is known as *coincidence* problem. The anthropic principle is often advocated in order to solve or alleviate the coincidence problem. The issues outlined above motivated the development of alternative explanations for the accelerated expansion that have a similar phenomenology to the cosmological constant. These alternative models can have, as a starting point, the introduction of a new form of fluid or field in our Universe (dynamical dark energy models) (see, for example, (Caldwell et al. 1998; Chiba et al. 2000; Khoury & Weltman 2004)), they can involve modification to the gravity theory (modified gravity models) (see (Clifton et al. 2012) for a general review) or can give up the Copernican principle (void models) (Clarkson 2012).

The current tests on the alternatives to the cosmological constant are based on one of the following approaches:

- Starting from first principles, one can derive the evolution equations for the background and the perturbations for a specific model and computes the observables that can be tested with data.

- Parametrizing the modifications to the evolution of the background and/or the perturbations (Kunz 2012). This phenomenological approach is certainly more general. However, it presents the drawback of including modifications of gravity that do not necessarily correspond to a physical model. Later in this thesis we will consider departure from General Relativity through a specific parametrization of the perturbations equations (see chapter 5).

- Describing in a unified framework a generic class of scalar theory in terms of unknown time-dependent function. This is the so called Effective Field Theory (EFT) approach in the context of dark energy (Gubitosi et al. 2013; Bloomfield et al. 2013; Piazza & Vernizzi 2013). A subset of the Effective Field theories is given by the Horndenski models (Horndeski 1974), which describes, in full-generality, the class of scalar-tensor theories, with second-order equations of motion in the fields which are universally coupled to gravity.

A comprehensive analysis of alternatives to the ΛCDM model based on CMB data can be found in Ref. (Ade et al. 2016b).

Chapter 2

THEORY OF COSMOLOGICAL PERTURBATIONS

The previous chapter was devoted to the description of the homogeneous and isotropic Universe, its geometrical properties and the evolution equations. However, the FLRW is a highly idealized model. We know today from the CMB anisotropies that since the very early stages of the Universe some deviations from this model were present and that these small perturbations grew as a consequence of their mutual gravitational interactions to form the structures that we see today: stars, galaxies, clusters, filaments, voids and so on.

The standard paradigm for cosmology prescribes that the initial conditions for the primordial fluctuations are generated during an early phase of the Universe's life which is known as inflation. Inflation is a phase of accelerated expansion, due to the dominance of the vacuum energy density associated with the potential of a scalar field, the so-called inflaton field. It was first proposed by Alan Guth in 1981 to solve the well-known flatness, the horizon and the monopole problems (Guth 1981) . However, the main merit of this paradigm is to provide a natural way to generate the primordial perturbations that break the perfect homogeneity and isotropy and allowed the observed structures to form (see (Bartolo et al. 2004) and references therein).

Since this thesis is mostly focused on the late-time Universe, we do not enter into the details of the inflationary paradigm, but we will just consider the statistical properties of the primordial fluctuations generated during the standard slow-roll inflationary model, which gives an almost scale invariant power spectrum. This aspect will be clarified in chapter 3. The aim of this chapter will be instead to outline how the tiny fluctuations generated during the inflation phase evolved from the early to the late-time Universe.

2.1 Relativistic perturbation theory

In GR, perturbation theory on the gravitational field consists of splitting the full space-time metric into a background $\bar{g}_{\mu\nu}$, which is a known solution of the Einstein equations, plus some perturbations $\delta g_{\mu\nu}$ which are assumed to be small,

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu}. \tag{2.1}$$

The metric perturbations are related to the stress-energy tensor perturbations by the Einstein equations,

$$\delta G_{\mu\nu} = 8\pi G \delta T_{\mu\nu}, \tag{2.2}$$

where $\delta T_{\mu\nu} = T_{\mu\nu} - \bar{T}_{\mu\nu}$ denotes the perturbations of the stress-energy tensor and we subtracted the background equations from the full system. From now on all the background quantities will be denoted with a bar.

Scalar-vector-tensor decomposition

In cosmology the background solution is the FLRW Universe described in the previous chapter. A general parametrization for the perturbed metric is the following

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + a^2(\eta) h_{\mu\nu}, \tag{2.3}$$

with

$$h_{\mu\nu} = -2A d\eta^2 - 2V_i d\eta dx^i + 2H_{ij} dx^i dx^j. \tag{2.4}$$

The metric perturbations V_i and H_{ij} can be decomposed into scalar, vector and tensor components

$$V_i = \nabla_i B + B_i, \qquad \text{with } \nabla_i B^i = 0, \tag{2.5}$$

$$H_{ij} = H_{\rm L}\delta_{ij} + \partial_{\langle i}\partial_{j\rangle} H_{\rm T} + \partial_{(i} H_{j)}^{(V)} + H_{ij}^{(T)}, \tag{2.6}$$

where

$$\partial_{\langle i}\partial_{j\rangle} H_{\rm T} \equiv \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\right) H_{\rm T}, \tag{2.7}$$

$$\partial_{(i} H_{j)}^{(V)} \equiv \frac{1}{2}(\partial_i H_j^{(V)} + \partial_j H_i^{(V)}). \tag{2.8}$$

In an analogous way, we can define the perturbations for the stress-energy tensor:

$$\delta T_0^0 = -\bar{\rho}\delta, \tag{2.9}$$

$$\delta T_0^i = (\bar{\rho} + \bar{P}) v^i, \tag{2.10}$$

$$\delta T_j^i = \delta P \delta_j^i + \Pi_j^i. \tag{2.11}$$

The 00 component represents the density perturbation: δ is the density contrast $\delta = (\rho - \bar{\rho})/\bar{\rho}$. The $i0$ component is the momentum density of the fluid, v^i is the peculiar velocity field, i.e. the fluid velocity with respect to the Hubble flow. The latter can be further decomposed into a scalar irrotational component and a divergenceless vector part:

$$v^i = -\nabla^i v + v^i_{\rm R} \qquad \text{with } \nabla_i v^i_{\rm R} = 0. \tag{2.12}$$

It is useful to introduce the divergence field θ and the vorticity ω, which are the divergence and the curl of the velocity field, respectively

$$\theta = \partial_i v^i, \tag{2.13}$$

$$\omega^i = \epsilon^{ijk}\partial_j v_k, \tag{2.14}$$

where ϵ^{ijk} denotes the Levi-Civita symbol. The ij components of the stress-energy tensor perturbations depend on the pressure perturbation δP and on the anisotropic stress Π^i_j. The latter can be further decomposed into $\Pi^i_j = \partial_{\langle i}\partial_{j\rangle}\Pi + \partial_{(i}\Pi^{(V)}_{j)} + \Pi^{(T)}_{ij}$.

The gauge issue in cosmology

The decomposition performed above splits the metric perturbations into 4 + 4 + 2 scalar, vector, tensor degrees of freedom, respectively. However, due to the general covariance property of the GR field equations, some of these degrees of freedom are redundant and not physical. In fact, the metric perturbations are not uniquely defined, they depend on the specific slicing of the spacetime and it is of crucial importance to distinguish the physical perturbations from the fictitious ones. There are two possible approaches, which are in a way complementary, to this issue:

1. To define gauge invariant perturbations and solve the corresponding gauge invariant equations. The formalism that follows this approach is the gauge invariant perturbation theory (Bardeen 1980; Kodama & Sasaki 1984; Mukhanov et al. 1992; Durrer 1994, 2008).

2. To fix a convenient gauge, i.e. a specific time slicing, and to compute observable quantities in this gauge. In fact, observables are physical and do not depend on the gauge. Therefore, even if the starting point is not fully general, the results are.

There are many gauges that are commonly adopted in cosmology, here a list of the gauges that will be employed in the following chapters:

- **Newtonian** or **longitudinal gauge**, which fixes $B = H_{\rm T} = 0$, and neglects vector and tensor perturbations. The metric can be then expressed as a function of the Bardeen potentials Ψ and Φ, which are the gauge invariant scalar perturbations:

$$ds^2 = a^2(\eta)[-(1+2\Psi)d\eta^2 + (1-2\Phi)\delta_{ij}dx^i dx^j]. \tag{2.15}$$

- **Synchronous gauge**, for which the metric components g_{00} and g_{0i} are unperturbed

$$ds^2 = ds^2 = a^2(\eta)[-(d\eta^2 + (\delta_{ij} + h_{ij})dx^i dx^j], \tag{2.16}$$

  and h_{ij} can be decomposed into two scalars, two vectors and two tensors degrees of freedom (see (Ma & Bertschinger 1995) for details).

- **Poisson gauge**, which is the generalization of the Newtonian gauge to include vector and tensor perturbations. The metric takes the form

$$ds^2 = a^2(\eta)[-(1+2\Phi)d\eta^2 - 2B_i d\eta dx^i + (1-2\Psi)\delta_{ij}dx^i dx^j + h_{ij}dx^i dx^j], \tag{2.17}$$

  where B_i is transverse and h_{ij} is transverse and traceless, i.e. $\delta^{ij}\partial_i B_j = \delta^{ij}h_{ij} = \delta^{jk}\partial_k h_{ij} = 0$.

The linearized Einstein equations in Newtonian gauge

The decomposition described above is particularly useful to study the linear evolution of perturbations since, at linear order in perturbations theory, the evolution of scalars, vectors and tensors decouples. Therefore, the three categories can be treated separately. In vacuum, both scalar and vector modes vanish, while the tensor modes, which represent gravitational waves, can propagate. For a non empty Universe, scalar perturbations represent the density and the gravitational potential sourced by irrotational matter and therefore they represent the seeds that grow to form structures. Vectors are related to the frame-dragging effect in the metric and the rotational velocity of the cosmological fluid. Most of the inflationary models do not contemplate the generation of vector modes in the primordial Universe and, even if they were present due to some unknown mechanism, at the linear level they are expected to decay in a matter dominated era due the expansion of the Universe. We will see in part III that non-linearities in the late-time Universe generate vector modes, i.e. vorticity, due to shell-crossing. However, in the linear treatment of cosmological perturbations they can safely be neglected and they will

be explicitly neglected up to chapter 7 in this thesis . Tensor modes are expected to be produced in the primordial Universe and they are expected to leave an imprint in the CMB polarization power spectrum. They have not been detected so far, but there is huge effort in the CMB community to hunt for their signal (see Akrami et al. (2018b) for the current status of the constraints on the tensor modes). However, they have not been investigated in the work presented in the next chapters and, therefore, they will be neglected throughout this thesis.

Focusing on the scalar modes alone, the linearized Einstein equations in Newtonian gauge, for a comoving Fourier mode k, can be written as follows

$$k^2\Phi + 3\mathcal{H}\dot{\Phi} + 3\mathcal{H}^2\Psi = -4\pi G a^2 \bar{\rho}\delta, \tag{2.18}$$

$$k^2(\Phi' + \mathcal{H}\Psi) = 4\pi G a^2(\bar{\rho} + \bar{P})\theta, \tag{2.19}$$

$$\ddot{\Phi} + \mathcal{H}(2\dot{\Phi} + \dot{\Psi}) + (2\dot{\mathcal{H}} + \mathcal{H}^2)\Psi + \frac{1}{3}k^2(\Phi - \Psi) = 4\pi G a^2 \delta P, \tag{2.20}$$

$$k^2(\Phi - \Psi) = 12\pi G a^2(\bar{\rho} + \bar{P})\sigma, \tag{2.21}$$

where the right-hand sides are sourced by all the contributors to the matter-energy content. The quantity σ is the so called anisotropic stress, related to the traceless part of δT^i_j as follows

$$\left(\partial^i\partial_j - \frac{1}{3}\delta^i_j\right)\delta T^i_j = (\bar{\rho} + \bar{P})\sigma. \tag{2.22}$$

The equations (2.18), (2.19), (2.20), (2.21) are derived from the time-time, the time-space, the trace of the space-space and the off-diagonal term of the space-space components of the Einstein equations, respectively. Only two of these equations are independent. By combining (2.18) and (2.19) we obtain the following equation

$$k^2\Phi = -4\pi G a^2 \bar{\rho} D, \tag{2.23}$$

where $D \equiv \bar{\rho}\delta + 3\mathcal{H}(\bar{\rho} + \bar{P})\theta/k^2$ is the density contrast in synchronous gauge, and equation (2.23) is the relativistic equivalent of the Poisson equation.

The set of equations discussed above is accompanied with the conservation of the stress-energy tensor $\nabla_\mu T^\mu_\nu = 0$, which applies separately to each component. However, the stress-energy conservation is a consequence of the Bianchi identity and it does not provide an independent equation. In conformal newtonian gauge, assuming that the perturbations are isentropic, the conservation law can be expressed as (Ma & Bertschinger 1995):

$$\dot{\delta} = -(1 + w)(\theta - 3\dot{\Phi}) - 3\dot{\mathcal{H}}(c_s^2 - w)\delta, \tag{2.24}$$

$$\dot{\theta} = -\mathcal{H}(1 - 3w)\theta - \frac{\dot{w}}{1 + w}\theta + \frac{c_s^2}{1 + w}k^2\delta - k^2\sigma + k^2\Psi, \tag{2.25}$$

where $c_s^2 \equiv \delta P/\delta\rho$ is the squared adiabatic sound speed, that is related to the background equation of state by $c_s^2 = w + \rho dw/d\rho$. For a single component in the stress-energy tensor we are provided with 4 independent equations and the following 6 degrees of freedom: $\{\Psi, \Phi, \delta, \theta, \sigma, c_s^2\}$. Therefore, in full generality the Einstein equations and the conservation law for the stress-energy tensor are not sufficient to fully characterise the system. For a perfect fluid we can reduce the number of degrees of freedom by setting to zero the velocity dispersion and by fixing the adiabatic sound speed of the fluid to be $c_s^2 = w$:

$$\sigma = 0, \qquad \delta P = c_s^2 \delta\rho. \tag{2.26}$$

However, if these conditions do not hold for the fluid under consideration one needs to consider the full Boltzmann equation for the phase space distribution of the fluid $f(\eta, \mathbf{x}, \mathbf{v})$ (Dodelson 2003)

$$\frac{df}{d\eta} = C[f_X], \tag{2.27}$$

where the right-hand side represents the collision term, which depends on the interaction between all the fluid components. In the case of photons, which is relevant for the CMB observations, the collision term is crucial to include the Compton scattering between photons and electrons before decoupling. Although very relevant for modern cosmology, we are mainly interested in the large scale structure of the Universe. At the time where structures start to form, baryons and dark matter are coupled through gravity: baryons fall into the dark matter potential well and form galaxies in the most overdense regions. Therefore, in first approximation we want to study the evolution of dark matter.

Assuming that no anisotropic stress is present ($\Psi \equiv \Phi$) at the linear level and assuming a constant equation of state for all the species $c_s^2 = w$, the evolution equations for the dark matter component in a perturbed FLRW Universe consist of one evolution equation for the gravitational potential and one evolution equation for the density contrast:

$$\ddot{\Phi} + 3(1 + c_s^2)\mathcal{H}\Phi + c_s^2 k^2 \Phi = 0, \tag{2.28}$$

$$\ddot{\delta} + \mathcal{H}\dot{\delta} = -k^2\Phi + 3\mathcal{H}\dot{\Phi} + 3\ddot{\Phi}. \tag{2.29}$$

The sound speed in equation (2.28) depends on which components is predominant in the cosmological era under consideration. Therefore, different stages of the evolution will present different solutions. The system of equations (2.28) and (2.29) can be solved analytically in two regimes, the super-horizon limit and the sub-horizon limit.

Super-horizon solution

The super-horizon limit assumes that the scales of interest are much larger than the Hubble horizon $1/\mathcal{H}$, i.e. $k << \mathcal{H}$. This condition is particularly relevant in the early Universe, when the super-horizon condition is fulfilled for most of the modes. Under this assumption, the third term in equation (2.28) can be neglected and the evolution of the gravitational potential is determined by

$$\ddot{\Phi} + 3(1 + c_s^2)\mathcal{H}\dot{\Phi} = 0. \tag{2.30}$$

Equation (2.30) admits a constant solution

$$\Phi(\mathbf{k}, \eta) = \Phi(\mathbf{k}, \eta_0), \tag{2.31}$$

therefore the metric perturbations are frozen outside the Hubble horizon. This is valid at all time for the scales outside the horizon, as long as the adiabatic sound speed is constant [1]. The equation for the density contrast simplifies to

$$\ddot{\delta} + \mathcal{H}\dot{\delta} = 0. \tag{2.32}$$

Therefore, the density contrast presents a constant solution as well.

Sub-horizon solution

As the Universe is expanding, the size of the Hubble horizon increases and modes start entering the horizon. There are two cases of interest: modes that enter the horizon in the radiation era and modes that enter the horizon in the matter dominated era. For a mode that enters the horizon during the radiation era, the sound speed is $c_s^2 = 1/3$ and the conformal Hubble rate is $\mathcal{H} = 1/\eta$. Therefore, the evolution equation for the potential is

$$\ddot{\Phi} + \frac{4}{\eta}\dot{\Phi} + \frac{k^2}{3}\Phi = 0. \tag{2.33}$$

This equation admits the following solution (Dodelson 2003)

$$\Phi(\mathbf{k}, \eta) \sim \left(\sin\left(k\eta/\sqrt{3}\right) - \left(k\eta/\sqrt{3}\right)\cos\left(k\eta/\sqrt{3}\right)\right)\left(\frac{\sqrt{3}}{k\eta}\right)^3. \tag{2.34}$$

From equation (2.34) we see that there are two physical effects that determine the evolution of the gravitational potential: the expansion of the Universe suppresses

[1] At the transition from the radiation to the matter dominated there is a change in adiabatic sound speed, such that the gravitational potential decreases by a factor 9/10

the gravitational potential, while the radiation pressure forces the potential to oscillate. The evolution equation for the density perturbations is given by equation (2.29), where the source is the solution of the evolution equation for the potential. According to this equation, the density contrast evolves as

$$\delta \sim \ln(Bk\eta). \tag{2.35}$$

where B is a constant. Thus, inside the horizon, in the radiation era, the density perturbations have a logarithmic growth. The growth results to not be very efficient due to the radiation pressure.

For a mode that enters the horizon during the matter dominated era, the evolution equation for the potential reduces to an equation in the same form as equation (2.30), since the sound speed is zero. Therefore, in this phase the potential stays constant due to the fact that the gravitational growth is balanced by the expansion of the Universe.

The growth of density perturbation in a matter dominated era can be derived from equation (2.29) and the Poisson equation. However, since this is the central topic in the large scale structure formation and evolution, we prefer to derive this result from a more general framework and this will be the topic of the next section.

2.2 The Vlasov equation

In full generality, the dark matter component is fully characterised by its phase space distribution $f(\eta, \mathbf{r}, \mathbf{v})$. The evolution of the phase space distribution is given by the collisionless Boltzmann equation, i.e. the (2.27) with no collisional terms, which is also known as Vlasov equation and expresses the conservation of density in phase space. The Vlasov equation can be explicitly written in the following way:

$$\frac{\partial f}{\partial \eta} + \mathbf{v} \cdot \nabla f - \nabla \phi \cdot \frac{\partial f}{\partial \mathbf{v}} = 0, \tag{2.36}$$

where $\mathbf{r}$ denotes the comoving coordinates, related to the physical coordinates $\mathbf{x}$ by $\mathbf{x} = a(\eta)\mathbf{r}$, and $\mathbf{v}$ is the peculiar velocity, related to the physical velocity $\mathbf{v}_{\text{phis}}$ by $\mathbf{v} = \mathbf{v}_{\text{phis}} - \mathcal{H}\mathbf{r}$.

The dark matter properties are fully determined by the solution of the Vlasov equation, coupled with the Poisson equation.

The relevant dynamical quantities such as density, momentum and stress tensor, can be derived from the phase space distribution by taking its moments, i.e. by

integrating the phase space distribution in velocity space. The 0th, 1st and 2nd moments are

$$\rho \equiv \int f(\eta, \mathbf{x}, \tilde{\mathbf{v}}) d^3\tilde{v}, \tag{2.37}$$

$$\rho v^i \equiv \int f(\eta, \mathbf{x}, \tilde{\mathbf{v}}) \tilde{v}^i d^3\tilde{v}, \tag{2.38}$$

$$\rho(v^i v^j + \sigma^{ij}) \equiv \int f(\eta, \mathbf{x}, \tilde{\mathbf{v}}) \tilde{v}^i \tilde{v}^j d^3\tilde{v}. \tag{2.39}$$

An infinite number of moments for the phase space distribution function can be defined similarly, and an infinite number of evolution equations can be obtained from the moments of the Vlasov equation. The 0th moment of the Vlasov equation is the continuity equation,

$$\dot{\delta} + \nabla \cdot \{[1 + \delta]\mathbf{v}\} = 0, \tag{2.40}$$

while the 1st moment gives the Euler equation,

$$\dot{\mathbf{v}} + \mathcal{H}\mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla\Phi - \frac{1}{\rho}\nabla_i(\rho\sigma^{ij}). \tag{2.41}$$

Note that the nth moment of the Vlasov equation is the evolution equation for nth moment of the phase space distribution function and it relates the latter to the $(n+1)$th moment of the phase space distribution. Therefore, in order to close the system, one needs to truncate the system hierarchy by setting to zero the $(n+1)$th moment and solving the system of n equations.

In cosmology it is common to truncate the hierarchy at the 2nd moment, i.e. to assume that the dark matter can be modelled as a perfect fluid such that $\sigma^{ij} = 0$. The resulting system of equations are known as Newtonian fluid equations in Eulerian coordinates. In the linear regime, they simplify to

$$\dot{\delta} = -\theta, \tag{2.42}$$

$$\dot{\theta} + \mathcal{H}\theta + \frac{3}{2}\Omega_{\rm m}(a)\mathcal{H}^2\delta = 0, \tag{2.43}$$

$$\dot{\omega} + \mathcal{H}\omega = 0, \tag{2.44}$$

where the Euler equation has been decomposed into its scalar and vector part, and $\Omega_{\rm m}(a)$ is the critical parameter as a function of the scale factor.

From equation (2.44), it follows that, at the linear level, a primordial vorticity, if present, would decay in an expanding Universe as $\omega \sim 1/a$. Therefore, within linear perturbation theory, it can be safely neglected.

The remaining equations, in terms of δ and θ, can be written as a single second-order differential equation:

$$\ddot{D}(a) + \mathcal{H}\dot{D}(a) = \frac{3}{2}\Omega_{\rm m}(a)\mathcal{H}^2 D(a), \tag{2.45}$$

where $D(a)$ is the linear growth factor, defined from $\delta(\mathbf{r}, a) = D(a)\delta(\mathbf{r}, 1)$. The equation has two independent solutions, $D_+(a)$ and $D_-(a)$, which represent a growing and decaying mode, respectively. The density contrast in terms of the growing and decaying modes reads

$$\delta(\mathbf{r}, a) = D_+(a)A(\mathbf{r}) + D_-(a)B(\mathbf{r}), \tag{2.46}$$

where A and B are two generic functions of the comoving coordinate $\mathbf{r}$, and the divergence field is

$$\theta(\mathbf{r}, a) = -\mathcal{H}[f(a)D_+(a)A(\mathbf{r}) + g(a)D_-(a)B(\mathbf{r})], \tag{2.47}$$

with

$$f(a) \equiv \frac{d\ln D_+}{d\ln a} \qquad\qquad g(a) \equiv \frac{d\ln D_-}{d\ln a}. \tag{2.48}$$

In a matter dominated Universe, i.e. for $\Omega_{\rm m}(a) = 1$, the equation (2.45) has a simple analytical solution: the decaying modes is $D_- = a^{-3/2}$ and the evolution of the density contrast is determined by its growing mode $D_+ = a$. The velocity divergence is simply related to the density contrast by $\theta = -\mathcal{H}f\delta$. In the present Universe both dark energy and dark matter are relevant and the present growth factor is

$$D_+(a) = \frac{5\Omega_{\rm m}}{2}\frac{H}{H_0}\int_0^a d\tilde{a}\left(\frac{H_0}{\tilde{a}H(\tilde{a})}\right)^3. \tag{2.49}$$

In this section, starting from the Vlasov equation, we showed that in a matter dominated era the linear density perturbations grow as the scale factor. Therefore, the growth of structures in this phase is significantly more effective with respect to the logarithmic growth in the radiation dominated era.

In the next section we will go beyond linear perturbation theory, with a particular focus on the *N-body method*, the method used to solve the Vlasov-Poisson system through numerical simulations.

2.3 Beyond linear perturbation theory: the N-body method

The theory of linear perturbation outlined in the previous sections is able to describe the evolution of density and metric perturbations on large scales and in the early phases of the Universe. However, its range of validity is limited when applied to the large scale structure, especially in the low-redshift regime. For this reason, many techniques have been explored in order to reach a high level of accuracy in modelling non-linear effects in the LSS. In this section we are interested in the evolution of the density perturbations beyond the linear perturbation theory, which can be studied through the following strategies:

- Including higher order corrections within a **perturbative method** (see (Bernardeau et al. 2002) for a detailed review). This approach has the advantage of being fully analytical. However, it cannot cover the highly non-linear regime, since a perturbative approach based on the expansion of the density contrast is valid as long as $\delta \leq 1$.

- The **Effective Field Theory** of LSS (Baumann et al. 2012; Carrasco et al. 2012)[2] , which models the dark matter as an effective fluid, i.e. the perturbative variables are smoothed on scales larger than the non-linear scale in order to assure that they remain small. This approach is semi-analytical, since it introduces in the theory some extra parameters that depend on the microscopical properties of the dark matter fluid and that need to be fitted from data or from small numerical simulations.

- The **halo model** (Cooray & Sheth 2002), a semi-analytical framework based on the assumption that the dark matter mass content in the Universe is distributed into discrete quantities, the dark matter *halos*. In this framework, the statistics of the dark matter density on non-linear scales can be inferred from the statistics of the halos, that can be studied through numerical simulations. The halo model is often employed to model the properties of the dark matter tracers, such as galaxies, by simply assuming that observed tracers are generally found inside the dark matter halos. This method will be employed in chapter 6 in order to model the relation between the bias and the shot-noise of a dark matter tracer. Therefore, more details will be given in chapter 3, where we will introduce the bias, and in chapter 6, where we will report more details on its concrete application.

[2]Effective field theories can be classified as well as perturbative methods. However, we prefer to distinguish them from the fully analytical approach.

- Solving the Vlasov-Poisson system with a **numerical simulation**. Numerical simulations are the most effective way to accurately study the non-linear regime of structure formation. Nevertheless, it is important to keep in mind that they are much more time consuming that the other options and suffer from finite-volume and finite-resolution effects.

In this section we will focus on the fourth option, and in particular on how it can be realized with the N-body method (Hockney & Eastwood 1988). The N-body method consists in simulating a finite comoving volume of the Universe $V \equiv L^3_{\rm box}$. The size of the box depends on the specific phenomenon that is under investigation. It is important to keep in mind that a simulation cannot study scales which are larger than the box-size and that the largest scales available in a simulation may have a large statistical uncertainty. The dark matter phase space distribution is sampled at some discrete locations in phase space, represented by the position and velocity of an N-body particle, inside the simulation box. The number of the N-body particles $N^3_{\rm part}$ defines the mass resolution of the simulation. i.e. the mass of the individual particles,

$$m_{\rm p} = \bar{\rho}(L_{\rm box}/N_{\rm part})^3, \tag{2.50}$$

where $\bar{\rho}$ is the homogeneous density and it is fixed by the cosmology. Each N-body particle represents a fluid element or an ensemble of particles of total mass $m_{\rm p}$, to which it is assigned a position $\mathbf{r}_i$ and a velocity$\mathbf{v}_i$ (each N-body particle is single-valued in velocity space). Therefore, the sampled phase space distribution for the dark matter fluid is

$$f(\eta, \mathbf{r}, \mathbf{v}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)\delta(\mathbf{v} - \mathbf{v}_i). \tag{2.51}$$

Note that the velocity field is in general not single valued, since several particles with different velocities may contribute to it, at a fixed position. The particles positions and velocities are initialized at a certain redshift (usually in the range $50 \leq z \leq 100$), such that at the corresponding cosmic time a perturbative approach is still meaningful and provides accurate initial conditions. Initial conditions can be set either by the linear solution for the displacement field or by a higher order solutions. The particles are evolved according to the theory of gravity implemented in the N-body code. This point allows us to make a first distinction between N-body codes that evolve particles according to Newtonian gravity or General Relativity.

Newtonian codes evolve the particles according to the equation of motion

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{v}_i, \qquad m_p \frac{d\mathbf{v}_i}{dt} = \mathbf{F}_i, \tag{2.52}$$

where $\mathbf{F}_i \equiv -\nabla\Phi$ is the gravitational force acting on the i-th particle. The gravitational potential Φ is sourced by the particles distribution through the Poisson equation. The crucial point for an N-body code is a fast and accurate algorithm to compute the forces between particles. There are several ways on the market to estimate the gravitational forces. Here we list the most common choices for dark matter only simulations:

- **Particle-particle** methods (PP). It estimates the force of each particle as the sum of the contribution of the other $(N-1)$ particles. Therefore, the computational time of the forces at each time-step scales as N^2.
- **Particle-mesh** methods (PM). The masses of the particles are projected onto a regular lattice in order to estimate on each point the density field. On the regular grid, the Fast Fourier Transform (FFT) algorithm is used to solve the Poisson's equation in an efficient way. This method drastically reduces the computational cost, which now depends also on the number of grid-points $N^3_{\rm grid}$. The price to pay is that the simulation resolution is limited by the grid-spacing $\delta x = L_{\rm box}/N_{\rm grid}$.
- **Tree** methods. For each particle, the force is computed as the sum of the contribution of several hierarchical groups such that only the neighboring particles are treated individually. The contribution of distant groups of particles are approximated as if all the particles in the group were located at their center of mass. The computational time scales as $N\ln(N)$, therefore this method is more efficient with respect to the PP method.

In the PP and Tree methods the spatial resolution of the simulation is given by the *softening lenght* ϵ, a numerical artifact that is needed in order to prevent forces between very close particles to diverge. Therefore, the Newtonian forces between two particles i and j are effectively computed as

$$\mathbf{F}_{ij} = G m_i m_j \frac{\mathbf{r}_i - \mathbf{r}_j}{(|\mathbf{r}_i - \mathbf{r}_j|^2 + \epsilon^2)^{3/2}}. \tag{2.53}$$

Some codes employ a hybrid method, i.e. a combination of the previously mentioned approaches (for example the code `GADGET-2` implement the TreePM method (Springel 2005)).

Most of the Newtonian codes in the market implement hydrodynamics in their simulations. The standard implementation of the hydrodynamics equations is usually based either on a Lagrangian approach, employing a **smoothed particle hydrodynamics** (SPH) technique (Monaghan 1992), or an Eulerian approach,

where the Eulerian equations for the fluid are solved on a Cartesian mesh with **adaptive mesh refinement** (AMR) (Berger & Colella 1989). Recently, **tessellation methods** have been proposed in order to improve the accuracy of the hydrodynamical cosmological simulations (Springel 2010). These methods build an irregular grid, which is adapted to the N-body particles distribution, and solve the hydrodynamics equations by applying finite-volume methods to this evolving irregular mesh. The most common examples of tessellations are the Voronoi tessellation and the Delaunay tessellation. The Voronoi tessellation splits the simulation domain in regions containing one and only one particle each. The distance between the particle and each point belonging to its region is not greater than the distance between the same point and any other particle in the simulation. Instead, the Delaunay tessellation splits the domain in tetrahedra such that their vertices correspond to the particles' positions. The Voronoi tessellation is employed by the moving-mesh N-body code Arepo (Springel 2010), which is currently the state-of-the-art of Newtonian cosmological simulations (Vogelsberger et al. 2014).

Newtonian N-body codes have a long and successful tradition in the literature and are now able to cover a wide dynamical range for the dark-matter and baryons clustering (Efstathiou et al. 1985; Frenk et al. 2000; Springel et al. 2005, 2006; Boylan-Kolchin et al. 2009; Shandarin et al. 2012; Springel 2010; Springel et al. 2018).

However, the Newtonian approximation neglects the horizon effects that naturally arise in GR and that may have an impact on the non-linear evolution. In the past years N-body methods have been developed in order to go beyond the Newtonian approximation and treat in a self-consistent way relativistic sources (Adamek et al. 2014, 2016a,b, 2017). These methods have been implemented in the N-body code *gevolution*. *gevolution* is a particle-mesh code which employs a weak-field expansion of the Einstein equations in Poisson gauge. Instead, the stress-energy tensor which sources the Einstein equations is evolved in the full non-linear regime and the particles move according to the relativistic geodesic equations. The relativistic approach has the advantage of being self-consistent in a cosmological framework, which is based on General Relativity. However, since this approach has been explored in the recent past, it does not offer all the tools that are currently available for the Newtonian codes (such as AMR and baryons effects). In this thesis, the code *gevolution* has been used to investigate the generation of vorticity in large-scale structure. Therefore, more details on the structure of the code will be given in part III of this thesis.

To summarize, in this chapter we outlined the theory of linear cosmological perturbation in a FLRW Universe and the numerical methods that are commonly

employed to go beyond the linear approximation. Up to this point, we know which are the theoretical predictions for the evolution of theoretical quantities such as the local density contrast of dark matter and the gravitational potential. The next chapter links these theoretical results to the observables in the LSS.

Chapter 3

LARGE SCALE STRUCTURE OBSERVABLES

In the previous chapters we described the framework to study the evolution of cosmological perturbations in a FLRW background. However, we do not know *exactly* the initial conditions of our Universe, since the process that has generated the initial conditions is stochastic. The consequence is that we can only model the evolution of the Universe statistically. In this chapter, we will briefly review the statistical description of the cosmological perturbations, which is necessary in order to link the theoretical model previously discussed with observations. In section 3.1 we will discuss the random field description for the cosmological perturbations and we will define the statistics of the perturbations in terms of the correlation function and the Fourier power spectrum. In section 3.2 we describe the scale dependence of the matter power spectrum and its baryon acoustic feature. Section 3.3 will introduce the bias for the LSS tracers and the redshift space distortions effect.

3.1 Statistics of random fields

In a statistical description of cosmological perturbations, the Universe that we observe is a single realization of a statistical ensamble. Each cosmological field is a stochastic field and it is *statistically homogeneous and isotropic*, i.e. the probability distribution function is invariant under translations and spatial rotations.

The 2-point correlation function ξ_{Ψ_r} for a random field Ψ_r is

$$\langle \Psi_r(\mathbf{x}_1)\Psi_r(\mathbf{x}_2)\rangle = \xi_{\Psi_r}(|\mathbf{x}_1 - \mathbf{x}_2|), \tag{3.1}$$

where the operator $\langle\rangle$ represents the average over the ensemble.

The correlation function depends only on the distance between the two coordinates $\mathbf{x}_1$ and $\mathbf{x}_2$ as a consequence of statistical homogeneity and isotropy. We will precise later that this symmetry is broken in the observed correlation

function, since we observe in redshift space. The random fields which are relevant in cosmology are generally assumed to be *gaussian*. Under this assumption, the gaussian random field is *ergodic*, i.e. the information on its correlators can be extracted from a single realization by simply interpreting the average operator over the ensemble as a spatial average.

In the case of the density contrast, the 2-point correlation function has an intuitive physical interpretation: it represents the excess probability with respect to a uniform distribution of finding two particles in the infinitesimal volumes dV_1 and dV_2. Indeed, for a random distribution one would expect this probability to be related to the number density $\bar{n}$ as $dP = \bar{n}^2 dV_1 dV_2$, while for a not random distribution we can write this probability in terms of the density 2-point function ξ,

$$dP = \bar{n}^2[1 + \xi(r)]dV_1 dV_2. \tag{3.2}$$

Another relevant statistical quantity is the Fourier counterpart of the correlation function, the power spectrum $P(k)$. Assuming statistical homogeneity and isotropy, the power spectrum is defined as

$$\langle \Psi_r(\mathbf{k}_1)\Psi_r(\mathbf{k}_2)\rangle = (2\pi)^3 \delta_{\mathrm{D}}(\mathbf{k}_1 + \mathbf{k}_2) P_{\Psi_r}(k), \tag{3.3}$$

where δ_{D} is the Dirac delta operator, and the $\Psi_r(\mathbf{k})$ is the Fourier transform of the field $\Psi_r(\mathbf{x})$.

3.2 Primordial power spectrum and matter power spectrum

In the previous section, we defined the statistical quantities which are relevant for any random field. In this section, we use these definitions to state which are the statistical prediction for the initial condition from the inflationary paradigm and we comment on the implication of these initial condition on the shape of the matter power spectrum.

As it was mentioned before, inflation is a phase in early Universe characterized by an exponentially accelerated expansion. The inflaton field, which is the scalar field responsible for the accelerated expansion, is not homogeneous but fluctuates due to quantum mechanical effects. These fluctuations follow the expansion of the Universe and, when they become larger than the Hubble horizon, they are frozen. At the end of the inflationary era, the inflaton field decays into particles and the fluctuations of the inflaton field, which were frozen outside the horizon, generate the primordial fluctuations in the energy density. The inflationary paradigm predicts

the statistics of the initial perturbations. In terms of the primordial gravitational potential $\Phi(\mathbf{k})$, the primordial power spectrum $P_\Phi(\mathbf{k})$ can be defined as

$$\langle \Phi_{\rm in}(\mathbf{k}_1)\Phi_{\rm in}(\mathbf{k}_2)\rangle = (2\pi)^3\delta_{\rm D}(\mathbf{k}_1+\mathbf{k}_2)P_\Phi(k). \tag{3.4}$$

The primordial power spectrum can be parametrized as follows:

$$P_\Phi(k) = \frac{A_s}{k^3}\left(\frac{k}{k_*}\right)^{n_{\rm s}-1}, \tag{3.5}$$

where $A_{\rm s}$ is the amplitude of the primordial spectrum, k_* is the pivot scale and $n_{\rm s}$ is the spectral index. Inflation predicts a primordial spectrum that is almost scale-invariant, i.e. $n_{\rm s}$ is very close to 1.

In chapter 2 we discussed the evolution of linear cosmological perturbations. In the light of what we previously discussed and given the primordial power spectrum in (3.5) we can predict the shape of the matter power spectrum $P_{\rm m}(k)$. In particular we can split the matter power spectrum in three regions:

- The low-k limit. These very large scales include modes that are **today outside the horizon**. Both potential and density modes are constant in time (in Newtonian gauge). Therefore, we expect the matter power spectrum to reproduce the scale-dependence of the primordial spectrum:

$$P_{\rm m}(k) \sim k^{-3} \qquad \text{for } k < \mathcal{H}. \tag{3.6}$$

- Modes that **enter the horizon in a matter dominated era**. The density fluctuation grows as the scale factor in this stage. Modes that enters the horizon at earlier times, i.e. the smaller modes, are expected to have a larger amplitude. Therefore, the matter power spectrum is expected to grow with k and to reach a maximum amplitude corresponding to $k = k_{\rm eq}$. The slope of the power spectrum in this range can be computed from the primordial spectrum and the Poisson equation, which relates the density and the gravitational potential. The result is

$$P_{\rm m}(k) \sim k \qquad \text{for } \mathcal{H} < k < k_{\rm eq}. \tag{3.7}$$

- Modes that **enter the horizon in a radiation dominated era**. Modes that enter the horizon in the radiation era have a logarithmic growth after the horizon crossing. Therefore, after the equality we expect the power spectrum to decay as the primordial spectrum, modulated by a logarithmic factor:

$$P_{\rm m}(k) \sim k^{-3}[\ln(\text{const}\cdot k)]^2 \qquad \text{for } k > k_{\rm eq}. \tag{3.8}$$

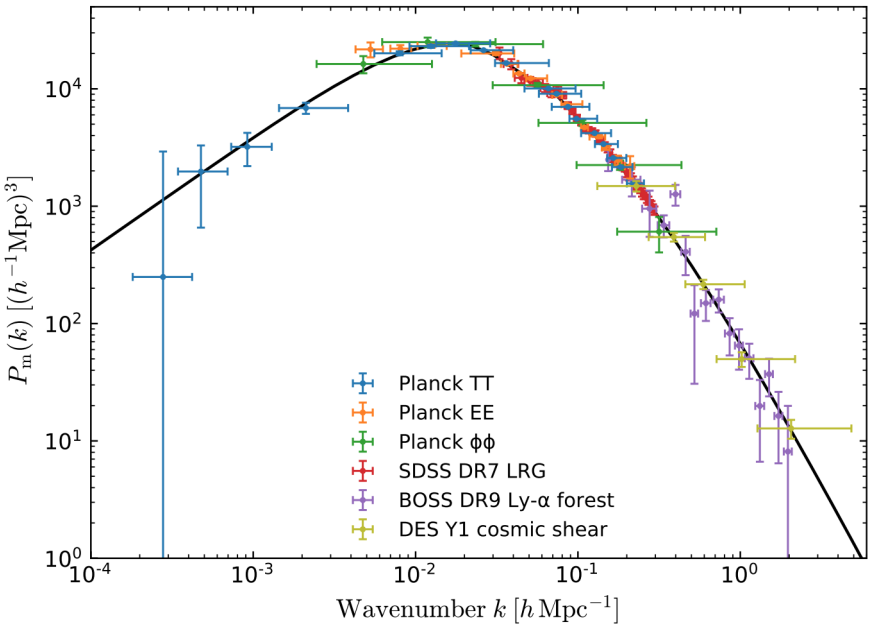


Figure 3.1: The linear matter power spectrum at $z = 0$, as measured from different cosmological probes. The figure is reprinted from (Akrami et al. 2018a). The largest scales are probed by the CMB temperature maps (blue data-points). Intermediate scales are tested through CMB polarization, CMB lensing and galaxy clustering, orange, green and red points, respectively. The small scales regime is constrained by the Lyman-α forest and cosmic shear measurements (denoted by violet and yellow data-points). The black line is the theoretical matter power spectrum (in synchronous gauge) for the best-fit parameters.

In the discussion above, we specified that super-horizon modes decay as k^{-3} in Newtonian gauge. This clarification is needed when discussing the properties of the Fourier space power spectrum since this quantity is not a direct observable and the super-horizon evolution depends on the gauge (Challinor & Lewis 2011; Bonvin & Durrer 2011). In synchronous gauge, the power spectrum does not show any divergence in the low-k limit.

The main feature of the matter power spectrum for modes that entered the horizon in the matter and radiation dominated era are displayed in figure 3.1, which was presented in the latest Planck data analysis (Akrami et al. 2018a).

In the discussion above, we did not mention the effect of the dark energy in the matter power spectrum. A cosmological constant is expected to not affect the shape of the matter power spectrum. However, the accelerated expansion is expected to contrast the dark matter clustering. Therefore, the amplitude of the power spectrum is slightly suppressed with respect to the prediction that only

consider matter as the full energy density content. Another effect that was not discussed is the impact of baryons.

3.2.1 Baryon Acoustic Oscillations

Baryons are a small fraction ($\sim 15\%$) of the total matter content. However, their presence leaves an important imprint in the matter power spectrum and correlation function: the *Baryon Acoustic Oscillations* (BAO) feature. In the radiation era, before electrons and protons combine to form neutral hydrogen (recombination), baryons are tightly coupled to photons and they behave as a single fluid. The gravitational attraction between the massive baryons tends to compress the fluid, while the radiation pressure due to photons forces the fluid to expand. The net effect is that the fluid oscillates and generate a sound wave which propagates in the early Universe. The sound wave propagates until recombination, when photons free-stream, while baryons stop to propagate and stay at the distance travelled by the photon-baryon fluid, which is called the *sound horizon*:

$$r_{\rm s} = \int_{\eta=0}^{\eta_{\rm rec}} c_{\rm s} d\eta = \int_{z_{\rm rec}}^{\infty} dz \frac{c_{\rm s}}{H(z)}, \tag{3.9}$$

where $c_{\rm s}$ is the sound speed and depends on the ratio between the baryon and the photon density: $c_{\rm s} = [3(1 + 3/4(\rho_b/\rho_\gamma))]^{-1}$. This preferred scale for baryons translates into an excess of clustering in the matter correlation function, the BAO peak, and into an oscillating feature in the matter power spectrum. The amplitude of the BAO peak depends on the ratio between the baryon density and the dark matter density, while the position of the BAO peak is given by the size of the sound horizon (equation (3.9)). The BAO scale is known at the 0.3% level from CMB measurements (Ade et al. 2016a; Aghanim et al. 2018a). Therefore, it is a *standard ruler*: combining the CMB measurements with the BAO measurements from galaxy clustering at different redshift, it is possible to reconstruct the expansion history. The BAO scale has been first measured from galaxy clustering in 2005 by the Sloan Digital Sky Survey (SDSS) at $z \approx 0.3$ (Eisenstein et al. 2005). Currently the state-of-the-art measurements come from the Baryon Oscillation Spectroscopic Survey (BOSS), both from galaxy clustering at $z < 1$ (Ross et al. 2017; Beutler et al. 2017; Vargas-Magana et al. 2016) and from the fluctuations in the Lyman-α flux and its cross-correlation with quasars at $z \sim 2.3$ (Bautista et al. 2017; du Mas des Bourboux et al. 2017). Future planned galaxy survey, such as Euclid (Amendola et al. 2013) and the Dark Energy Spectroscopic Instrument (DESI) (DESI 2015) will measure the BAO peak with a much better statistics, making of the BAO one of the major probe for the expansion of the Universe and the dark energy.

To summarize, in this section we introduced the linear matter power spectrum and the correlation function. In the next section we will show how the matter power spectrum in related to the power spectrum of an observed LSS tracer, i.e. we will discuss how to model the bias and the how the peculiar velocities of the tracer affect the measurements of the power spectrum and the correlation function.

3.3 Bias and redshift space distortions

In the previous section we defined the statistical quantities relevant to understand the properties of the dark matter fluid. However, the dark matter density and velocity cannot be directly observed. In fact, cosmological surveys detect galaxies or other tracers, whose formation is the result of complex and not-fully understood physical process. For simplicity, in this section we will assume galaxies to be our reference tracer. We denote with $\Delta_{\rm gal}$ the observable quantity for a galaxy survey, while $\delta_{\rm gal}$ refers to the local galaxy overdensity as it would be measured by an observer comoving with the source. However, the discussion can be easily generalized to any LSS tracer.

A crucial point is understanding which is the relation between the observable $\Delta_{\rm g}$ and the dark matter density contrast that is predicted from our theory. We will focus here on two effects: the galaxy bias and the redshift space distortions (RSD).

Bias

The galaxy bias relates the fluctuation in the dark matter density to the local fluctuation of the observable $\delta_{\rm gal}$ (see (Desjacques et al. 2018) for a recent review). The simplest model consists of assuming a non-linear relation in the form $\delta_{\rm gal} = g(\delta)$, that can be Taylor expanded as follows (Fry & Gaztanaga 1993):

$$\delta_{\rm gal}(\mathbf{x}, z) = \sum_{n}^{\infty} \frac{b_n(z)}{n!} \delta^n(\mathbf{x}, z). \tag{3.10}$$

This model for the bias is known as *local* bias. On very large scales, the bias is usually assumed to be *local* and *linear*, i.e. the Taylor expansion is truncated at the linear order, leading to the simple relation

$$\delta_{\rm gal}(\mathbf{x}, z) = b(z)\delta(\mathbf{x}, z). \tag{3.11}$$

On small scales, the relation (3.11) is expected to break down and the bias to exhibit a scale dependence. A semi-analytical approach to model the non-linear bias is the already mentioned *halo model* (see for example (Cooray & Sheth 2002)).

Within the *halo model*, dark matter tracers are found in the so called *halos*, regions that were dense enough to collapse into a virialised structure. Halos are biased tracers of the dark matter distribution, but unlike galaxies or any other observed tracer, their properties must be studied through N-body simulations. At a given redshift, fits from simulations provide the *halo mass function*, i.e. the halos comoving number density, and the *halo bias*, as a function of the halo mass alone. In this framework, it is possible to link the properties of the halos to the properties of a galaxy population, provided that we can model or measure from data the average number of galaxies within a halo of a certain mass. In this framework, density and bias of a given tracer depend on how massive is the halo that hosts them. A comprehensive review of the halo model is beyond the scope of this thesis, however more details on the concrete application of the halo model to study the shot-noise and the bias of a tracer will be given in chapter 6.

Redshift Space Distortions

As we mentioned in the above lines, the bias relates the dark matter distribution to the local galaxy overdensity. However, the galaxy overdensity $\delta_{\rm g}$ is not a direct observable. This is a consequence of the fact that we do not observe comoving distances directly, but a redshift survey collects photons coming from a source at a given redshift and those photons have crossed the foreground large scale structure before reaching the survey's detector. In the specific case of a galaxy survey, the true observable is the number of galaxies that are seen at a certain angle in the sky and at a given redshift. In this section we will focus on the linear redshift space distortions (see (Hamilton 1997) for a general review), i.e. the impact of the peculiar velocities on the observed overdensity. In part II we will see that redshift space distortions (RSD) is the dominant effect, but not the only one. We will give the full expression of the galaxy number count at the linear order and we will detail the several physical effects that affect the observable.

In 1987 Kaiser was the first to point out that statistical properties of galaxies are sensitive to both the dark matter distribution and the galaxies peculiar velocities (Kaiser 1987). The linear redshift space distortions effect is displayed in figure 3.2. The peculiar motion of the galaxies is a certain region toward a central overdensity will induce an additional redshift to the galaxies in front, while galaxies behind the overdensity will appear blue-shifted with respect to the Hubble flow. The net results in the observed galaxy overdensity is that the galaxies distribution will appear squashed and the measured density is enhanced. In order to see how this effect quantitatively impact the observed galaxy correlations, we denote with $\mathbf{r}_s$

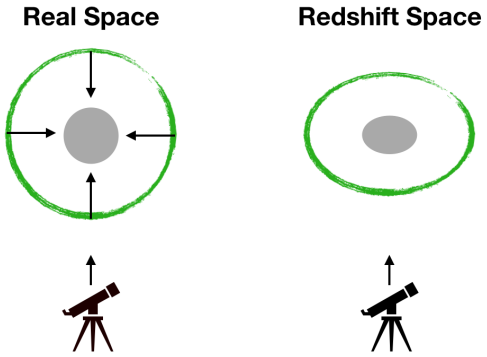


Figure 3.2: The linear RSD. A spherical overdensity in real space appears squashed in redshift-space due to the peculiar motion of galaxies.

the observed distance in redshift space and with $\mathbf{r}$ the corresponding distance in real space. An underscript s marks all redshift-space quantities.

The number of galaxies must be conserved in real and redshift space, i.e.

$$\bar{n}_{\rm g,s}(\mathbf{r}_{\rm s})d^3r_{\rm s} = \bar{n}_{\rm g}(\mathbf{r})d^3r, \tag{3.12}$$

where $\bar{n}_{\rm g}$ and $\bar{n}_{\rm g,s}$ are the galaxies number density in real and redshift space, respectively. Therefore, the galaxies overdensity in real space $\delta_{\rm gal}$ is related to the galaxies overdensity in redshift space $\delta_{\rm gal,s}$ as follows:

$$(1+\delta_{\rm gal,s}) = (1+\delta_{\rm gal})\left|\frac{d^3r}{d^3r_s}\right|, \tag{3.13}$$

where $|d^3r/d^3r_s|$ is the Jacobian of the coordinates transformation from real to redshift space.

The observed redshift has a contribution coming from the Hubble flow $\bar{z} = \mathcal{H}r$, where $r = ||\mathbf{r}||$, and a contribution coming from the peculiar velocity projected along the line of sight $-\mathbf{n}$ [1]. Thus, the observed redshift can be written as

$$z_{\rm obs} = \bar{z} - \mathbf{v}\cdot\mathbf{n}, \tag{3.14}$$

and the relation between redshift space and real space coordinates is

$$r_{\rm s} = r - \frac{\mathbf{v}\cdot\mathbf{n}}{\mathcal{H}}. \tag{3.15}$$

From (3.15), the Jacobian of the transformation is

$$\left|\frac{d^3r}{d^3r_s}\right| = \frac{r^2}{r_{\rm s}^2}\frac{dr}{dr_{\rm s}} = \Big[1-\partial_r(\mathbf{v}\cdot\mathbf{n}/\mathcal{H})\Big]^{-1}\Big[1-(\mathbf{v}\cdot\mathbf{n})/(\mathcal{H}r)\Big]^{-2}. \tag{3.16}$$

[1] We denote with $\mathbf{n}$ the direction of propagation of the incoming photons.

The second term in the Jacobian is usually neglected. This is due to the fact that, inside the Hubble horizon, derivative of the velocity field dominates over the purely velocity correction. For the moment, we neglect the second term in in (3.16). However, the expression found in this section will be generalized in the part II of this thesis. Assuming that the gradient of the radial peculiar velocity is small, we can write the Jacobian as follows

$$\left|\frac{d^3r}{d^3r_s}\right| \simeq \left[1 - \partial_r(\mathbf{v}\cdot\mathbf{n}/\mathcal{H})\right]^{-1} \simeq 1 + \partial_r(\mathbf{v}\cdot\mathbf{n}/\mathcal{H}). \tag{3.17}$$

In this approximation, the relation between the overdensity in real and redshift space is

$$\delta_{\rm gal,s} = \delta_{\rm gal} + \partial_r(\mathbf{v}\cdot\mathbf{n}/\mathcal{H}). \tag{3.18}$$

The comoving galaxy overdensity, within the linear approximation, is related to the dark matter density contrast in synchronous gauge D through a linear and local bias $\delta_{\rm g} = b_{\rm gal}D$. Therefore, in first approximation, the relation between the redshift space galaxy overdensity and the dark matter density contrast is

$$\delta_{\rm gal,s} = b_{\rm gal}D + \partial_r(\mathbf{v}\cdot\mathbf{n}/\mathcal{H}). \tag{3.19}$$

Similarly to the statistical properties of the overdensity in real space, we can compute the statistics of the overdensity in redshift space. The galaxy correlation function in redshift space, at fixed redshift, is

$$\xi_{\rm s} = \langle [b_{\rm gal}D(\mathbf{r}_1) + \partial_{r_1}(\mathbf{v}_1\cdot\mathbf{n}_1/\mathcal{H})], [b_{\rm gal}D(\mathbf{r}_2) + \partial_{r_2}(\mathbf{v}_2\cdot\mathbf{n}_2/\mathcal{H})]\rangle\,, \tag{3.20}$$

where $\mathbf{v}_1 = \mathbf{v}(\mathbf{r}_1)$ and $\mathbf{v}_2 = \mathbf{v}(\mathbf{r}_2)$.

Here we will report the results for the linear correlation function in redshift space, in the flat-sky approximation. The redshift-space correlation function can be expanded in multipoles $\xi_\ell(r)$,

$$\xi(r,\mu) = \sum_\ell \xi_\ell(r) L_\ell(\mu), \tag{3.21}$$

where L_ℓ are the Legendre polynomials of degree ℓ and the coefficients of the expansion $\xi_\ell(r)$ can be written as

$$\xi_\ell(r) = (-i)^\ell \int \frac{k^2 dk}{2\pi^2} S_\ell(z)\, j_\ell(k\,r) P_{\rm m}(k), \tag{3.22}$$

where j_ℓ are the spherical Bessel function of order ℓ and the redshift dependent coefficients S_ℓ are

$$S_0(z) = (b_{\rm gal}^2 + 2/3 f b_{\rm gal} + f^2/5), \tag{3.23}$$

$$S_2(z) = (4/3 f b_{\rm gal} + 4/7 f^2), \tag{3.24}$$

$$S_4(z) = 8/3 f^2. \tag{3.25}$$

The parameter f is the growth factor, i.e. the logarithmic growth rate of the density perturbation,

$$f \equiv \frac{d \ln \delta}{d \ln a}. \tag{3.26}$$

The redshift space corrections break the isotropy of the correlation function: not only they affect the monopole of the correlation function, but they also source a quadrupole and a hexadecapole. The monopole and the quadrupole of the correlation function have been measured from the Baryon Oscillation Spectroscopic Survey (BOSS) (Samushia et al. 2014). Since monopole and quadrupole depends on different combination of cosmological parameters, an anisotropic analysis allows to break the degeneracy between the bias and the growth factor (Samushia et al. 2014).

RSD measurements are extremely relevant for modern cosmology. From one side, they introduce an anisotropy in the correlation function which must be taken into account in order to correctly interpret the observations. Furthermore, they provide a powerful of test for the growth of structure f. In GR, the growth of structure is scale-independent and can be written as a function of the matter critical density at a given redshift,

$$f_{\rm GR}(z) = \Omega_m(z)^\gamma, \tag{3.27}$$

where $\gamma = 4/7$ (Lahav et al. 1991; Carroll et al. 1992), while alternative gravity theories predict a different relation between the growth factor and Ω_m. Therefore, through RSD measurements it is possible to test GR on cosmological scales (Guzzo et al. 2008).

As we mentioned previously, the redshift space galaxy overdensity in (3.18) is an approximation to the observed fluctuation in the galaxies count number $\Delta_{\rm g}$. In the next part of the thesis we will focus on the corrections beyond the RSD.

Part II

Relativistic effects in the Large Scale Structure

In the last section of the introduction, we discussed the linear RSD effect, the perturbation to the observed galaxy overdensity due to the peculiar velocity gradient at a given direction in the sky. As we already mentioned, RSD is not the only relativistic effect which perturbs our observable. In this part of the thesis we will introduce the full expression for the galaxy survey observable, in linear perturbation theory.

Redshift surveys measure the redshift of the source and its angular position in the sky, estimate the number of sources at a certain coordinate $N(\mathbf{n}, z)$ and its fluctuation with respect to the average number of galaxies over all the sky $\langle N \rangle (\mathbf{n}, z)$. The relative fluctuation gives the galaxy number count:

$$\Delta_{\rm gal}(\mathbf{n}, z) = \frac{N(\mathbf{n}, z) - \langle N \rangle (z)}{\langle N \rangle (z)}. \tag{3.28}$$

The observed number of galaxies and its average over all directions depend both on the density and the volume: $N(\mathbf{n}, z) = \rho_{\rm gal}(\mathbf{n}, z) V(\mathbf{n}, z)$ and $\langle N \rangle (z) = \bar{\rho}_{\rm gal}(z) \bar{V}(z)$. By expanding (3.28) at first order in perturbation theory, we obtain

$$\Delta_{\rm gal}(\mathbf{n}, z) = \frac{\delta\rho_{\rm gal}(\mathbf{n}, z)}{\bar{\rho}_{\rm gal}(z)} + \frac{\delta V(\mathbf{n}, z)}{V(z)}. \tag{3.29}$$

The first term in (3.29) is the galaxy overdensity at the observed redshift. In order to relate this quantity to the dark matter density contrast, which is predicted from theory at the background redshift $\bar{z}$, we can Taylor-expand the density fluctuation

$$\frac{\delta\rho_{\rm gal}(\mathbf{n}, z)}{\bar{\rho}_{\rm gal}(z)} \simeq \frac{\delta\rho_{\rm gal}(\mathbf{n}, \bar{z})}{\bar{\rho}_{\rm gal}(\bar{z})} + \frac{1}{\bar{\rho}_{\rm gal}} \frac{d\bar{\rho}_{\rm gal}}{dz} \delta z. \tag{3.30}$$

Since the background comoving density is conserved, i.e. $\bar{\rho}_{\rm gal} \sim (1+z)^3$, (3.29) can be expressed in terms of the galaxies overdensity as

$$\Delta_{\rm gal}(\mathbf{n}, z) = \delta_{\rm gal}(\mathbf{n}, \bar{z}) - \frac{3}{(1+\bar{z})} \delta z + \frac{\delta V(\mathbf{n}, z)}{V(z)}. \tag{3.31}$$

In order to relate the redshift and volume perturbation to the metric and velocity perturbation, one need to solve the geodesic equations for the photons trajectories. The redshift perturbation can be computed by comparing the background redshift and the redshift due to the energy shift that the photon experience by propagating from the source to the observable. The volume perturbation can be computed from the Jacobian of the coordinates transformation, which relates the position in the sky of the source and the observer's position.

The full expression has been computed by several independent groups (Yoo et al. 2009; Yoo 2010; Bonvin & Durrer 2011; Challinor & Lewis 2011; Schmidt

& Jeong 2012; Bertacca et al. 2012; Raccanelli et al. 2016b, 2014). Here we will only report the final result,

$$\begin{aligned}\Delta_{\rm gal}(\mathbf{n},z,m^*) &= b_{\rm gal}(z,m^*)D + \frac{1}{\mathcal{H}(z)}\partial_r(\mathbf{V}\cdot\mathbf{n}) \\ &+ (5s(m^*,z)-2)\int_0^{r(z)}\frac{r(z)-r}{2r(z)r}\Delta_\Omega(\Phi+\Psi)dr \\ &+ \left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2}+\frac{2-5s(m^*,z)}{r\mathcal{H}}+5s(m^*,z)-f_{\rm evo}^{\rm gal}(m^*,z)\right)(\mathbf{V}\cdot\mathbf{n}) \\ &+ (f_{\rm evo}^{\rm gal}-3)\mathcal{H}V+(5s-2)\Phi+\Psi+\frac{1}{\mathcal{H}}\dot{\Phi}+\frac{2-5s}{r(z)}\int_0^{r(z)}dr(\Phi+\Psi) \\ &+ \left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2}+\frac{2-5s}{r(z)\mathcal{H}}+5s-f_{\rm evo}^{\rm gal}\right)\left(\Psi+\int_0^{r(z)}dr(\dot{\Phi}+\dot{\Psi})\right), \end{aligned} \tag{3.32}$$

where $b_{\rm gal}(z,m^*)$ is the galaxy bias of the sources whose magnitudes are smaller than the magnitude limit of the survey m^* [2], assumed to be linear and local; D is the dark matter density fluctuation in synchronous gauge; $\mathbf{V}$ is the peculiar velocity in longitudinal gauge; V the velocity potential defined by $\mathbf{V}=-\nabla V$; η_o is the present time and $r(z)=\eta_o-\eta$ is the conformal distance at redshift z; Δ_Ω represents the angular laplacian operator.

The first line in equation (3.32) includes the local density fluctuation of galaxies and the redshift space distortions discussed in section 3.3. We will refer to these contributions to be the *standard* terms, since all current clustering analysis take them into account in the physical model for the observable.

The second line is the gravitational lensing, which includes two effects. The first effect the volume distortion due to the foreground overdensity, which bend the photons trajectories and modify the observed solid angle. The second effect is the lensing magnification, due to fact that galaxy surveys are limited in flux: they are able to detect a galaxies whose flux is above a detection threshold. Since gravitational lensing caused by the presence of foreground structures magnifies the apparent brightness, the lensing effect can have an impact in the detected number of galaxies whose flux is close to the flux limit. This effect is parametrized by the magnification bias $s(m^*,z)$, the slope of the cumulative luminosity function of the survey (Di Dio et al. 2013; Alonso & Ferreira 2015)

$$s(m^*,z)=\frac{\partial\log_{10}\bar{N}(z,m<m^*)}{\partial m^*}, \tag{3.33}$$

[2]The magnitude limit of the survey is related to the flux limit through $m^*=-\frac{5}{2}\log_{10}\left[\frac{F^*}{F_0}\right]$, where F_0 is a reference value for the flux.

where $\bar{N}(z, m < m^*)$ denotes the cumulative luminosity function

$$\bar{N}(z, m < m^*) = \int_{\ln L*}^{\infty} \phi(\eta(z), \ln L) d \ln L, \tag{3.34}$$

and $\phi(\eta(z), \ln L)$ is the luminosity function and L^* is the luminosity threshold of the survey that is related to the flux threshold by $L^* = 4\pi(1+z)^2 r^2(z) F^*$. The lensing convergence effect was a well known effect before the fully relativistic computation of the number count was presented and it was first detected in 2005 from SDSS data by cross-correlating quasars and their foreground galaxies (Scranton et al. 2005). However, most of the clustering analysis neglect its contribution to the power spectrum and correlation function.

The third line in equation (3.32) is the Doppler effect, i.e. the Doppler shift on the mean redshift induced by the peculiar velocity of galaxies at a given position in the sky. In the fourth and fifth line we collected all the local and integrated terms involving the gravitational and velocity potential.

The bias factor in the Doppler and potential correction $f_{\rm evo}(m^*, z)$ is the evolution bias, which describes the departure from a sample of sources conserved in a comoving region

$$f_{\rm evo}^{\rm gal}(m^*, z) = \frac{\partial \ln \bar{N}(z, m < m^*)}{\mathcal{H} \partial \eta} = -(1+z) \frac{\partial \ln \bar{N}(z, m < m^*)}{\partial z}. \tag{3.35}$$

Modelling both magnification and the evolution biases of a galaxy catalogue requires a prior knowledge on the luminosity function of the targeted galaxy population. The expression for the number counts in (6.3) assumes that galaxies follow geodesics

$$\partial_r \Psi = \mathbf{V}' \cdot \mathbf{n} + \mathcal{H} \mathbf{V} \cdot \mathbf{n}, \tag{3.36}$$

therefore it is valid for any metric theory of gravity.

In GR, we can compare the different contributions to the number count by writing them in terms of metric perturbations, i.e. using the relations between density, velocities and gravitational potential. From the Poisson equation we have that the density is related to the potential through two spatial derivatives, i.e. in Fourier space $D \sim (k/\mathcal{H})^2 \Phi$. From the Euler equation, the velocity along the line of sight scales as $\mathbf{V} \cdot \mathbf{n} \sim (k/\mathcal{H}) \Phi$. Therefore, the density, the redshift space distortions and the lensing corrections enter in equation (3.32) at the same parametrical order $\sim (k/\mathcal{H})^2 \Phi$. The Doppler corrections enter a the order $\sim (k/\mathcal{H}) \Phi$, while the fourth and fifth line in (3.32) are proportional to the gravitational potential, i.e. $\sim \Phi$. Therefore, assuming GR and for scales well-inside the horizon $(k/\mathcal{H}) >> 1$ we expect density, redshift space distortions and lensing to be the dominant contributions.

Moreover, the gravitational lensing is an integrated effect: we expect it to be suppressed at low redshift and become relevant at high-z. It is not surprising that all the not-standard corrections to the number count were neglected in the past without disastrous consequences: in fact past and present galaxy survey cover a small volume compared to the horizon size. However, this may not be the case for the future generation of redshift surveys, which are expected to cover a large fraction of the sky, a wider redshift range and to have a greatly improved statistics.

In this part of the thesis we investigate the relevance of the non-standard corrections to the galaxy number count, focusing mainly on three aspects. First, the relativistic formalism offers the possibility to perform model independent analysis and therefore relax the prior assumption on the cosmological parameters that are needed in standard analysis to convert angles and redshift into distances. Furthermore, neglecting the non-standard correction to the galaxy number count may affect the estimation of cosmological parameters. This issue have been studied in the context of the minimal ΛCDM model and beyond. Last but not least, the relativistic corrections are not just a systematic contaminant to our observations, but contain by themselves cosmological information. Therefore, it is crucial to investigate the perspective that we will have with the future observations to isolate them from the dominant density and redshift space distortions contributions, and eventually detect them. The first two aspects will be investigated in chapter 4 and 5, while the latter will be studied in chapter 6.

Chapter 4

THE ALCOCK PACZYŃSKI TEST WITH BARYON ACOUSTIC OSCILLATIONS: SYSTEMATIC EFFECTS FOR FUTURE SURVEYS

4.1 Introduction

The upcoming galaxy surveys will map the distribution of galaxies on a large fraction of the sky up to redshift $z \sim 2$. In order to exploit this huge amount of incoming data, an accurate model for what we will be measuring in galaxy surveys is required. Moreover, having much better statistics, we may release some of the assumptions and proceed in a more model independent way. In the past few years, the galaxy number counts have been computed including all the relativistic effects at the linear order (Yoo et al. 2009; Yoo 2010; Bonvin & Durrer 2011; Challinor & Lewis 2011) and at second order (Yoo & Zaldarriaga 2014; Bertacca et al. 2014; Di Dio et al. 2014a) in perturbation theory. Besides the known redshift space distortions correction to the local overdensity of galaxies (Kaiser 1987), other terms contribute to galaxy number counts, e.g. Doppler corrections and gravitational lensing. Not taking into account these effects in our theoretical model may bias the analysis (Cardona et al. 2016), therefore their relevance in any cosmological observable should be tested. In this chapter we investigate the relevance of these corrections for the Alcock Paczyński (AP) test (Alcock & Paczynski 1979). We start from the method proposed in (Montanari & Durrer 2012), where the AP test is performed on the Baryon Acoustic Oscillation (BAO) feature of the galaxy 2-point correlation function without any prior assumption on the cosmological parameters. Nevertheless, we will show that some prior information about the galaxy bias can improve the accuracy in determining the BAO scale from the observable quantities.

In section 4.2 we introduce the notation we employ throughout the chapter for the observed galaxy number density. In section 4.3 we present the AP test and we introduce the AP parameter. In section 4.4 we summarize the method we use to compute the galaxy 2-point correlation function and to determine the position of the acoustic peak and we outline the general strategy to investigate observational distortions of the test. More details on the methodology are given in appendix A.1. In section 4.5 we report our results: in 4.5.1, we discuss the relevance of the relativistic correction to the galaxy number count for the AP test, in 4.5.2 we introduce a linear local bias and we show how it affects our method; in 4.5.3 we study projection effects induced by a radial window function and we compute the corrective factors that must be applied to the estimated AP parameter for three future planned galaxy surveys, i.e. Euclid (Amendola et al. 2013), the Dark Energy Spectroscopic Instrument (DESI) (Levi et al. 2013) and the Square Kilometer Array (SKA) (Abdalla et al. 2015). We also analyze the impact of shot-noise and cosmic variance on the accuracy of the BAO peak determination in section 4.5.4. In section 4.6 we summarize the results of this chapter.

4.2 Relativistic formalism for galaxy correlations

We consider a Friedmann universe with linear scalar perturbations. We work in Newtonian gauge so that the line element is given by

$$ds^2 = a(\eta)^2 \left(-(1+2\Psi)\, d\eta^2 + (1-2\Phi)\, d\mathbf{x}^2 \right) , \tag{4.1}$$

where $a(\eta)$ is the scale factor, η is conformal time and the metric perturbations, Ψ and Φ, are the Bardeen potentials. We remark that working solely with observational quantities, we can fix the gauge without loss of generality.

The observed galaxy number count $\Delta_{\rm gal}(\mathbf{n}, z)$ can be schematically expressed as a sum of different contributions

$$\Delta_{\rm gal} = \Delta_g + \Delta_{\rm RSD} + \Delta_\kappa + \Delta_{\rm rel}, \tag{4.2}$$

where Δ_g is the local galaxy overdensity,

$$\Delta_g = b\ \delta \tag{4.3}$$

proportional to the density contrast in comoving gauge δ. For the sake of simplicity, we assume the bias b to be linear and local. The term $\Delta_{\rm RSD}$ in our notation includes

the linear redshift space distortions contributions, due to the peculiar motion of galaxies, plus other subdominant Doppler corrections (Di Dio et al. 2014b)

$$\Delta_{\rm RSD}(\mathbf{n},z)=\frac{1}{\mathcal{H}(z)}\partial_r(\mathbf{V}\cdot\mathbf{n})+\left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2}+\frac{2}{r\mathcal{H}}\right)(\mathbf{V}\cdot\mathbf{n})-3\mathcal{H}V, \tag{4.4}$$

where $\mathbf{V}$ is the peculiar velocity in longitudinal gauge, V the potential velocity defined by $\mathbf{V}=-\nabla V$ and r is the conformal distance.

Δ_κ is the gravitational lensing term,

$$\Delta_\kappa=-\int_0^{r(z)}\frac{r(z)-r}{r(z)r}\Delta_\Omega(\Phi+\Psi)dr, \tag{4.5}$$

where Δ_Ω is the Laplace operator on the sphere. The last term in Eq. (4.2), $\Delta_{\rm rel}$, includes subdominant local and integrated combinations of the Bardeen potentials:

$$\begin{aligned}\Delta_{\rm rel}=&-2\Phi+\Psi+\frac{1}{\mathcal{H}}\dot{\Phi}+\frac{2}{r(z)}\int_0^{r(z)}dr(\Phi+\Psi)+\\&+\left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2}+\frac{2}{r(z)\mathcal{H}}\right)\left(\Psi+\int_0^{r(z)}dr(\dot{\Phi}+\dot{\Psi})\right).\end{aligned} \tag{4.6}$$

Throughout this chapter we will neglect the effect of magnification bias and the possible evolution of the number of counts (Challinor & Lewis 2011). Even if present, these bias factors multiply terms which are subdominant in the AP test.

In terms of observational coordinates, the relevant statistical quantities are the angular correlation function $\xi(\theta,z_1,z_2)$ or the redshift dependent angular power spectra $C_\ell(z_1,z_2)$. The galaxy correlation function, under the assumption of statistical isotropy reduces to

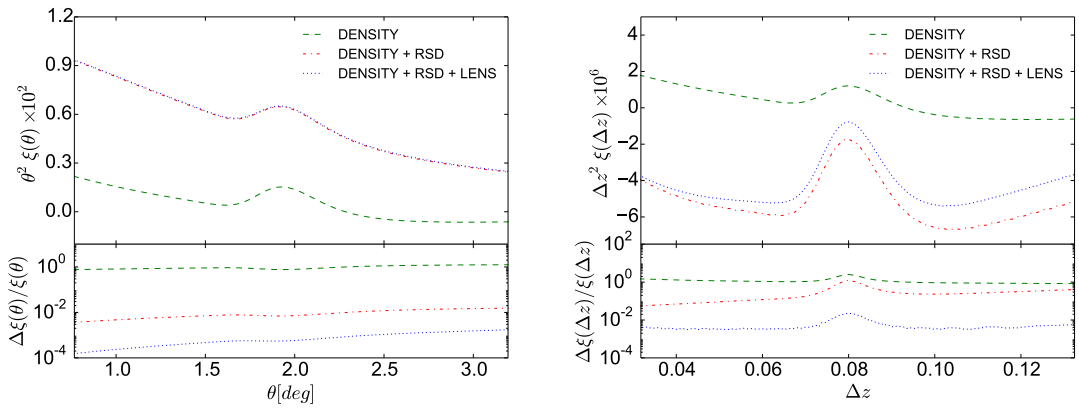
$$\xi(\theta,z_1,z_2)=\langle\Delta_{\rm obs}(\mathbf{n}_1,z_1)\Delta_{\rm obs}(\mathbf{n}_2,z_2)\rangle\,,\qquad\cos\theta\equiv\mathbf{n_1}\cdot\mathbf{n_2}, \tag{4.7}$$

where $\langle..\rangle$ denotes the ensemble average over several realizations. Observationally, this is replaced by an average over directions at fixed observed redshift and opening angle θ. In a similar way, the angular power spectrum is defined as (Bonvin & Durrer 2011)

$$C_\ell(z_1,z_2)=\langle a_{\ell m}(z_1)a^*_{\ell m}(z_2)\rangle\,, \tag{4.8}$$

where a star denotes the complex conjugate and $a_{\ell m}$ are the coefficients of the spherical harmonic expansion for $\Delta(\mathbf{n},z)$

$$\Delta_{\rm obs}(\mathbf{n},z)=\sum_{\ell m}a_{\ell m}(z)Y_{\ell m}(\mathbf{n}),\qquad a_{\ell m}\,(z)=\int d\Omega_{\mathbf{n}}Y^*_{\ell m}(\mathbf{n})\Delta_{\rm obs}(\mathbf{n},z). \tag{4.9}$$



(a) Trasverse correlation function **(b)** Longitudinal correlation function

Figure 4.1: Transverse and longitudinal correlation function computed at $z_{\rm mean} = 1.5$, partially including the corrective terms in Eq. (4.2). In the bottom panels, we show the relative differences between the exact correlation function (including all the terms) and a partial correlation function computed including only the local density term (green, dashed line), density and redshift space distortion correction (red, dash-dotted line), the previous terms plus the lensing (blue, dotted line).

The angular power spectra and the correlation functions are related by

$$\xi(\theta, z_1, z_2) = \frac{1}{4\pi}\sum_{\ell=0}^{\infty}(2\ell+1)C_\ell(z_1, z_2)P_\ell(\cos\theta), \tag{4.10}$$

where $P_\ell(\cos\theta)$ is the Legendre polynomial of degree ℓ. From Eq. (4.10) it is straightforward to define the correlation function along the line-of-sight direction by setting $\theta = 0$ and in the transverse direction by setting $z_1 = z_2$.

The correlation function at a given mean redshift $z_{\rm mean}$ along the line-of-sight is given by

$$\xi_{\parallel}(\Delta z, z_{\rm mean}) = \frac{1}{4\pi}\sum_{\ell=0}^{\infty}(2\ell+1)C_\ell(z_{\rm mean} - \Delta z/2, z_{\rm mean} + \Delta z/2) \tag{4.11}$$

while the transverse correlations are

$$\xi_{\perp}(\theta, z_{\rm mean}) = \frac{1}{4\pi}\sum_{\ell=0}^{\infty}(2\ell+1)C_\ell(z_{\rm mean}, z_{\rm mean})P_\ell(\cos\theta). \tag{4.12}$$

Figure 4.1 represents transverse and longitudinal correlation function at fixed mean redshift $z_{\rm mean} = 1.5$. In each plot the correlation function is computed gradually adding subdominant contributions. Redshift-space distortions affect considerably the correlations in both directions, while lensing does not change sensitively the transverse correlations. In the longitudinal direction, the lensing

term enhances the amplitude of the correlation function by up to 30% for $\Delta z \lesssim 0.13$, but it does not modify its shape at the BAO scale. We notice that the lensing effect is more important for pairs of galaxies with large radial separation. The contribution from the relativistic terms, i.e. Eq. (4.6), is completely subdominant and it would not be visible in figure 4.1. Therefore we neglect it in the analysis performed in the rest of the chapter. We remark that the largest relativistic correction, namely the Doppler term, is included in the redshift space distortions, see Eq. (4.4).

4.3 The Alcock Paczyński test

The Alcock Paczyński test (Alcock & Paczynski 1979), proposed for the first time in 1979, is a purely geometrical test of the cosmic expansion history performed by measuring the shape of an object expanding with the Hubble flow. When we observe an astrophysical object, we measure its shape in terms of its angular size θ and its radial extent in redshift space Δz. These two quantities depend on the comoving sizes of the source, $L_\parallel$ and $L_\perp$, and on a conversion factor: Δz is related to the Hubble expansion rate $H(z)$ by

$$\Delta z = L_\parallel H(z), \tag{4.13}$$

while the dependence of θ on the angular diameter distance is given by

$$\theta = \frac{L_\perp}{(1+z)D_A(z)}. \tag{4.14}$$

If the object is known to be spherically symmetric, i.e. $L = L_\perp = L_\parallel$, then the ratio of these two measured quantities,

$$F_{\rm AP} \equiv \frac{\Delta z}{\theta}, \tag{4.15}$$

does not depend on the physical size of the object, but only on the redshift and on the spacetime geometry

$$F_{\rm th}(z) = (1+z)D_A(z)H(z). \tag{4.16}$$

Here we explicitly distinguished the measured AP parameter $F_{\rm AP}$ from its theoretical (background) value $F_{\rm th}$.

In realistic applications, the spherical symmetry is not required to hold for single objects because the test can be applied to the galaxy clustering statistics. In fact, statistical isotropy of space implies statistical spherical symmetry for the correlation function. Furthermore, the correlation function naturally offers a robust feature for

the application of the AP test, the BAO scale. BAO high-precision measurements are one of the main target of future spectroscopic surveys (see (Aubourg et al. 2015) for a recent overview of the cosmological implication of BAO measurements).

The application of the AP test on the BAO feature in the galaxy correlation function has been proposed in (Montanari & Durrer 2012), where it has been found that the peak position of the transverse correlation function is significantly affected by the binning in redshift. This effect is due to the fact that a finite window function in the transverse direction introduces a spurious radial component in the correlation function and the BAO peak estimation must be properly corrected for. In this chapter we will extend the work presented in (Montanari & Durrer 2012), addressing also other effects that may distort the result of the test.

4.4 Methodology

We assume a ΛCDM cosmology consistent with the best fit parameters from the Planck 2015 data (Ade et al. 2016a): $h = 0.6774$, $\Omega_{cdm}h^2 = 0.1188$, $\Omega_b h^2 = 0.0223$, $\Omega_\Lambda = 1 - \Omega_m$, $\Omega_k = 0$. The primordial amplitude and spectral index are set to $A_s = 2.142 \times 10^{-9}$ and $n_s = 0.9667$.

The computation of the angular power spectrum for the observed galaxy number counts is implemented in the publicly available Boltzmann code CLASSgal (Di Dio et al. 2013), a modified version of the Cosmic Linear Anisotropy Solving System (CLASS) code (Lesgourgues 2011; Blas et al. 2011) optimized to compute accurately and efficiently the relativistic large scale observables to linear order (Di Dio et al. 2014b). The angular power spectrum is computed by running CLASSgal for the fiducial cosmology. Unless otherwise stated, we do not include any window function in the model and we do not include non-linearities.

The radial and transverse correlation functions are computed from (4.11) and (4.12) summing over a finite number of multipoles $\ell \leq \ell_{max}$. In the radial direction, the value ℓ_{max} can be set to be large enough in order to avoid spurious numerical oscillations induced by a sharp cutoff in ℓ-space. In the transverse direction we set $\ell_{max} = 20000$ and we introduced a cutoff W_ℓ to smooth numerical spurious oscillations, so that we have

$$\xi_\perp(\theta) = \frac{1}{4\pi}\sum_{\ell=0}^{\ell_{max}}(2\ell+1)C_\ell(z_{mean}, z_{mean})P_\ell(\cos\theta)W_\ell(\ell_s, \ell_x), \tag{4.17}$$

where

$$W_\ell(\ell_s, \ell_x) = \frac{1}{2}\Big(1 - \tanh\{(\ell - \ell_s + 3\ell_x)/\ell_x\}\Big). \tag{4.18}$$

The cutoff parameters ℓ_s and ℓ_x are set to be respectively $\ell_s \approx \ell_{max}$ and $\ell_x = 2000$. The estimation of the BAO peak is not affected by small variation of these parameters. We remark that the correlation function can be computed directly for density, redshift space distortions and for local terms in general, see e.g. Ref. (Matsubara 2000, 2004; Montanari & Durrer 2012). Nevertheless we prefer to use Eq. (4.11) and (4.12) to handle the integrated terms. In this way we can use the precise results obtained with CLASSgal code, instead of relying on some uncontrolled approximation.

We model the correlation functions with the following parameterization

$$\xi(x) = A \cdot e^{-(x-x_{BAO})^2/2\sigma^2} + \sum_{n=0}^{N} K_n \cdot x^n, \tag{4.19}$$

where $x = \theta$ for the transverse correlation function and $x = \Delta z$ for the radial one. In Eq. (4.19) a polynomial of degree N models the shape of the correlation function on scales unaffected by the BAO peak, while a Gaussian describes the BAO feature. We fit the data points with the template model, where the free parameters are A, x_{BAO}, σ, K_n, with $n = 0, 1, ..., N$. The BAO scale is estimated as the best fitting value of θ_{BAO} and Δz_{BAO}.

The non-linear least squares fitting is performed using the Python version of MPFIT [1], which implements the Levenberg-Marquardt method (Moré 1978).

Once the position of the BAO feature has been estimated in both radial and transverse directions, we compute the AP parameter in Eq. (4.15) and we compare it with its theoretical value $F_{\rm th}$, computed for the same fiducial cosmology. We perform the same analysis for different values of the mean redshift $z_{\rm mean}$ in the range between $z = 0.3$ and $z = 2$. The two values are always expected to coincide, within error-bars. Violations of the consistency relation

$$F_{\rm AP}(z_{\rm mean}) = F_{\rm th}(z_{\rm mean}) \tag{4.20}$$

indicates an inaccurate method for the estimation of the BAO scale. We have performed several tests of our methodology: we tested the accuracy of the AP test for different degrees of the polynomial and different template functions. The details of these tests are reported in Appendix A.1. Our results show that the sufficient accuracy in the parameterization (4.19) is reached with $N = 10$.

4.5 Results

In the previous section we summarized the method to recover the BAO peak position from the computed correlation function in both radial and transverse directions. In the next paragraphs we use this method to study possible sources that may affect the accuracy in the measurement of the AP parameter F_{AP}.

[1] http://cars9.uchicago.edu/software/python/mpfit.html

4.5.1 Relativistic effects on the correlation function

In this section we aim to understand to which extent redshift space distortions and lensing corrections in the observed galaxy overdensity affect the measurement of the AP parameter. We compare the results of the AP test considering first density perturbations only, and then adding subsequently redshift space distortions and lensing corrections. These effects change the observed correlation function in radial and transverse directions, but clearly they do not change the intrinsic BAO scale. For simplicity, in this section we assume the galaxies in our sample to be unbiased tracers of the dark matter distribution, i.e. $b = 1$.

Figure 4.2 and figure 4.3 show the radial and the transverse correlations and the corresponding fits for the three cases of interest, at different redshifts, respectively. As we have already seen in the previous sections, redshift space distortions significantly modify both the shape and the amplitude of the correlation functions in the radial and transverse directions. We remark that the ratio between the density and the redshift space distortions terms is only weakly sensitive to the mean redshift. Indeed, in linear theory only the growth factor has a weak redshift dependence. The estimated values of the position of the BAO peak are slightly affected in both directions: when we do not include RSD corrections the radial BAO peak is shifted toward larger scales, while the angular peak position $\theta_{\rm BAO}$ is shifted to smaller scales. Although both effects are small (~ 0.5 %), they sum up when applied to the AP test. This result should not be interpreted as a physical shift of the acoustic scale, rather as a numerical effect related to the model assumed for the numerical fit, which is calibrated to be valid in redshift space and not in real space.

The gravitational lensing correction does not affect the transverse correlation function, while in the radial direction it changes the amplitude, but not the shape of the correlation function. Therefore, it does not appreciably shift the peak position. In figure 4.2 we show that the lensing term reduces the amplitude of the radial correlation function at low redshifts, while it increases the amplitude at high redshifts. This is due to the fact that the lensing contribution is dominated by the negative correlation between density and lensing at low redshifts, while is dominated by the auto-correlation at high redshifts.

In figure 4.4 we show the result of the consistency test for the three cases of interest here. We find that gravitational lensing does not affect the AP test, while redshift space distortions enhance the result by about 1% if the analysis is performed redshift space compared real space.

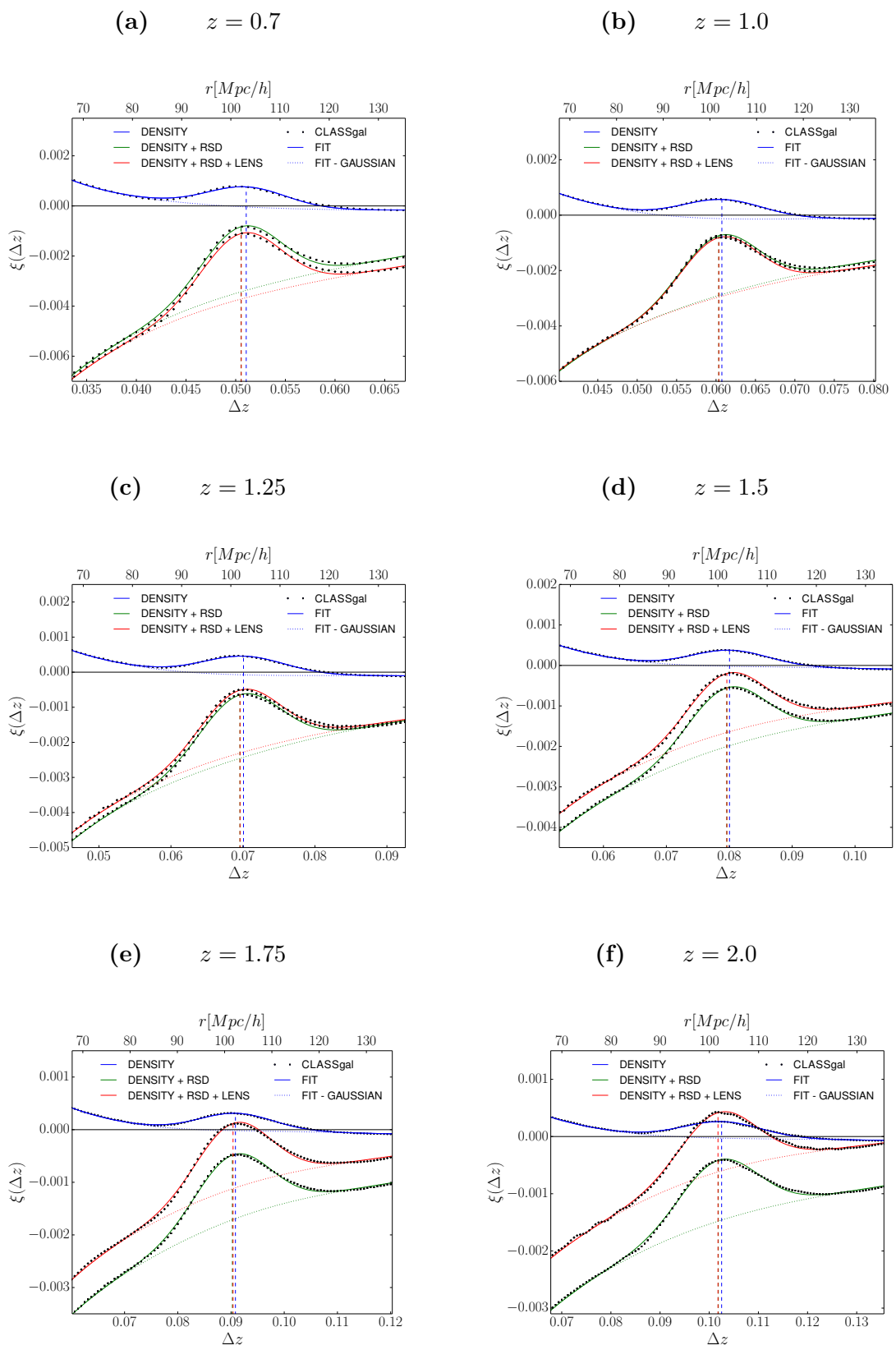


Figure 4.2: The radial correlation function and the BAO peak position computed including different contributions to the observed over density. The blue line refers to density correlations only, the green line includes also the redshift space distortions, and the red line takes into account the previous plus the lensing terms. The continuous lines refer to the best-fit model, while the dots are the CLASSgal output. The dotted lines are obtained by subtracting the Gaussian term from the best-fit model. The vertical dashed lines identify the estimated peak positions.

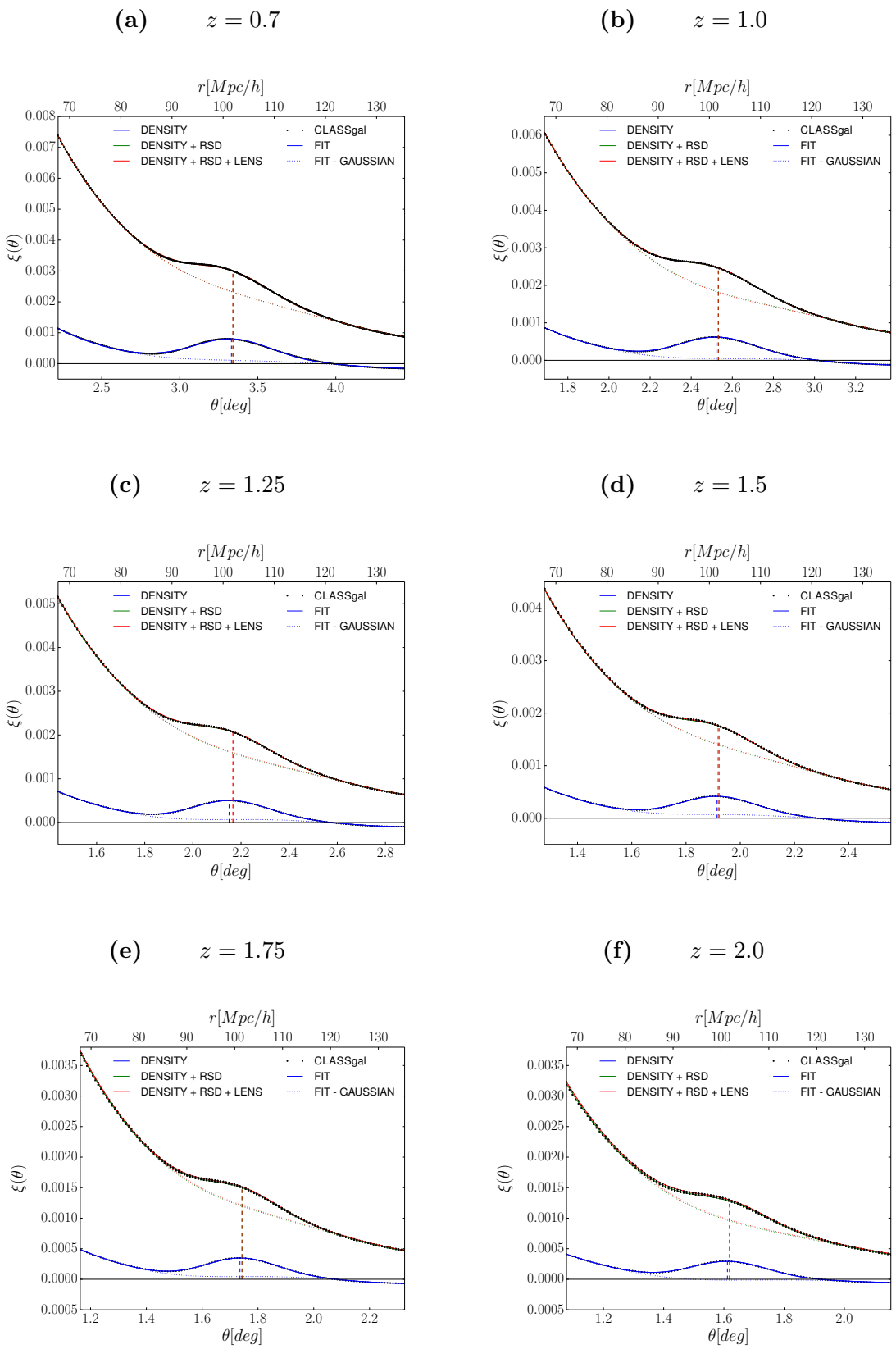


Figure 4.3: Transverse correlation function and BAO peak position computed including different contributions to the observed over density. We used the same colours and line styles as in figure 4.2.

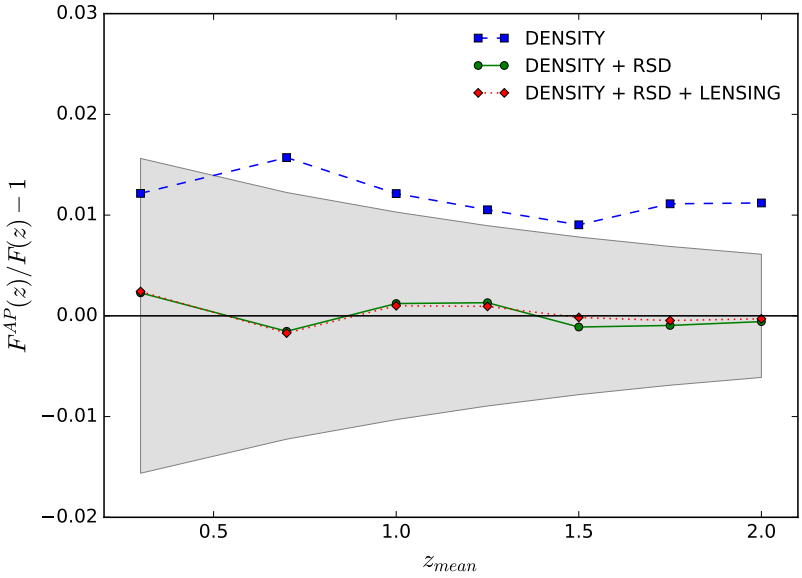


Figure 4.4: AP consistency test for the three cases analysed in this section. The shaded region represents the expected errors around zero. The errors on $F_{\rm AP}$ are estimated from the propagation of uncertainties in redshift and angular resolution. Uncertainties on $\Delta z_{\rm BAO}$ and $\theta_{\rm BAO}$ are assumed to be due to the resolution of the survey: we set $\Delta z = 6.25 \cdot 10^{-4}$; this is the resolution used to compute radial correlations, while for the angular resolution we assume the resolution of a Euclid-like survey, i.e. $\Delta(\theta) = 0.1\,$arcsec. The error on $F_{\rm AP}$ results to be dominated by the uncertainty on Δz.

4.5.2 Galaxy bias

In the previous section we assumed observed galaxies to be unbiased tracers of the underlying dark matter distribution. In this section we relax this assumption in order to investigate the implication of galaxy bias for the BAO measurements and for the AP test. The effect of galaxy bias on BAO measurement was investigated in (Mehta et al. 2011), where it was found that for the most biased tracers ($b > 3$) a non-linear shift on the acoustic scale occurred at the percent level. However, it has been demonstrated that applying BAO reconstruction (Eisenstein et al. 2007) compensates for this effect.

In this section we aim to study how a simple local and linear bias model influences the AP test. In fact, the bias affects the computation of the correlation function, but does not enter in the theoretical value of the AP parameter. Hence, any deviation from the theoretical value, is due to our ability of determining the BAO scale in terms of the truly observable quantities $\xi_{\parallel}(\Delta z, z_{\rm mean})$ and $\xi_{\perp}(\theta, z_{\rm mean})$. Since in the previous section we have found that lensing does not affect the measurement of the AP parameter, we neglect the lensing contribution here. We follow an analogous procedure as in the previous section, but here we vary the bias parameter b, that we assume to be redshift and scale independent (this assumption can also be easily generalized).

In figure 4.5 we show the result of the AP consistency test for different values of the bias parameter. We see that the bias causes an offset which grows with increasing bias. This offset is due to the fact that the BAO position is recovered by using the phenomenological parameterization (4.19), which has been calibrated for unbiased sources. From Figs. 4.2 and 4.3 we note that the shape changes considerably when including redshift space distortions in the correlation function. Therefore, correlation functions for different galaxy bias parameters ($b > 1$) range between the blue and the green lines in Figs. 4.2 and 4.3 and they affect the precision of the parameterization. This indicates, that even when working only with directly observable quantities, we need to assume some cosmological prior to be able to determine accurately the BAO position. Nevertheless, to perform an AP test we do not need to know the physical scale of BAO, and so we can use radial and transverse correlation functions to self-calibrate the parameterizations.

In figure 4.6 we show how the offset varies as a function of the bias parameter, at fixed redshift. When increasing the bias, this offset approaches the offset found for an unbiased tracer, when only the local density term is taken into account. Indeed, for large galaxy bias parameters the redshift space distortions contribution tends to be negligible. As expected there is only a marginal redshift dependence,

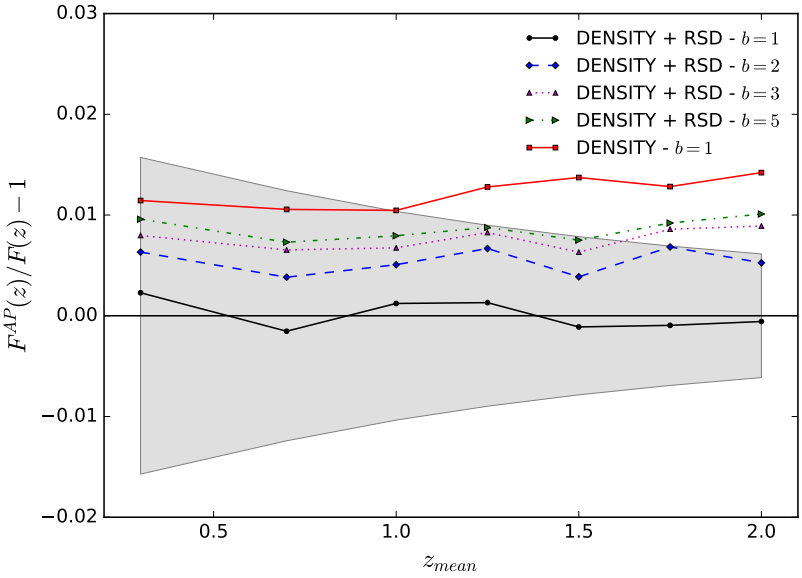


Figure 4.5: AP consistency test for different values of the bias parameter b. The red line is computed including the only contribution from the local density field. The shaded region highlights the typical error-bars around zero.

since we have considered a redshift independent bias and there is only a weak redshift dependence in the relation between the density and the velocity transfer functions, given by the growth factor. In a given survey, it is therefore important to reconstruct the bias as good as possible and to marginalize over the remaining uncertainty in order to obtain reliable results from an AP test.

4.5.3 The radial window function

In the previous section we assumed that the redshifts of the sources are exactly known. In realistic applications, though, a redshift bin has finite thickness that can be modeled by a window function (usually a Gaussian or a top-hat) centered at some mean redshift.

To be able to locate the radial BAO peak position a spectroscopic precision (typically $\sigma_z = 0.001(1+z_{\rm mean})$ or smaller) for the redshift determination is required, whereas photometric redshift resolution is not sufficient. However, in the transverse direction, the presence of a window function does not affect the resolution, but it smears out the BAO feature. As shown in Fig. 4.7, redshift bins of a typical width of photometric survey are sufficient to locate the peak position, but the position depends on the bin-width, see also (Montanari & Durrer 2012).

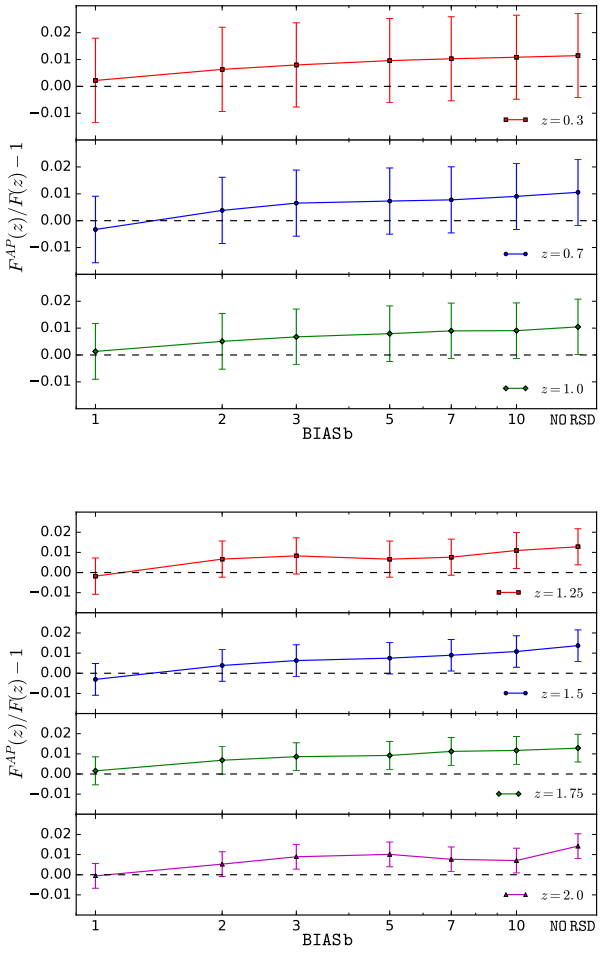


Figure 4.6: AP consistency test, at fixed redshift, for different values of the bias parameter.

Here we are interested in studying the effect of a window function on the transverse correlation function and thereby on the AP test. Since we consider window function with width much larger than the redshift resolution, we assume a top-hat window function in the transverse direction and we refer to σ_z as the half-width of the top-hat. In the radial direction, given the fact that a window function with spectroscopic width ($\sigma_z = 0.001(1 + z_{\rm mean})$ or smaller) does not substantially affect the correlation function, we used a Delta Dirac window.

In figure 4.7 we compare the transverse correlation function computed applying three window functions of different widths. In this example, we assume bias $b = 1$ and, use the analytic expression from (Crocce et al. 2011) for the galaxy

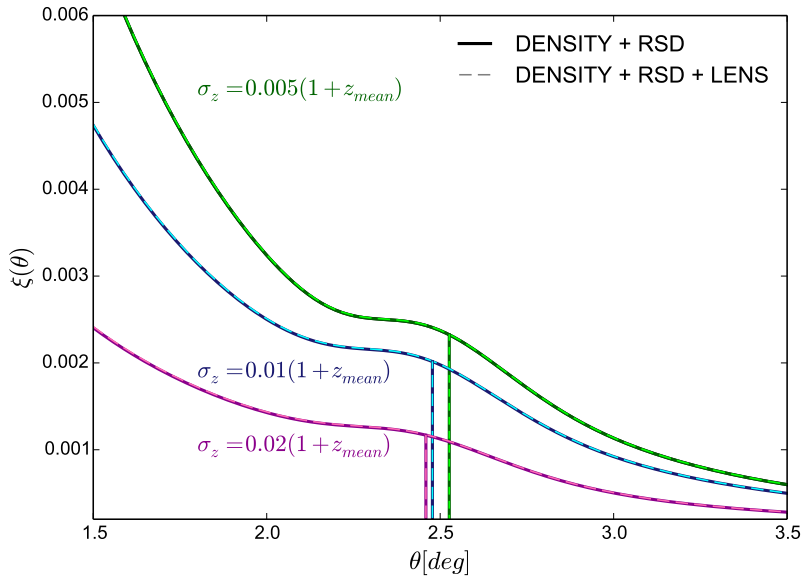


Figure 4.7: Transverse correlation function, at $z = 1$, for different width of the window function. For the thick continuous lines only density and redshift space distortions are included, while in the thin dashed lines also the lensing term is taken into account.

density distribution,

$$\frac{dN}{dz} \propto \left(\frac{z}{0.5}\right)^2 \exp\left[-\left(\frac{z}{0.5}\right)^{1.5}\right]. \tag{4.21}$$

The window function suppresses the correlations and it shifts the peak position toward smaller angular scales, which is in agreement with the result presented in (Montanari & Durrer 2012). Furthermore, we test the modifications induced by gravitational lensing in the transverse direction when a window function is employed. We find that gravitational lensing slightly modifies the amplitude of the correlation function, but this effect is negligible with respect to the suppression induced by the window and leaves the peak position unaffected.

In order to recover the physical BAO scale, the shift induced by the window function can be modeled as longitudinal component in the measured BAO scale. More precisely, the BAO scale can be computed as

$$L = \sqrt{L_\perp^2 + \left(\frac{\delta z}{H(z)}\right)^2}, \tag{4.22}$$

where $L_\perp$ is estimated from the BAO peak in transverse correlation function, while the corrective term δz is a function of the width σ_z of the window function. We

use the parameterization introduced in (Montanari & Durrer 2012), where the corrective term is simply proportional to σ_z

$$\delta z = \sqrt{\gamma}\sigma_z. \tag{4.23}$$

Here we predict the value of the corrective factor γ from the AP test, for three future galaxy redshift surveys: Euclid, the Square Kilometer Array (SKA) and the Dark Energy Spectroscopic Instrument (DESI). As for DESI, we consider the two classes of galaxies that will be targeted by this experiment: bright Emission Line Galaxies (ELGs), that will be observed up to $z = 1.7$, and Luminous Red Galaxies (LRGs), that will be observed up to $z = 1.0$. We introduce also a realistic model for the redshift dependence of galaxy bias.

Even though the three surveys will be able to measure galaxy redshifts with spectroscopic precision, a window function with larger width is used to maximize the signal-to-noise ratio

$$\left(\frac{S}{N}\right)_\theta = \frac{\xi_\perp(\theta)}{\sigma_{\xi_\theta}}, \tag{4.24}$$

where the noise σ_{ξ_θ} is given by the root mean square of the diagonal elements of the covariance matrix $\mathtt{COV}$

$$\mathtt{COV}_{\theta\theta} = \frac{2}{f_{sky}} \sum_{\ell=0}^{\ell_{max}} \frac{2\ell+1}{(4\pi)^2} \Big[P_\ell(\cos\theta)\Big]^2 \left(C_\ell + \frac{1}{n_i}\right)^2, \tag{4.25}$$

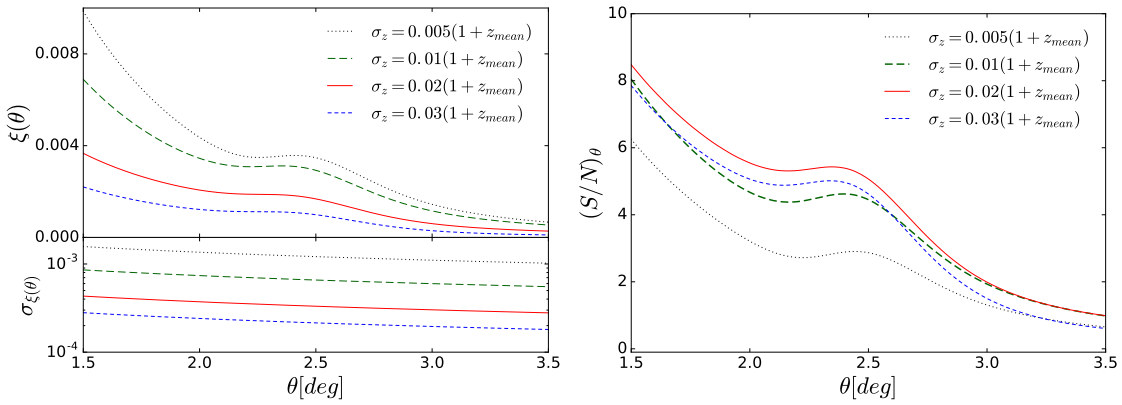
$$\sigma_{\xi_\theta} = (\mathtt{COV}_{\theta\theta})^{1/2}. \tag{4.26}$$

In the expression (4.25), f_{sky} is the fraction of the sky covered by the survey. The shot noise is computed as the inverse of the number of galaxies per steradian n_i inside the i-th redshift bin.

In order to find the width of the window function which maximizes the signal-to-noise ratio, we compute this quantity for different values of σ_z

$$\sigma_z = \mathtt{w} \cdot (1 + z_{\rm mean}), \qquad \mathtt{w} = 0.005, 0.01, 0.02, 0.03, \tag{4.27}$$

at different redshifts in the range $z_{\rm mean} = [0.7{-}2.0]$ and for the different specifications of the survey. In Appendix A.2 we present the survey specifications we use in the computation of the signal-to-noise ratio. In figure 6.36 we show, as an example, the comparison of the signal-to-noise ratio for a Euclid-like galaxy survey at redshift $z = 1$. From figure 4.8a we see that, for a large window function, the signal is suppressed, but the correspondent noise also decreases. We find that the signal-to-noise ratio, around the BAO peak, is optimized for $\sigma_z = 0.02(1+z_{\rm mean})$ (figure 4.8b).



(a) Transverse correlation function and noise. **(b)** Signal-to-noise ratio.

Figure 4.8: In figure a and figure b we show, respectively, the transverse correlation function with its corresponding noise and the signal-to-noise ratio for different widths of the Gaussian window function. Both figures refer to the case of an Euclid-like survey at redshift $z = 1$.

For all the four sets of survey specifications analyzed here, we find the same result. Therefore, from now on, we set our window function to $\sigma_z = 0.02(1 + z_{\rm mean})$.

We perform the AP test, by introducing the window function in the transverse direction and modeling a redshift dependent bias in both radial and transverse directions. For all four considered surveys, we find that redshift binning generates an offset between 3% and 5% (see figure 4.9, dash lines). Minor differences between the four cases are due to the different galaxy density distributions and the redshift dependence of the galaxy biases.

In order to correct the result of the test from the distortions induced by the window function, we model the true AP parameter, as follows from Eqs. (4.22, 4.23) and as already suggested in (Montanari & Durrer 2012)

$$F_{\rm AP}(z_{\rm mean}) = \frac{\Delta z_{\rm BAO}}{\theta_{\rm BAO}} \sqrt{1 - \gamma \cdot \left(\frac{\sigma_z}{\Delta z_{\rm BAO}}\right)^2}\,, \tag{4.28}$$

where γ is the parameter which quantifies the offset due to the window function. We estimate the parameter γ by minimizing the quantity

$$\chi^2 = \sum_i \left(\frac{F_{\rm AP}(z_i) - F_{\rm th}(z_i)}{\Delta F_{\rm AP}(z_i)}\right)^2, \tag{4.29}$$

where $F_{\rm th}$ is the expected value for the AP function and $\Delta F_{\rm AP}$ is the error on $F_{\rm AP}$, determined by the redshift and angular uncertainties, which we assume to be statistically independent at different redshifts. The values of γ which minimizes

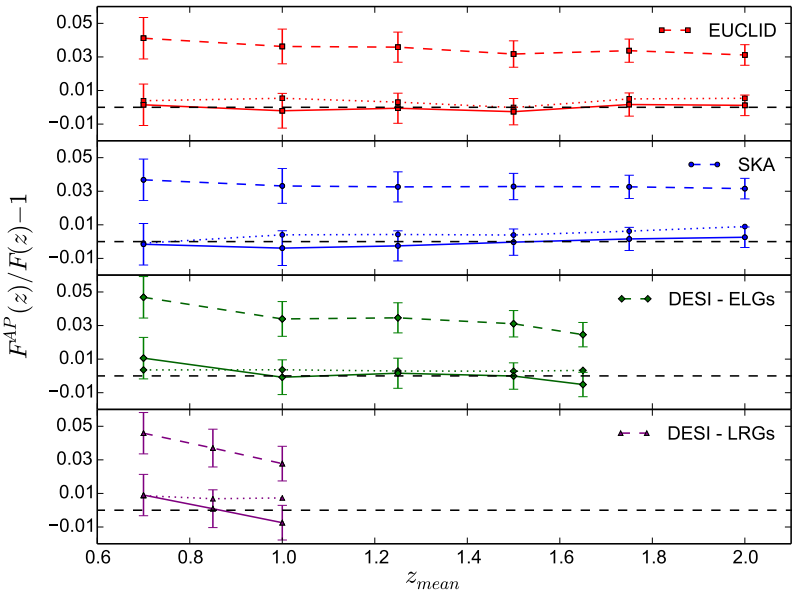


Figure 4.9: The result of the AP test for the four survey specifications considered. The dash lines refer to the cases in which a window function with $\sigma_z = 0.02(1 + z_{\rm mean})$ is used and no correction is applied in order to take into account the effect of the window function. The dotted lines refer to the case in which no window function is employed (the small deviations are due to the redshift dependent bias). The continuous lines represents the cases with window function in the radial direction and which are corrected using the expression (4.28), where γ are the best-fit values, reported in Table 4.1.

	Bias	γ
Euclid-like	$\sqrt{1+z}$	0.166
SKA	$c_4 \exp(c_5 z)$	0.161
DESI - ELGs	$0.84/D(z)$	0.151
DESI - LRGs	$1.7/D(z)$	0.154

Table 4.1: Best-fit values of the parameter γ, for the four different survey specifications. The parameters which specify the bias dependence on redshift for SKA are taken from Table 4 in (Santos et al. 2015), telescope SKA (phase 2). The galaxy bias model for DESI depends on the linear growth factor $D(z)$.

the difference between the measured AP function and its theoretical values are reported in Table 4.1.

In figure 4.9 we show, as solid lines, the results of the test properly corrected for the window function effect. We find that, within the error, the AP function computed from angular and radial positions of the BAO peak in the correlation function agrees with the theoretical value.

4.5.4 Impact of shot-noise on the AP test

In the previous section, we tested the precision of the AP test for different sources of uncertainties. In particular, in section 4.5.2 we found that the galaxy bias introduces an extra 1% error in the estimation of F_{AP}. In this Section we quantify the impact of the shot-noise in the computation of the peak position and on the AP function. We estimate the uncertainty on the cosmological parameters obtained from the AP test including shot-noise and cosmic variance in the error budget. Our estimate is approximate, the true errors are somewhat larger.

We consider an Euclid-like survey, with specifications as described in Appendix A.2. We first compute the errors induced by shot-noise and cosmic variance as the statistical 1-σ errors on our fitting parameters for the correlation functions. We include the error due to the shot-noise and cosmic variance in the fitting procedure by minimizing

$$\chi^2 = \sum_i \left(\frac{\xi(\theta_i) - \hat{\xi}(\theta_i, \mathbf{p})}{\sigma_{\xi_{\theta_i}}} \right)^2, \tag{4.30}$$

where $\hat{\xi}(\theta, \mathbf{p})$ is the parameterization described in Eq. (4.19), $\mathbf{p}$ is the vector of our fitting parameters and $\sigma_{\xi_{\theta_i}}$ are the errors of the correlation function, defined in (4.26). The errors on the parameters $\mathbf{p}$ are estimated by the fit as

$$\sigma_{\mathbf{p}} = \sqrt{\mathrm{diag}[\mathbf{J^T W J}]^{-1}}, \tag{4.31}$$

where $\mathbf{J}$ is the Jacobian of the transformation

$$J^{ij} = \frac{\partial \hat{\xi}(\theta_i, \mathbf{p})}{\partial p_j}, \tag{4.32}$$

and $\mathbf{W}$ is the diagonal matrix

$$W^{ij} = \frac{\delta^{ij}}{\sigma^2_{\xi_{\theta_i}}} \tag{4.33}$$

Using the definition (4.30) we are implicitly assuming the errors $\sigma_{\xi_{\theta_i}}$ to be uncorrelated, which is not true for the angular correlation function. Therefore the statistical errors on the angular position of the BAO will be underestimated.

We use a tophat window function of half-width $\sigma_z = 0.02(1 + z_{\mathrm{mean}})$, which maximizes the signal-to-noise ratio as shown in section 4.5.3.

In Table 4.2 we report the errors on θ_{BAO} computed at different redshifts for two different parameterizations. The polynomial model is the 10 degrees *polynomial+Gaussian* we tested in the previous sections, see Eq. (4.19), which fits

	$\sigma_{\theta_{BAO}}$ [%] Cosmic variance + shot-noise		$\sigma_{\theta_{BAO}}$ [%] Cosmic variance	
z	Polynomial	Power-law	Polynomial	Power-law
0.7	2%	1.3%	2%	1%
1.0	3%	1.4%	1.9%	1%
1.25	3%	1.3%	1.7 %	0.8%
1.5	4%	1.8%	1.7%	0.7%
1.75	6%	3%	1.4%	0.7%
2.0	13%	6%	1.5%	0.7%

Table 4.2: 1-σ statistical error on θ_{BAO} estimated by the fit for two different parameterizations. The polynomial parameterization is the *polynomial+Gaussian* we adopted throughout the chapter, where the polynomial degree is set to be $N = 10$, while the power-law is the 6-parameter *power-law+Gaussian* model considered in Appendix A.1.

14 parameters. The *power-law+ Gaussian* model refers the parameterization tested in Appendix A.1, see Eqs. (A.1, A.2), which is based on 6 parameters. In both cases, we find that the statistical errors dominate the uncertainties due to the angular resolution of the survey. At low redshift ($z < 1.5$), the statistical error is dominated by cosmic variance, while shot-noise dominates at high redshifts ($z \geq 1.5$), where the error rapidly increases up to $\sim 10\%$ (see Table 4.2). Somewhat surprisingly, the errors for the *power-law + Gaussian* model are smaller than the ones estimated for the *polynomial+Gaussian* model, as shown in Table 4.2. This suggests that the large numbers of parameters we introduced to better model the correlation functions are not all necessary, hence we have somewhat 'over fitted' the correlation functions. This has, however, no appreciable effect on the peak positions we have found and which are the only quantities relevant for the AP test.

In order to estimate the impact of these errors on the AP function we compute the total error by summing the systematic errors, due to the resolution of the survey and on the bias, and the statistical errors, due to the shot-noise and cosmic variance,

$$\frac{\Delta F_{\rm AP}}{F_{\rm AP}} = \sqrt{\left(\frac{\Delta F_{\rm AP}}{F_{\rm AP}}\right)^2_{\rm sys} + \left(\frac{\Delta F_{\rm AP}}{F_{\rm AP}}\right)^2_{\rm stat}} \,. \tag{4.34}$$

In figure 4.10 we show the impact of the error due to the shot-noise on the AP function. When only the systematic error is taken into account, the error is dominated by the resolution in redshift in the radial correlation function. The shot-noise contribution is, however, not negligible and becomes even dominant at high redshifts, due to the lower galaxy density. This indicates that with a higher number of galaxies, one could improve the AP test significantly. If both sources of

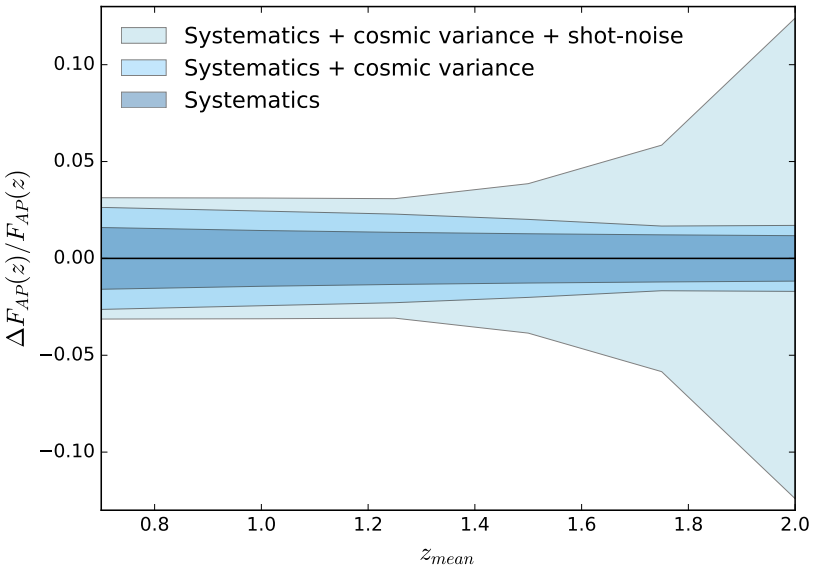


Figure 4.10: Impact of shot-noise on the error of the AP function.

uncertainty are taken into account we find a relative error between 3% and 4% at $z \leq 1.5$, while at higher redshift the error increases up to 12% at $z = 2$.

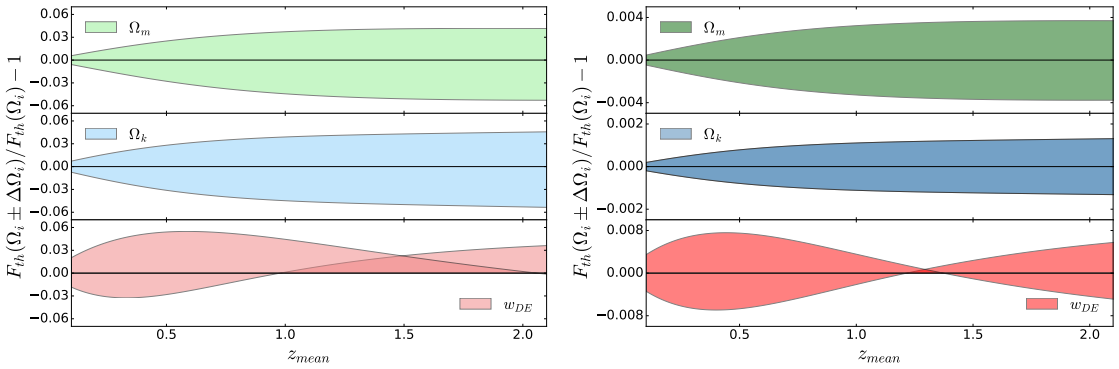
Here we give an approximative estimation of the expected errors on $\Omega_{\rm m}$, $\Omega_{\rm k}$ and w_{DE} from the $F_{\rm AP}$ test. First we study the dependence of the AP function on cosmological parameters. In figure 4.11 we show the sensitivity of the AP function to variations of the cosmological parameters when considering BOSS error-bars (left panel) and the Planck 2015 error-bars (right panel).

For each parameter, we also naively compute its change over the allowed range of variation of the AP function when all the other parameters are kept fixed,

$$\Delta\Omega_i(z) = \left(\frac{\partial F_{\rm AP}}{\partial \Omega_i}\right)^{-1} \Delta F_{\rm AP}. \tag{4.35}$$

The error $\Delta F_{\rm AP}$ is the one obtained above. We compare these errors to the constraints presently available from galaxy clustering data, more precisely, from the Baryon Oscillation Spectroscopic Survey (BOSS) (Chuang et al. 2013), and from the state-of-the-art constraints, given by Planck, *CMB + External Data* analysis (Ade et al. 2016a) in figure 4.12, light shaded regions.

The equation of state of dark energy, for this simple parameterization, is poorly constrained in the AP test around $z \sim 1.3$. This reflects the fact that the AP parameter is not sensitive to w around $z \sim 1.3$, as we see in figure 4.11. The light



(a) BOSS **(b)** Planck 2015

Figure 4.11: Sensitivity of the AP function to different cosmological parameters. $\Delta\Omega_i$ is taken to be the error on the parameter Ω_i from galaxy clustering only (top panel) and from CMB analysis (bottom panel). The constraints from galaxy clustering come from the Baryon Oscillation Spectroscopic Survey (BOSS) of SDSS-III, when only the CMASS-Large dataset is employed (Chuang et al. 2013). The constraints from CMB are the 1-σ limits ($\Omega_{\rm m}$) and 2-σ limits ($\Omega_{\rm k}$ and w_{DE}) in Planck 2015 analysis (Ade et al. 2016a), when *CMB plus external data* are used. The sign change in $\frac{\partial F_{\rm AP}}{\partial w_{DE}}$ at $z \simeq 1.3$ is well visible in the precise Planck data 'smeared out' in the less accurate BOSS data.

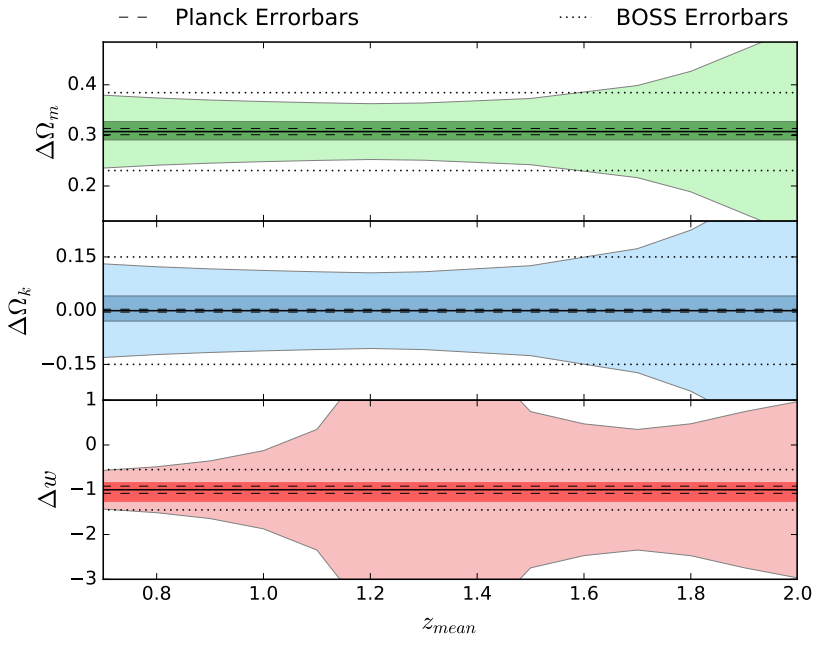


Figure 4.12: Constraints on cosmological parameters from a single measurement of the AP function (light shaded region) and from $N = 14$ measurements (assumed to be independent) (dark shaded region). The Planck and BOSS limits are indicated by dashes and dotted horizontal lines.

shaded regions in figure 4.12 represent the error on a single measurement of $F_{\rm AP}$, the overall statistical errors on the parameters can be reduced by performing N independent measurements at different redshifts.

We also estimate the errors from $N = 14$ independent measurements in the range $[0.7, 2.0]$ with redshift resolution $\Delta z = 0.1$. The errors are estimated by fitting a set of hypothetical data points for the AP function F_{DATA}. The data points are computed by summing the value of F_{AP} for our fiducial cosmology to a randomly generated scattering term proportional to the error. The fit is performed by running `emcee` (Foreman-Mackey et al. 2013), a Markov Chain Monte Carlo (MCMC) code. The χ^2 inferred from the simulated data is

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{F_{DATA}(z_i) - F_{AP}(z, \Omega)}{\Delta F_{AP}} \right)^2, \tag{4.36}$$

where we vary only one parameter Ω at a time, keeping the others fixed to their fiducial value.

We obtain the following results for Ω_m, Ω_k and w_0:

$$\Omega_m = 0.309^{+0.019}_{-0.018}, \qquad \Omega_k = 0.005^{+0.036}_{-0.034}, \qquad w_0 = -1.02^{+0.18}_{-0.24}, \tag{4.37}$$

where the quoted errors are the $1-$sigma limits.

The dark shaded regions in Fig. 4.12 shows the error regions given in (4.37). The error-bars from the AP test are still significantly larger than the Planck error bars but smaller that present errors obtained from BOSS. This is not surprising since the AP test uses just one single scale in the power spectrum, while parameter estimation from the full CMB power spectrum, has the advantage of using all available modes.

4.6 Summary

In this chapter we have presented a model independent method to perform the Alcock Paczyński test. We applied the test to the BAO feature in the galaxy 2-point correlation function, in the radial and transverse directions. In order to perform the test without prior assumption on the cosmological parameters, we model the shape of the correlation function with a polynomial, while the acoustic feature is modeled as a Gaussian. We have shown that both redshift space distortions and galaxy bias must be taken into account to improve the accuracy of the estimated BAO scale, while gravitational lensing does not significantly affect the result of the AP test. Finally, we have investigated the projection effect induced by the finite width of the radial window function. In fact, in the transverse direction a

redshift bin of a typical width of a photometric survey is sufficient to locate the peak position and it can be employed to maximize the signal-to-noise. However, the projection effect shifts the measured acoustic peak toward smaller angular scale. If not corrected for, this introduces a systematic error of up to 5% in the AP test. We have computed the correction that has to be applied in order to adjust the result of the test for this effect for three planned galaxy surveys, Euclid-like, SKA and DESI. We have also estimated the effect of shot-noise and cosmic variance in the precision of the BAO peak detection and on the AP test. We have found that the precision of the test could be improved significantly by increasing the number of galaxies in the survey, especially at high redshift. Finally, we have shown that cosmological parameters can be estimated in a model independent approach, by the AP test, with roughly the same accuracy as the BAO analysis performed with the BOSS survey. The latter, however, has the disadvantage to use the power spectrum in Fourier space, $P(k)$, which itself depends in a non-trivial way on cosmological parameters via the conversion of angles and redshifts into distances.

Chapter 5

LENSING CONVERGENCE IN GALAXY CLUSTERING IN ΛCDM AND BEYOND

5.1 Introduction

The predictions of General Relativity (GR) have been successfully tested with great precision in the near Universe, in the weak-field regime of the solar system (Everitt et al. 2011, 2015) and in the strong-field regime of binary pulsar systems (Taylor & Weisberg 1989; Will 2014), black hole (Abbott et al. 2016a,b, 2017c) and neutron star (Abbott et al. 2017d) mergers. Cosmological observations allow GR to be investigated on entirely new length and time scales. Testing GR on cosmological scales is further motivated by the lack of a convincing physical explanation for the accelerated expansion and by the purpose of comparing data with Dark Energy or modified gravity models. Moreover, several modified gravity models give rise naturally to some screening mechanisms, which due to non-linearity freeze the additional degree(s) of freedom on small scales recovering the GR predictions. This provides a further motivation to test GR on cosmological scales, i.e. well beyond the screening radius, where it is possible to discriminate between GR and different modified gravity models.

The success of precision cosmology depends not only on precise observations, but also on the theoretical modelling, which must be understood to at least to the same level of accuracy to avoid theoretical systematics that may bias the measurements at a level that cannot be neglected nowadays.

The observed galaxy number count receives contributions from the density perturbation, from the peculiar velocity of the sources, the well-known redshift space distortions and the Doppler term (Kaiser 1987), and from the fact that the photons travel on perturbed geodesics, leading to perturbations in the observed

redshift and in the spacetime volume. These additional contributions, beyond density perturbations and redshift space distortions, are the so-called relativistic effects and they are due to gravitational lensing, Doppler, Shapiro time-delay and integrated Sachs-Wolfe effects on the path of photons. See (Yoo et al. 2009; Yoo 2010; Bonvin & Durrer 2011; Challinor & Lewis 2011; Schmidt & Jeong 2012; Bertacca et al. 2012; Raccanelli et al. 2016b, 2014) for the fully relativistic computation at linear order in the perturbations. In this chapter we focus on the lensing magnification effect, which enters at the same parametrical order (in terms of an expansion in $\mathcal{H}/k$) of the standard Newtonian contributions: density and redshift space distortion. It accounts for two competing modifications to the galaxy number counts due to gravitational lensing. At one side the number of observed galaxies is decreased because of the stretching of the solid angle of observation. At the other side, it can be increased (or decreased) since the observed luminosity is affected by matter perturbations in a way that the signal of a galaxy behind overdense regions is magnified (or de-magnified for galaxies behind underdense regions) and therefore can be detected, even if the intrinsic luminosity is below the threshold of the survey. The last effect, which takes into account that a real galaxy survey is luminosity-limited, is called magnification bias and depends on the specifications of the survey. The lensing magnification effect has been predicted more than two decades ago (Bartelmann 1995; Dolag & Bartelmann 1997; Sanz et al. 1997) and first detected in 2005 (Scranton et al. 2005). Whereas other relativistic effects has been shown to be detectable by cross-correlating different probes (McDonald 2009; Yoo et al. 2012; Bonvin et al. 2014; Alonso & Ferreira 2015; Fonseca et al. 2015; Iršič et al. 2016; Bonvin et al. 2016; Gaztanaga et al. 2017; Dai et al. 2016; Borzyszkowski et al. 2017; Abramo & Bertacca 2017) (see however (Raccanelli et al. 2016a)), their impact on a single tracer analysis can be safely neglected, apart for the measurement of primordial non-Gaussianity (see (Camera et al. 2015b,a; Raccanelli et al. 2016c; Alonso et al. 2015)) that we do not address.

The study of the impact of gravitational lensing on cosmological observables, in particular the distance-redshift relation, has a long history starting more than thirty years ago, see e.g (Dyer & Roeder 1972, 1973, 1974; Sasaki 1987; Kantowski 2003) and successively the effect of cosmological perturbations was addressed by several groups with different methods, see e.g. (Bonvin et al. 2006; Barausse et al. 2005; Gasperini et al. 2011; Ben-Dayan et al. 2012; Umeh et al. 2014a,b; Ben-Dayan et al. 2013b; Fanizza et al. 2013; Ben-Dayan et al. 2013a) and references therein. This issue has recently received attention also in the context of CMB cosmology, that could be altered from a biased estimation of the distance to last

scattering surface. In (Kaiser & Peacock 2016; Bonvin et al. 2015) it is clarified the role of lensing in the CMB analysis.

While lensing magnification effect has been known by decades, ongoing galaxy clustering surveys are not including its contribution in the data analysis process. In view of the next generation of surveys which will reach higher redshifts, we are interested in questioning the validity of this approximation and to estimate possible biases.

This issue has been considered already in (Namikawa et al. 2011; Camera et al. 2015b,a; Raccanelli et al. 2016c; Montanari & Durrer 2015; Di Dio et al. 2016b; Cardona et al. 2016; Lepori et al. 2017). The main result is that including the lensing contribution to the galaxy clustering analysis is essential in order to avoid biased estimations of the best-fit values and constraints, especially for some cosmological parameters. In this chapter we focus on the impact on parameters beyond ΛCDM: massive neutrinos and modified gravity. To test GR, and consequently to constrain the cosmological parameters which describe the deviation from GR, observations need to measure both the Bardeen potentials. While the standard Newtonian contribution to galaxy clustering is sensitive only the gravitational potential described by the time component of the metric, which determines the motion of non-relativistic object as galaxies, the lensing magnification measures the Weyl potential, hence the sum of the two Bardeen potentials. Therefore, it is clear that the lensing magnification carries the useful information to test for deviation of GR and it is worth to investigate if neglecting its contribution to the galaxy number counts may bias the results. This issue was addressed in (Alonso et al. 2017), where the authors study the impact of lensing and the other relativistic effects on the constraints for the case of a sub-class of Horndeski scalar-tensor theories, also in combination with other probes. For a study of relativistic effects within Horndeski gravity see (Renk et al. 2016).

We will consider the parametrized description where modifications of gravity are encoded in two parameters that can be in general functions of time and scale: they modify the Poisson equation and the relation between the two Bardeen potentials prescribed by GR. We will study the impact of neglecting lensing magnification in the constraints and also in the best-fit value of the cosmological and modified gravity parameters inferred from galaxy clustering measurements. We will also analyse the dependence of the lensing contribution on the type of survey and on the redshift binning configuration. Other works (Raccanelli et al. 2016c; Montanari & Durrer 2015; Di Dio et al. 2016b; Cardona et al. 2016) have described the lensing magnification impact in terms of few very broad redshift bins, which can adapt

well to photometric surveys. We will extend this analysis to modified gravity. Moreover we will study how the lensing impact changes by increasing the number of redshift bins. While we can not reach the same redshift accuracy of a spectroscopic survey, we intend to find a hint of the real effect on spectroscopic surveys. In this analysis we choose to use the photometric specifications for Euclid[1] and the spectroscopic specifications for SKA[2].

The chapter is organized as follows. We introduce the galaxy number counts and describe the parametrization we use for modified gravity in section 5.2 and describe the methodology and the survey specifications that we use in section 5.3. We then present the results in section 5.4 and we summarise the main findings of this chapter in section 5.5.

5.2 Galaxy Number Counts

In a galaxy clustering experiment cosmologists measure the number of galaxies or sources $N(\mathbf{n}, z)$ in terms of the observed direction $\mathbf{n}$ and measured redshift z. We can therefore naturally define the galaxy number count observable as

$$\Delta_{\rm gal}(\mathbf{n}, z) = \frac{N(\mathbf{n}, z) - \langle N \rangle(z)}{\langle N \rangle(z)} \tag{5.1}$$

where $\langle .. \rangle$ denotes the angular average at fixed measured redshift z. The galaxy number counts at first order in perturbation theory reads

$$\begin{aligned}
\Delta_{\rm gal}(\mathbf{n}, z) \;=\; & b_{\rm gal} D + \frac{1}{\mathcal{H}(z)} \partial_r (\mathbf{V} \cdot \mathbf{n}) \\
& + (5s - 2) \int_0^{r(z)} \frac{r(z) - r}{2 r(z) r} \Delta_\Omega (\Phi + \Psi) dr \\
& + \left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2} + \frac{2 - 5s}{r\mathcal{H}} + 5s - f^{\rm gal}_{\rm evo} \right) (\mathbf{V} \cdot \mathbf{n}) \\
& + (f^{\rm gal}_{\rm evo} - 3)\mathcal{H} V + (5s - 2)\Phi + \Psi + \frac{1}{\mathcal{H}} \dot{\Phi} + \frac{2 - 5s}{r(z)} \int_0^{r(z)} dr (\Phi + \Psi) \\
& + \left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2} + \frac{2 - 5s}{r(z)\mathcal{H}} + 5s - f^{\rm gal}_{\rm evo} \right) \left(\Psi + \int_0^{r(z)} dr (\dot{\Phi} + \dot{\Psi}) \right),
\end{aligned} \tag{5.2}$$

where D is the dark matter density fluctuation in synchronous gauge, $\mathbf{V}$ the peculiar velocity in longitudinal gauge, V the potential velocity related through $\mathbf{V} = -\nabla V$ to the peculiar velocity and Φ and Ψ are the gauge-invariant Bardeen potentials.

[1] http://sci.esa.int/euclid/
[2] http://skatelescope.org/

The Eq. (5.2) introduces three bias parameters to relate dark matter perturbations to galaxies, namely a galaxy bias $b_{\rm gal}$, a magnification bias defined as

$$s(z, F_*) = -\frac{2}{5}\frac{\partial \ln \bar{n}(z, F > F_*)}{\partial \ln F_*} \tag{5.3}$$

where $\bar{n}$ is the background cumulative number density of galaxies with flux above the threshold F_*, namely

$$\bar{n}(z, F > F_*) = \int_{\ln F_*}^{\infty} \bar{n}(z, \ln F)\, d\ln F\,, \tag{5.4}$$

and an evolution bias

$$f_{\rm evo}^{\rm gal}(z, F > F_*) = \frac{\partial \ln \bar{n}\,(z, F > F_*)}{\mathcal{H}\partial\eta}. \tag{5.5}$$

Being the galaxy number counts a function of a unit vector $\mathbf{n}$ and a redshift z, it is natural to expand them in terms of spherical harmonics for different redshifts

$$\Delta_{\rm gal}(\mathbf{n}, z) = \sum_{\ell m} a_{\ell m}(z)\, Y_{\ell m}(\mathbf{n}) \tag{5.6}$$

with

$$a_{\ell m}(z) = \int d\Omega_{\mathbf{n}} Y^*_{\ell m}(\mathbf{n})\, \Delta_{\rm gal}(\mathbf{n}, z) \tag{5.7}$$

This leads to the redshift-dependent galaxy power spectra

$$\langle a_{\ell m}(z)\, a^*_{\ell' m'}(z')\rangle = C_\ell(z, z')\, \delta_{\ell\ell'}\delta_{mm'} \tag{5.8}$$

where the Kronecker symbols are a consequence of angular isotropy. It is worth remarking that, despite working with angular statistics, the power spectra $C_\ell(z, z')$ contain the full 3-dimensional information, which can be fully recovered by optimally cross-correlating spectra at different redshifts, see e.g. Refs. (Asorey et al. 2012; Di Dio et al. 2014b). In order to cross-correlate different redshifts we need to introduce a redshift binning, hence we defined the binned spectra as

$$C_\ell^{ij} = \int dz_1' dz_2' \frac{dN}{dz_1'}\frac{dN}{dz_2'} W_i(z_1'; z_1, \sigma_1)\, W_j(z_2'; z_2, \sigma_2)\, C_\ell(z_1', z_2') \tag{5.9}$$

where dN/dz is the survey selection function and W_i is the window function for the $i-$th redshift bin centred at z_i with a width of σ_i normalised to the unity. In our work we will consider two different shapes of window functions: tophat and gaussian bins, see section 5.3.3.

In our analysis we consider the first three terms of the galaxy number counts in Eq. (5.2): beside the contribution from the density, we include only redshift-space

distortions and lensing which physically represent the leading perturbations of the space-time volume, in the radial and transverse direction, respectively. The lensing magnification effect enters at the same parametrical order (in terms of an expansion in $\mathcal{H}/k$) of the standard Newtonian contributions - density and redshift-space distortions - and is expected to dominate over the other relativistic effects for typical current and upcoming galaxy surveys. It is well known that it dominates radial cross-correlations between different redshift bins, while the contributions due to the Newtonian terms drops rapidly. The lensing effect can be described easily in terms of angular power spectra, as done for CMB or shear weak lensing, but including lensing in a 3-dimensional Fourier analysis is a complicate and severe problem (Mandel & Zaldarriaga 2006). Indeed lensing gets contributions along the whole line of sight and therefore it mixes non-trivially different Fourier scales. As we see from Eq. (5.2), lensing magnification is the angular gradient of the lensing potential to the source position. Hence the lensing magnification is a pure transverse effect, see Ref. (Di Dio 2017) for the impact of lensing on the transversal modes of the matter power spectrum, and by adding more redshift correlations we do not measure more physical modes induced by lensing. We therefore expect that by increasing the number of redshift bins, or having a better redshift resolution, will not increase the number of modes induced by lensing but only the ones dominated by density and redshift space distortions. This simple picture neglect the information in the redshift evolution, but to first approximation we do expect that the relative impact of lensing magnification decreases by measuring more radial modes, apart for the cosmological parameters which require the lensing information. In the next sessions we perform a Fisher analysis to provide a quantitative answer.

We compute the angular power spectra by using a modified version[3] of CLASS (Lesgourgues 2011; Di Dio et al. 2013; Baker & Bull 2015) in order to include modifications of GR. We consider first-order scalar metric perturbations about a flat Friedmann-Lemaître-Robertson-Walker (FLRW) background. The line element in the Newtonian gauge reads

$$\mathrm{d}s^2 = a^2(\eta)\left[-(1+2\Psi)\,\mathrm{d}\eta^2 + (1-2\Phi)\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j\right], \tag{5.10}$$

where η is the conformal time, a is the scale factor, and Ψ and Φ are the Bardeen potentials. In GR the linearized Einstein equations set the two Bardeen potentials Ψ and Φ to be equal (in absence of anisotropic stress) and relate them to the matter density in comoving gauge through the Poisson equation. Modified theories of

[3] https://gitlab.com/philbull/mgclass

gravity can change these relations. It is therefore common to parametrize deviations from GR by the following relations (Zhao et al. 2009):

$$\frac{\Phi}{\Psi} = \gamma \tag{5.11}$$

$$k^2\Psi = -\frac{3}{2}\frac{\mathcal{H}_0^2\Omega_{m_0}}{a}\mu D\,, \tag{5.12}$$

where Ω_{m_0} and $\mathcal{H}_0$ are the density and Hubble parameter today, respectively. The "clustering parameter" μ parametrizes modification of the Poisson equation relating the gauge invariant density perturbation D to the potential Ψ and the "slip parameter" γ parametrizes anisotropic stress which makes the ratio between the two scalar potentials different from unity. In this chapter we focus on the effects of lensing magnification which depends the Weyl potential $\Phi + \Psi$. Thus we will consider the equivalent parametrization in terms of $\{\mu, \Sigma\}$ instead of the one in $\{\gamma, \mu\}$, where Σ is defined by

$$\Sigma = -\frac{k^2\,(\Phi+\Psi)}{3\mathcal{H}_0^2\Omega_{m_0}D} = \frac{\mu\,(1+\gamma)}{2}\,. \tag{5.13}$$

In GR $\mu = \gamma = \Sigma = 1$ whereas in modified gravity theories, in full generality, they can depend on time and wave number. Throughout this chapter we consider the most simple case, where μ and Σ assume constant values μ_0 and Σ_0 at late times, starting from $\mu = \Sigma = 1$ at early times. We fix the threshold redshift at which gravity starts to be modified from GR at $z = 5$.

5.3 Methodology

In order to estimate the bias on the constraints and on the best fit of the cosmological parameters we use the Fisher matrix formalism, see (Fisher 1935; Tegmark et al. 1998). The Fisher matrix for the galaxy angular power spectra is given by

$$F_{\alpha\beta} = \sum_{\ell=\ell_{\rm min}}^{\ell_{\rm max}} \sum_{(ij)(pq)} \frac{\partial C^{ij}_{\ell,\rm th}}{\partial\Phi_\alpha}\frac{\partial C^{pq}_{\ell,\rm th}}{\partial\Phi_\beta}\sigma^{-2}_{C_\ell\,[(ij),(pq)]}\,, \tag{5.14}$$

where Φ_α denotes the α-th parameter, the second sum over runs over the indices (ij) and (pq) with $i \leq j$ and with $p \leq q$ which range from 1 to the total number of redshift bins $N_{\rm bin}$, and all bin auto- and cross-correlations are included, unless stated differently. For Gaussian fluctuations the error in the measured power spectra is given by, (Di Dio et al. 2014b)

$$\sigma^2_{C_\ell\,[(ij),(pq)]} = \frac{C^{(ip)}_{\ell,\rm obs}C^{(jq)}_{\ell,\rm obs} + C^{(iq)}_{\ell,\rm obs}C^{(jp)}_{\ell,\rm obs}}{(2\ell+1)f_{\rm sky}}\,. \tag{5.15}$$

The observed correlation multipoles $C^{ij}_{\ell,\rm obs}$ including shot-noise are

$$C^{ij}_{\ell,\rm obs} = C^{\rm Nwt}_\ell + C^{\rm Lens}_\ell + \frac{\delta^{ij}}{N_i} \tag{5.16}$$

where $N_i = dN_{z_i}/d\Omega$ denotes the number of sources per steradian in the i-th bin. We remark that we always include the lensing convergence contribution in the observed power spectra in the covariance. We choose $\ell_{\rm min} = 2$ and $\ell_{\rm max} = 300$, and we comment on the dependence on $\ell_{\rm max}$ in Section 5.4.2. It is worth remarking that more sophisticate redshift dependent non-linear cutoffs can be chosen, see e.g. (Di Dio et al. 2014b; Alonso et al. 2017). Nevertheless our choice, namely to use the same $\ell_{\rm max}$ for any redshift-bin is conservative. Indeed this includes a larger amount of small transversal scales at low redshift compare to high redshift. Therefore it includes more modes in the regime where the lensing magnification is less relevant.

Regarding the theoretical power spectra we will instead assume two different models. As in (Alonso et al. 2015; Di Dio et al. 2016b) we introduce an additional parameter ϵ_L and write the power spectra as

$$\begin{aligned} C_{\ell,\rm th}(\theta_\alpha, \epsilon_L) &= C^{\rm g}_\ell(\theta_\alpha) + C^{\rm RSD}_\ell(\theta_\alpha) + \epsilon_L C^{\rm Lens}_\ell(\theta_\alpha) \\ &= C^{\rm Nwt}_\ell(\theta_\alpha) + \epsilon_L C^{\rm Lens}_\ell(\theta_\alpha) \end{aligned} \tag{5.17}$$

where the first two terms represent galaxy clustering and redshift-space distortions and we have explicitly shown the dependence on the cosmological parameters $\{\theta_\alpha\}$ and on the lensing parameter ϵ_L. When the galaxy number counts are modelled by the Newtonian terms only, i.e. $\epsilon_L = 0$, we just consider the first two terms in Eq. (5.17). When instead we include the lensing contribution $\epsilon_L = 1$ and the power spectra are given by

$$C_{\ell,\rm th}(\theta_\alpha) = C^{\rm Nwt}_\ell(\theta_\alpha) + C^{\rm Lens}_\ell(\theta_\alpha)\,. \tag{5.18}$$

For convenience we split the parameter set as $\{\Phi\} = \{\theta_\alpha, \epsilon_L\}$. The Fisher matrix F for the full set of parameters $\{\theta_\alpha, \epsilon_L\}$ has then the form

$$F^{\Phi\Phi} = \begin{pmatrix} F^{\theta\theta} & F^{\theta\epsilon_L} \\ F^{\theta\epsilon_L} & F^{\epsilon_L\epsilon_L} \end{pmatrix}. \tag{5.19}$$

The derivatives with respect to the cosmological parameters $\{\theta_\alpha\}$ are approximated using a five-point stencil, where for the step size $\delta\theta_\alpha$ we take the 10% of the fiducial value and we have verified that our results do not depend on the particular choice for $\delta\theta_\alpha$. The derivatives with respect to the lensing parameter ϵ_L are calculated analytically and are simply given by $C^{\rm Lens}_\ell$.

We assume a fiducial flat ΛCDM consistent with Planck, (Ade et al. 2016a), including massive neutrinos. The fiducial values of the cosmological parameters are as follows. The amplitude of the curvature fluctuations is $\ln 10^{10} A_s = 3.094$ and the spectral index is $n_s = 0.9645$. The matter content of the Universe is parametrized by $h^2\Omega_b = \omega_b = 2.225 \times 10^{-2}$, $h^2\Omega_m = \omega_m = 0.1198$, $M_\nu = 0.06\,\mathrm{eV}$ for the baryon and cold dark matter density parameters and for the sum of the neutrino masses, respectively. The dimension-less Hubble parameter is set to be $h = 0.6727$. For the modified gravity parameters the fiducial values are $\mu = \Sigma = 1$. We also consider the galaxy bias at redshift $z = 0$, b_0, as an additional parameter with fiducial values consistent with the specifications for Euclid and SKA that we assume, see eqs. (5.28) and (5.33). We will marginalize all our results with respect to the galaxy bias parameter. We do not impose any prior from other experiments nor combine galaxy clustering with other probes: our purpose is to focus on a galaxy clustering analysis only and to study the relative impact of the lensing contribution in the determination of the cosmological parameters and in the constraining power of the surveys under consideration. Shear weak lensing surveys and lensing magnification will provide the same information about the lensing potential, but the two measurements suffer by different systematic effects: on one hand we have intrinsic alignment and on there other one galaxy and magnification biases.

5.3.1 Bias on the parameter constraints

Our aim is to quantify the amount of information that is misinterpreted if the lensing contribution is neglected in the galaxy clustering analysis. We have therefore introduced in Eq. (5.17) the lensing parameter ϵ_L but for our purpose here we do not treat it as the other cosmological parameters. This is because in our approach ϵ_L is not free to vary over any value and it is not supposed to be measured by the experiments. Rather we treat the lensing contribution in the galaxy power spectra as a theoretical systematic meaning that the parameter ϵ_L is fixed to 1 or to 0, depending on which is the model we assume with or without the lensing term respectively. We consider the sub-matrix $F^{\theta\theta}$ in Eq. (5.19) for the cosmological parameters only

$$F^{\theta\theta}_{\alpha\beta} = \sum_{\ell=2}^{\ell_{\max}} \sum_{(ij)(pq)} \frac{\partial C^{ij}_{\ell,\mathrm{th}}}{\partial\theta_\alpha} \frac{\partial C^{pq}_{\ell,\mathrm{th}}}{\partial\theta_\beta} \sigma^{-2}_{C_\ell\,[(ij),(pq)]}\,, \tag{5.20}$$

and estimate the standard marginalized errors for the cosmological parameters from

$$\sigma_{\theta_\alpha} = \sqrt{\left[(F^{\theta\theta})^{-1}\right]_{\alpha\alpha}} \tag{5.21}$$

We then compare the marginalized errors that we get from (5.21) with the spectra given by $C_{\ell,\mathrm{th}}(\theta_\alpha) = C_\ell^{\mathrm{Nwt}}(\theta_\alpha) + C_\ell^{\mathrm{Lens}}(\theta_\alpha)$ with the ones that we get with the Newtonian spectra only. In both the analyses we include the lensing contribution in the observed power spectra in the covariance. The change in the marginalized errors is due to the actual dependence of the lensing term on the cosmological parameters.

The Fisher matrix (5.20) has dimension $N \times (N+1)/2$, where N is the number of cosmological parameters under consideration: we have 5 parameters for ΛCDM, 6 parameters when we include massive neutrinos and 8 parameters including also modified gravity, parametrized by μ and Σ. The errors in Eq. (5.21) on each parameter are calculated marginalizing over all the others.

5.3.2 Bias on the parameter estimation

Our approach for the calculation of the bias on the estimation of the cosmological parameters is based on the Fisher matrix method applied to the case of nested models as introduced in (Taylor et al. 2007; Knox et al. 1998; Heavens et al. 2007; Kitching et al. 2009). Suppose to have two models such that the parameter space of one model is contained in the one of the other model, in a way that both the two models have some parameters in common, e.g. the cosmological parameters $\{\theta_\alpha\}$, that are supposed to be measured from some experiment. The difference between the two models is that in the "wrong" model some other parameters are fixed to some known values whereas in the "correct" model they are free to vary as the other parameters. This framework can be adapted to the situation where the additional parameters represent the amplitude of some systematic effects, i.e. they are set to zero in the wrong model, where systematics are not taken into account or they are set to 1, which corresponds to the correct model, where the effect of systematics is included. This is the case we are studying here: under the general assumption that the Universe is described by the power spectra including the lensing contribution, in the correct model we set $\epsilon_L = 1$ whereas in the wrong Newtonian model we have $\epsilon_L = 0$.

The wrong model assumption leads to a shift in the best-fit values of the cosmological parameters $\{\theta_\alpha\}$ which follows simply from the fact that the dependence of the likelihood on the additional parameters in the correct model displaces the value of the maximum, (Taylor et al. 2007). The shift in the best-fit value of a parameter is given by

$$\Delta\theta_\alpha = \sum_\beta \left(F^{\theta\theta}\right)^{-1}_{\alpha\beta} F_\beta^{\theta\epsilon_L}\,, \tag{5.22}$$

where both the terms are subset of the Fisher matrix $F^{\Phi\Phi}$ of the full set of the parameters of Eq. (5.19): $\left(F^{\theta\theta}\right)^{-1}_{\alpha\beta}$ is the inverse of the sub-matrix of the $\{\theta_\alpha\}$ parameters and $F^{\theta\epsilon_L}$ is the column corresponding to the $\{\theta_\alpha\}$ and ϵ_L parameters[4]. The estimation of Eq. (5.22) is based on a first-order Taylor expansion of the likelihood of the correct model and relies on two crucial assumptions. Firstly the difference between the two models has to be small, and secondly the approximated expression in (5.22) holds locally, i.e. for small shifts $\Delta\theta_\alpha$: for large shifts the Taylor approximation is not accurate meaning that the results can only be trusted qualitatively. Keeping in mind this caveat, it is worth to discuss two more subtle issues related to the terms appearing in the Fisher matrix in Eq. (5.22). The first concerns which spectra have to be used for the derivatives with respect to the cosmological parameters $\{\theta_\alpha\}$, where we have two possibilities: the spectra with lensing $C_{\ell,\mathrm{th}}(\theta_\alpha) = C_\ell^{\mathrm{Nwt}}(\theta_\alpha) + C_\ell^{\mathrm{Lens}}(\theta_\alpha)$ - fiducial model $\epsilon_L = 1$ - or just the Newtonian ones - fiducial model $\epsilon_L = 0$, where in practice the second choice neglects the fact that the lensing term depends on the cosmological parameters. If we use the spectra with lensing we are calculating the bias in the parameter estimation due to neglecting a systematic effect that depends itself on such parameters. This is indeed the case for lensing and actually we have exploited this dependence to estimate the bias in the constraints, see the discussion of our method in section 5.3.1. However this approach would of course lead to larger shifts $\Delta\theta_\alpha$: we have calculated that in this case the shifts normalized to their fiducial values are much greater than one and roughly one order of magnitude larger than the results using Newtonian spectra for almost all the cosmological parameters and for both Euclid and SKA. Since, as remarked above, the approximated formula (5.22) is accurate for small shifts, we prefer to be more conservative and use the Newtonian spectra for the derivatives with respect to the parameters $\{\theta_\alpha\}$. The second issue is related to the errors $\sigma^2_{C_\ell}$. Again we have two choices: to include the lensing contribution in the measured spectra or not. In this case we do take into account the lensing contribution simply because, even if it is not included in the theoretical description of the galaxy clustering, it is nevertheless observed.

To summarize, for the terms in Eq. (5.22) we take

$$F^{\theta\theta}_{\alpha\beta} = \sum_{\ell=2}^{\ell_{\max}} \sum_{(ij)(pq)} \frac{\partial C_\ell^{\mathrm{Nwt}\,ij}}{\partial\theta_\alpha} \frac{\partial C_\ell^{\mathrm{Nwt}\,pq}}{\partial\theta_\alpha} \sigma^{-2}_{C_\ell\,[(ij),(pq)]}\,, \tag{5.23}$$

[4]We stress that for the correct calculation of the shift in Eq. (5.22) one should first consider the subset $F^{\theta\theta}$ of the full Fisher matrix $F^{\Phi\Phi}$ and then invert it. This was first pointed out in Ref. (Camera et al. 2017).

and

$$F^{\theta\epsilon}_{\alpha\beta} = \sum_{\ell=2}^{\ell_{\rm max}} \sum_{(ij)(pq)} \frac{\partial C_\ell^{{\rm Nwt}\,ij}}{\partial\theta_\alpha} C_\ell^{{\rm Lens}\,pq} \sigma^{-2}_{C_\ell\,[(ij),(pq)]}\,, \tag{5.24}$$

where we remind the reader that $C_\ell^{\rm Lens}$ is the derivative of the spectra including lensing with respect to ϵ_L and we include the lensing contribution in the errors $\sigma^2_{C_\ell}$. Our choices follow the approach of (Kitching et al. 2009) and are slightly different to the one used in previous literature for the same effect, i.e. neglecting lensing (and the other relativistic) effects, see (Camera et al. 2015a; Di Dio et al. 2016b; Cardona et al. 2016): in these works the lensing contribution was neglected in the modelled (as we do) but also in the observed power spectra.

5.3.3 Surveys specifications

In this section we present the specifications that we assume for the galaxy surveys we consider: Euclid and SKA. They are summarized in figure 5.1.

We consider Euclid photometric specifications as presented in (Amendola et al. 2013): the number of galaxies per redshift and per steradian, the galaxy bias and the magnification bias are given by

$$\frac{{\rm d}N}{{\rm d}z d\Omega} = 3.5\times10^8 z^2 {\rm exp}\left[-\left(\frac{z}{z_0}\right)^{3/2}\right] \tag{5.25}$$

$$\text{for} \quad 0.1 < z < 2 \tag{5.26}$$

$$f_{\rm sky} = 0.375 \tag{5.27}$$

$$b(z) = b_0\sqrt{1+z}\,, \tag{5.28}$$

where $z_0 = z_{\rm mean}/1.412$, the median redshift is $z_{\rm mean} = 0.9$ and we set $b_0 = 1$. We assume the magnification bias as computed in (Montanari & Durrer 2015):

$$s(z) = s_0 + s_1 z + s_2 z^2 + s_3 z^3\,, \tag{5.29}$$

where the coefficients are $s_0 = 0.1194$, $s_1 = 0.2122$, $s_2 = -0.0671$ and $s_3 = 0.1031$. Galaxy and magnification biases are assumed to be constant in each redshift bin, with value determined at the mean redshift. We consider two configurations, with 5 and 10 equally spaced redshift bins in the redshift range $0.1 < z < 2$ with $\Delta z_5 = 0.38$ for 5 bins and $\Delta z_{10} = 0.19$ for 10 bins. The photometric redshift uncertainty is modelled with Gaussian bins with standard deviation $\Delta z_i/2$, $i = 5, 10$, for the observed power spectra in the errors $\sigma^2_{C_\ell}$ whereas we use top-hat bins with half-width $\Delta z_i/2$ for the theoretical power spectra in the derivatives with respect to the parameters in the Fisher matrix.

For the SKA spectroscopic survey we assume specifications as in (Camera et al. 2015b; Santos et al. 2015) consistent with the 5 μJy sensitivity:

$$\frac{\mathrm{d}N}{\mathrm{d}z d\Omega} = \left(\frac{180}{\pi^2}\right) 10^{c_1} z^{c_2} \exp\left(-c_3\right) \tag{5.30}$$

$$\text{for} \quad 0.1 < z < 1.5 \tag{5.31}$$

$$f_{\rm sky} = 0.73 \tag{5.32}$$

$$b(z) = b_0 \exp\left(b_1 z\right) , \tag{5.33}$$

where $c_1 = 6.7767$, $c_2 = 2.1757$, $c_3 = 6.6874$, the fiducial value for the galaxy bias at $z = 0$ is $b_0 = 0.5887$ and $b_1 = 0.8130$.

We compute the magnification bias according from the fitting formula for the cumulative number density of galaxies for SKA in Eq. (5.4) as given in (Camera et al. 2015b). From Eq. (5.3) we find

$$s(z) = s_0 + s_1 z + s_2 z^2 + s_3 z^3 , \tag{5.34}$$

with coefficients $s_0 = -0.106875$, $s_1 = 1.35999$, $s_2 = -0.620008$ and $s_3 = 0.188594$. Note that, physically, the magnification bias assumes positive values at all redshifts[5]. Galaxy and magnification biases are assumed to be constant in each redshift bin, with value determined at the mean redshift.

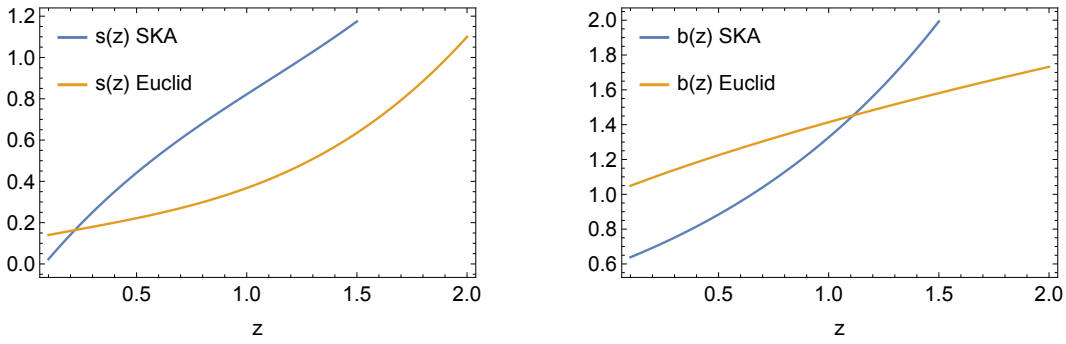


Figure 5.1: Magnification bias (left panel) and galaxy bias (right panel) for the EUCLID and SKA specification considered in this chapter. A constant galaxy and magnification bias is assumed in each redshift bin, with value evaluated at the mean redshift.

For SKA, we consider five configurations, with 5, 10, 20, 30 and 40 equally spaced redshift bins in the redshift range $0.1 < z < 1.5$. Given the spectroscopic

[5]We notify the reader that the fitting formula for the magnification bias for SKA appearing in Refs. (Raccanelli et al. 2016a,c; Montanari & Durrer 2015; Di Dio et al. 2016b) is incorrect. This is due an error in fitting formula for the cumulative number density of galaxies in Ref. (Camera et al. 2015b) which was corrected afterwards.

redshift determination, we use top-hat redshift bins with half-width $\Delta z_i/2$, $i = 5, 10, 20, 30, 40$, both for the observed and for the theoretical power spectra in the Fisher matrix. For SKA we also consider the case where the cross-correlations between far redshift bins including lensing are neglected. To be specific we discard all cross-cross-correlations with separation $\Delta z > \Delta z_5 = 0.28$. The reason is twofold: firstly in this way our results mimic to some extend what one finds with a $P(k)$ analysis by neglecting long range correlations, even if we are using the angular power spectra. The second reason is that in this configuration we can completely trust the robustness of the method we use to calculate the bias in the parameter estimation and our results: as we remark above, the Fisher matrix method is based on the Taylor expansion of the likelihood and thus it is implicitly based on the assumption that the difference between the two models, here angular power spectra with or without lensing, is small. This is not the case for far cross-bin correlations with a spectroscopic survey, where the amplitude of the Newtonian spectra is completely negligible and the signal is completely dominated by the lensing convergence contribution. But it is true for cross-correlations between adjacent bins, at all scales. To recover the 3-dimensional information with a spectroscopic survey one needs order of 10^2 bins. Here we limit our analysis to a maximum of 40 redshift bins because of computational convenience.

5.4 Results

The correlation between the cosmological parameters and the lensing parameter ϵ_L is one of the most important quantity to understand the behaviour of the effects of lensing that we want to study. It is defined by

$$\rho_{\epsilon_L}(\theta_\alpha) = \frac{\left[\left(F^{\Phi\Phi}\right)^{-1}\right]_{\alpha\epsilon_L}}{\sqrt{\left[\left(F^{\Phi\Phi}\right)^{-1}\right]_{\alpha\alpha}\left[\left(F^{\Phi\Phi}\right)^{-1}\right]_{\epsilon_L\epsilon_L}}}\,, \tag{5.35}$$

where the index α refers to the cosmological parameters $\{\theta_\alpha\}$ and we have included lensing in the Fisher matrix, both in the covariance and in the angular power spectra we use for the derivatives. The more this quantity is close to 1 the more the correspondent parameter is correlated with lensing and therefore we would expect the impact of lensing to be more important for the constraints and for the shift of the best-value of that parameter.

In figure 5.2 and in figure 5.3 we plot this correlation for the ΛCDM model including massive neutrinos and including also modified gravity, respectively. By

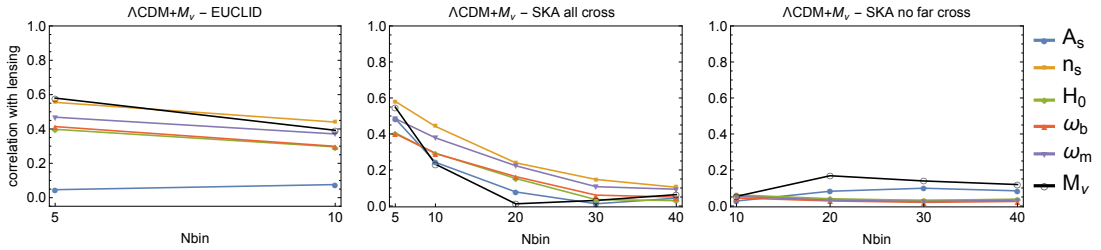


Figure 5.2: The correlation of the lensing parameter ϵ_L with the parameters of the ΛCDM model with massive neutrinos. The correlation is given for different binning configurations for Euclid (left panel) and for SKA considering all the redshift bin cross-correlations (middle panel) and neglecting the cross-correlations between far redshift bins (right panel).

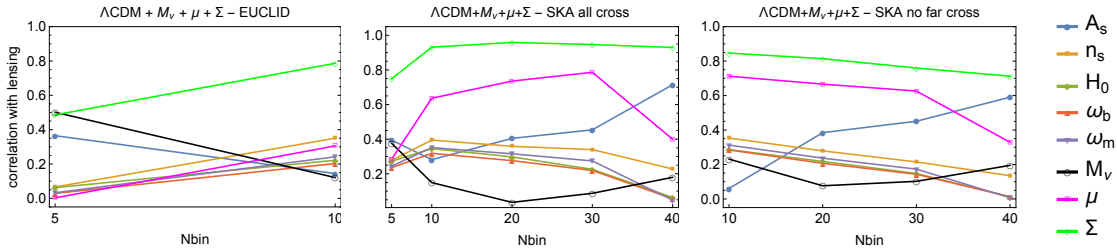


Figure 5.3: The correlation of the lensing parameter ϵ_L with the parameters of the ΛCDM model including massive neutrinos and modified gravity. The correlation is given for different binning configurations for Euclid (left panel) and for SKA considering all the redshift bin cross-correlations (middle panel) and neglecting the cross-correlations between far redshift bins (right panel).

looking first at figure 5.2 we see that all the ΛCDM parameters, including the sum of neutrino masses, are all not much sensitive to lensing. From the results for SKA when we include all the possible cross-correlations we also see that the correlation with ϵ_L decreases as the number of redshift bins increases, as one would expect since the lensing signal becomes relatively weaker. Indeed, by increasing the number of bins, we consider more radial modes which are dominated by the standard Newtonian contributions, being the lensing a pure transversal effect. In addition, in the case where we neglect far-bin cross-correlations the correlation with ϵ_L is in general smaller for all the parameters with values which do not exceed 0.2 and are independent of the number of bins. When we include modified gravity, the correlation of the ΛCDM parameters and the neutrino mass M_ν behaves in the same way as for the ΛCDM + M_ν case, with the only exception of the amplitude of scalar perturbations A_s, which becomes more correlated with lensing especially for SKA, as shown in figure 5.3. The parameter Σ which parametrizes the sum of the two Bardeen potentials in modified gravity (see Eq. (5.13)) is the parameter most correlated with lensing, as one would expect since lensing is crucial to probe

this quantity. This can be clearly seen from figure 5.3, especially for SKA, where the correlation between Σ and ϵ_L stays constant and close to 1 as the number of redshift bins increases (middle panel of figure 5.3), also for the configuration where we drop the far-bin correlations (right panel of figure 5.3). Note that as the number of redshift bins increases the correlations with ϵ_L for SKA including all the cross-correlations and those neglecting far-bin cross-correlations become almost the same, for all the three models we consider. This indicates that as one increases the number of bins the correlations with lensing is not controlled by the far-bin cross-spectra, despite the lensing signal dominates over the Newtonian terms. This is a first indication that for most of the cosmological parameters, lensing magnification is not carrying crucial information for a spectroscopic survey.

5.4.1 The effect of lensing magnification: constraints on the cosmological parameters

In this section we quantify how much information is present in the lensing contribution in galaxy clustering. Our aim here is not to provide precise forecast constraints, rather to estimate the relative impact of lensing compared to the standard Newtonian analysis. We therefore present our results in terms of the ratio of the marginalized errors calculated with two different Fisher analyses, the difference being the contributions we consider in the model for the galaxy power spectra: intrinsic clustering and redshift-space distortions only in one case and including also lensing convergence in the other case. If this ratio is close to 1 then the constraining power of lensing is negligible. On the contrary, for values significantly less than 1, the improvement on the errors coming from lensing is important. The errors are marginalized over the entire set of the parameters for each model we consider: we have 5 parameters for ΛCDM, 6 parameters when we include massive neutrinos and 8 parameters if we also include modified gravity. We address to section 5.3.1 for a detailed discussion of our approach. We also analyse the dependence of the constraining power of lensing on the number of redshift bins. The relative behaviour of the errors on a parameter - with or without lensing - is due to the dependence of the lensing magnification on that parameter and is also due to the dependence of the lensing signal on the survey and the binning configuration.

Figure 5.4 shows the results for Euclid and figure 5.5 for SKA. Again we look first to the ΛCDM model and its extension with massive neutrinos. From left and middle panels in figure 5.4 and in figure 5.5 we see that the information contained in the lensing contribution does not improve substantially the constraints: the ratios of the errors for all the parameters rapidly approach 1 as the number of redhsift

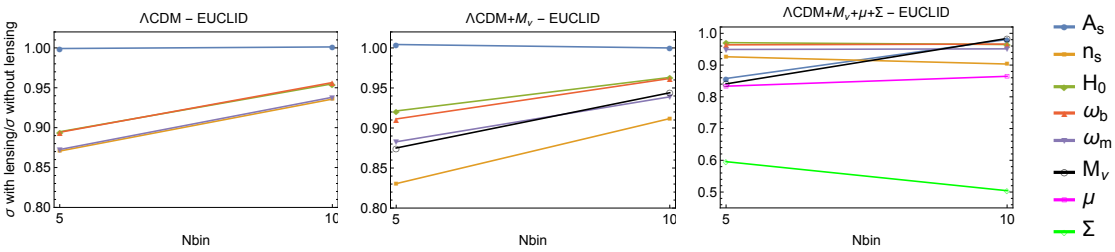


Figure 5.4: Results for Euclid: the ratio of the errors on the cosmological parameters calculated including lensing to the ones calculated neglecting it. In both analyses we include lensing contribution in the covariance matrix.

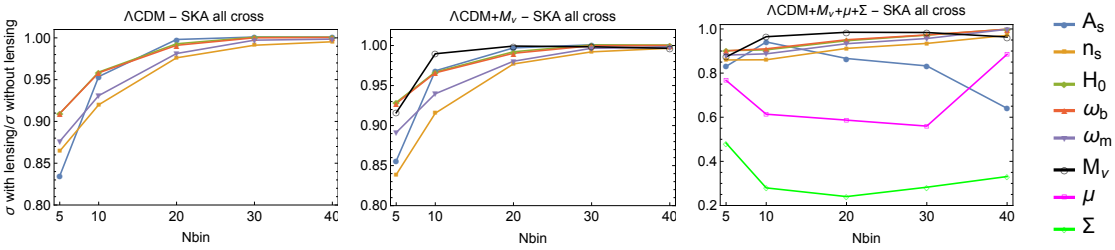


Figure 5.5: Results for SKA: the ratio of the errors on the cosmological parameters calculated including lensing to the ones calculated neglecting it. In both analyses we include lensing contribution in the covariance matrix.

bins increases. The results are also shown in table 5.1 in section 5.5. In the right panel of figure 5.4 and 5.5 we show the results for modified gravity. We see that lensing improves the constraints on the modified gravity parameters, the strongest improvement being on Σ: for Euclid with 10 bins the error with lensing is one half the error without lensing and for SKA with 40 bins and considering all the cross-correlations the ratio of the errors is 0.33, see table 5.2 in section 5.5. In addition, for SKA the right panel of figure 5.5 tells us that adding lensing magnification in modified gravity improves the constraints on the amplitude of primordial fluctuations A_s, because the lensing potential breaks degeneracy between these parameters.

It is worth discussing in more detail the behaviour of the relative error of the modified gravity parameters μ and Σ for the case of the spectroscopic survey SKA. By inspecting the right panel of figure 5.5 we note that the improvement on the error of Σ is significant for the 30 and 40 bins configuration, the ratio of the errors with and without lensing being lower than 0.4. On the contrary for the parameter μ the ratio of the errors pass from 0.5598 at 30 bins to 0.8853 at 40 bins - approaching the value of the other ΛCDM parameters - indicating that the constraining power of lensing on μ decreases as the number of redshift bins increases. Indeed by increasing the number of redshift bins, we start considering shorter radial modes. With a better radial resolution we resolve the redshift space distortion contributions. This

leads to a measurement of the growth rate and it provides constraints on the galaxy clustering parametrized by μ. Therefore the constraints on μ are driven by the standard newtonian contributions in the limit of a spectroscopic survey.

To corroborate this statement we isolate the effect of redshift-space distortions by performing an additional Fisher analysis to examine their constraining power on the modified gravity parameters. In figure 5.6 we show on the right panel the ratio of the errors calculated including density and redshift-space distortion to the ones calculated with density alone and we report on the left panel the ratio of the errors given by including to neglecting lensing. It is clear the different behaviour of the two modified gravity parameters. From the right panel we see that redshift-space distortions improve the error on μ whereas they do not add useful information for Σ starting from 30 bins, the ratio of the errors on Σ becoming closer to 1 for 40 bins. On the contrary, from the left panel we see the opposite behaviour: starting from 30 bins adding lensing stops improving the error on μ whereas the improvement of the errors on Σ stay roughly constant by increasing the number of bins. Our results also validate the choice of the parametrization $\{\mu, \Sigma\}$ to test deviation from GR on large scales. Indeed with this choice we decouple the effect induced by structure formation on μ and the lensing effect on Σ.

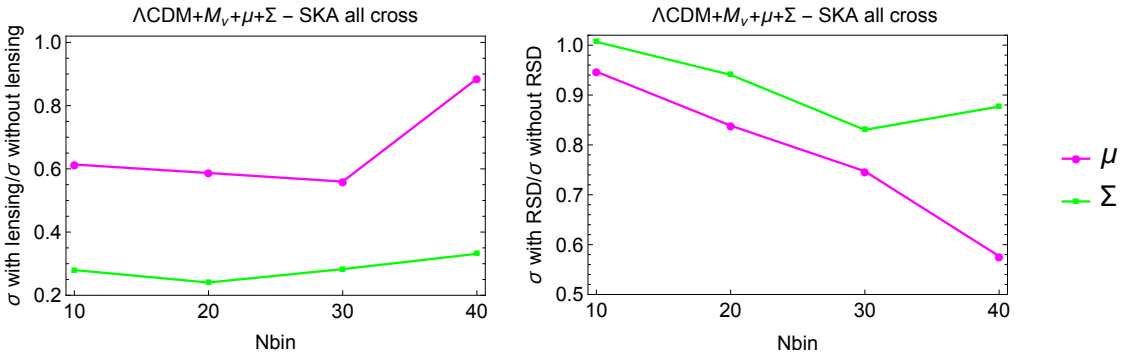


Figure 5.6: Left panel: the ratio of the errors on the modified gravity parameters calculated including lensing to the ones calculated neglecting it. Right panel: the ratio of the errors on the modified gravity parameters calculated including density and redshift-space distortion to the ones calculated with density alone. The errors are calculated by marginalizing over the other parameters of the $\Lambda\text{CDM} + M_\nu + \mu + \Sigma$ model. Here we consider all the cross-correlation between the redshift bins for SKA.

5.4.2 The effect of lensing magnification: bias on the cosmological parameters estimation

In this section we present our results for the shift in the best-fit values of the ΛCDM and modified gravity parameters. We report in the same plots in figure 5.7 and

figure 5.8 for every parameter: its correlations with the lensing parameter, the ratio of the errors with lensing to the one obtained without and the results for the shift in the best-fit values. This is to highlight the correspondence that we expect between these three different quantities: the parameters whose correlation with the lensing parameter is higher are those whose estimation is more biased by neglecting lensing and also those whose constraints are more improved by including lensing. We express the shifts, given by Eq. (5.22), in units of the error we get from the standard Newtonian analysis. From the left and middle panels of Figs. 5.7 and 5.8, it is evident the one-to-one correspondence between improvement on the errors, correlation with lensing and shifts, both Euclid and SKA, for the ΛCDM model and for the extension including massive neutrinos. When we add modified gravity the values of the shifts is not completely determined by the correlation with the lensing parameter because it is influenced also by the correlations between the parameters of the model itself, which are altered in modified gravity.

For Euclid, and in general for photometric surveys, we find shifts larger than 1σ, and even more for the modified gravity parameters. We again explain this behaviour considering that photometric surveys have a much better angular resolution than radial, therefore they are more sensitive to the lensing effects and by neglecting its contribution in the analysis will affect the cosmological parameters process stronger. Note that for the ΛCDM and for the ΛCDM + M_ν models the shifts on the cosmological parameters are all around 1σ, reaching 2σ for the spectral index n_s and neutrino mass M_ν, whereas they drop to smaller values when adding modified gravity, see tables 5.1 and 5.2 in section 5.5.

For SKA, for the ΛCDM and for the extension with neutrinos the shifts on all the cosmological parameters are below 1σ and sistematically lower than the ones for Euclid, as expected, see left and middle panels of figure 5.8 compared with those of figure 5.7. When including modified gravity, the bias on the cosmological parameters and neutrino mass are lower than in the ΛCDM and the ΛCDM + M_ν models and we find a bias of more than 1σ on μ and Σ only for the 10 bins configuration. In this respect, it is worth remarking the dependence of the shifts on the modified gravity parameters on the number of bins: for 10 bins the lensing signal is such that the shifts on μ and Σ are more than 1σ. But by increasing the number of redshift bins the shifts on μ and Σ reflect the relative behaviour of the constraining power of lensing and redshift-space distortions that we have already remarked in section 5.4.1. In particular we find that when we approach a number of bins where the angular analysis is competitive with respect to the full 3-dimensional analysis, i.e. when the width of the bin is comparable to the

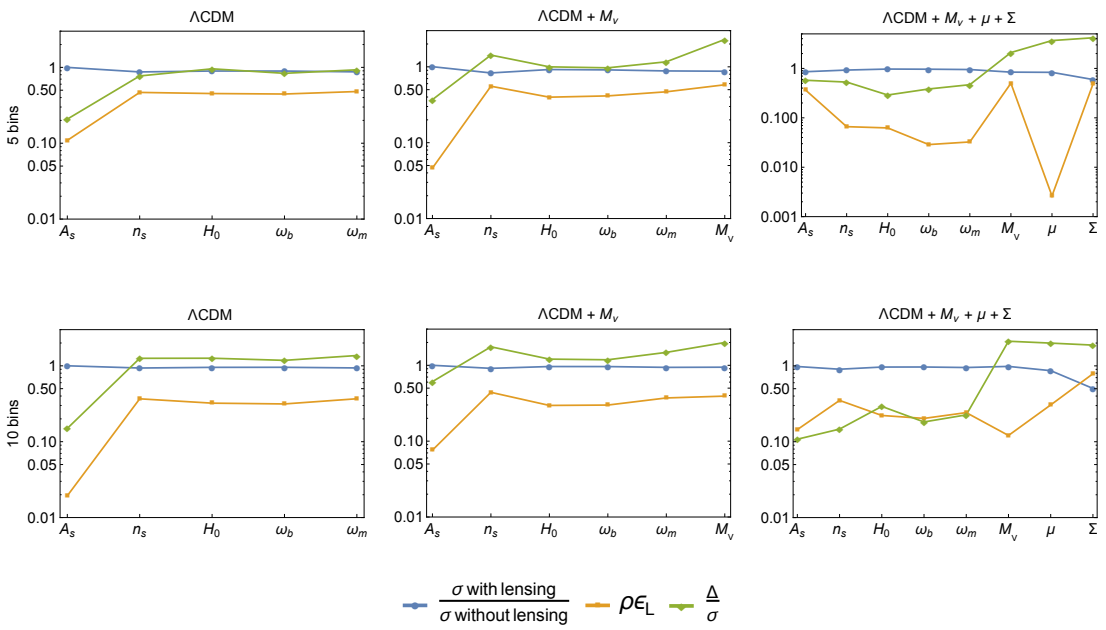


Figure 5.7: Results for Euclid: we show the shifts in the best fit values in units of the errors from the Newtonian analysis (green line), the correlation with the lensing parameter (orange line) and the ratio between the errors with lensing to the one without lensing (blue line).

non-linear scale (Asorey et al. 2012; Di Dio et al. 2014b), the shift are below 1σ. We may conclude that SKA is not sensitive to any bias induced by neglecting lensing magnification. Nevertheless we want to stress that the shift on the parameter Σ if expressed in units of the error obtained including lensing is about 1σ, whereas the shifts on the other cosmological parameters remain unchanged.

As part of our analysis for SKA we also compare the shifts that we get by considering all the cross-correlations with those obtained by setting to zero those between far redshift bins. From figure 5.9 we see that we basically obtain the same results with both configurations and for all the parameters, including μ and Σ. This means that the lensing contribution from far-bin cross-correlations is sub-dominant for the estimation of the bias of the best-fit value, despite the lensing convergence dominates the signal of the angular power spectra of these cross-correlations.

Let us finally comment on the dependence of the bias in parameter estimation on the choices for the maximum multipole. In this case we consider Euclid, which exhibits the most significant shifts. In figure 5.10 we show the dependence of the errors (left panel) and the dependence of the shifts (in units of the errors, right panel) on the maximum multipole $\ell_{\rm max}$. For both the quantities it is clear the separation in two groups: neutrino mass and modified gravity parameters at one side and all the parameters of the standard ΛCDM model on the other side. Given

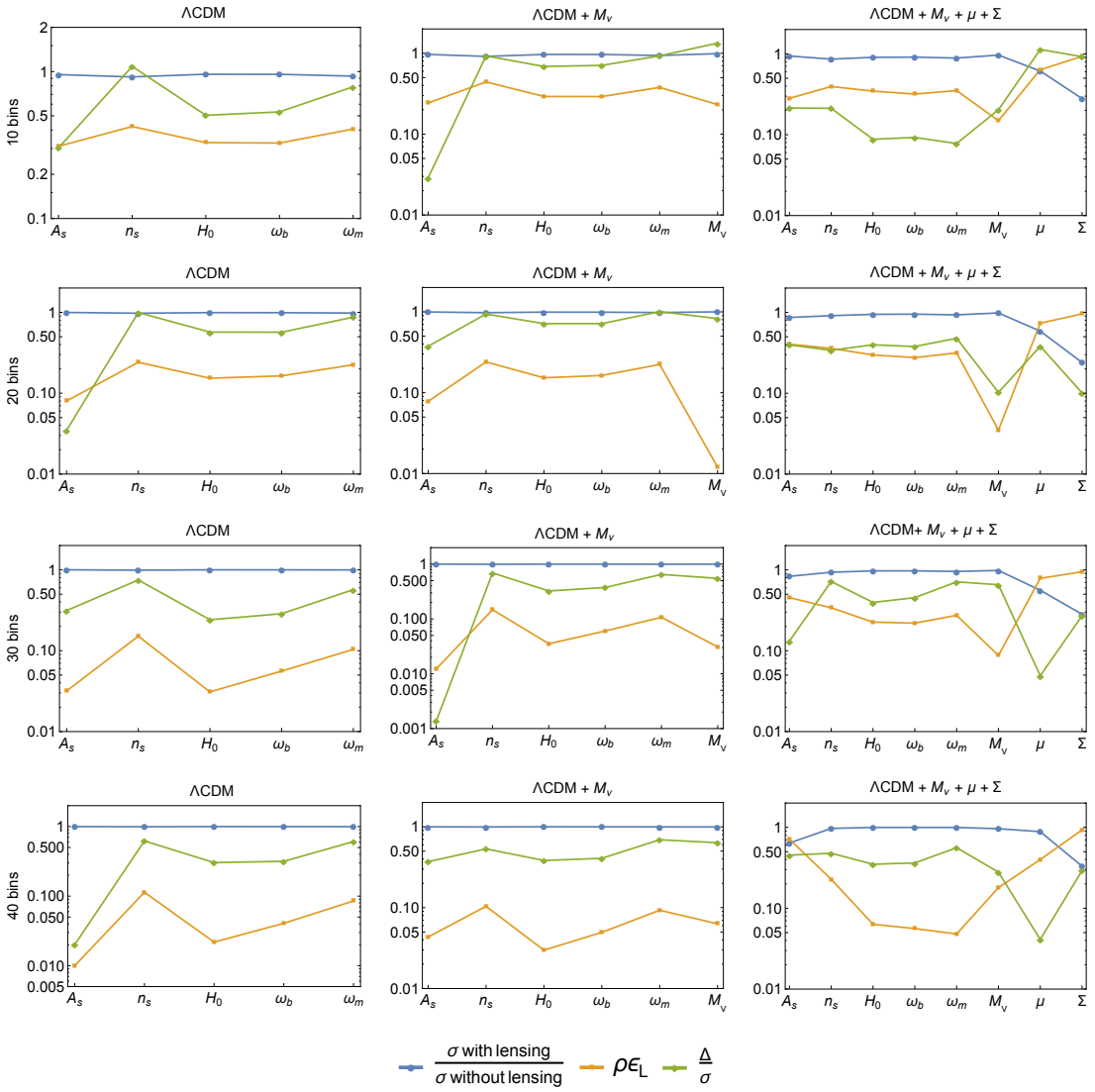


Figure 5.8: Results for SKA: we show the shifts in the best-fit values in units of the errors from the Newtonian analysis (green line), the correlation with the lensing parameter (orange line) and the ratio between the errors with lensing to the one without lensing (blue line). Here we consider all the cross-correlations.

that if we include more multipoles the errors decreases for each parameter, from the left panel we see that for neutrino mass, μ and Σ the information coming from including larger multipoles is less significant, meaning that they are more sensitive to larger scales. We also remark that on large scale there is a more relevant contribution due to the pure lensing term, while on smaller scales the lensing signal is dominated by its correlation with the standard Newtonian terms. This behaviour is mirrored in their shifts: the major contribution comes from larger scale and they do not change significantly as $\ell_{\rm max}$ increases. Therefore, once the shifts are expressed in terms of the errors, the information gained from smaller scales does

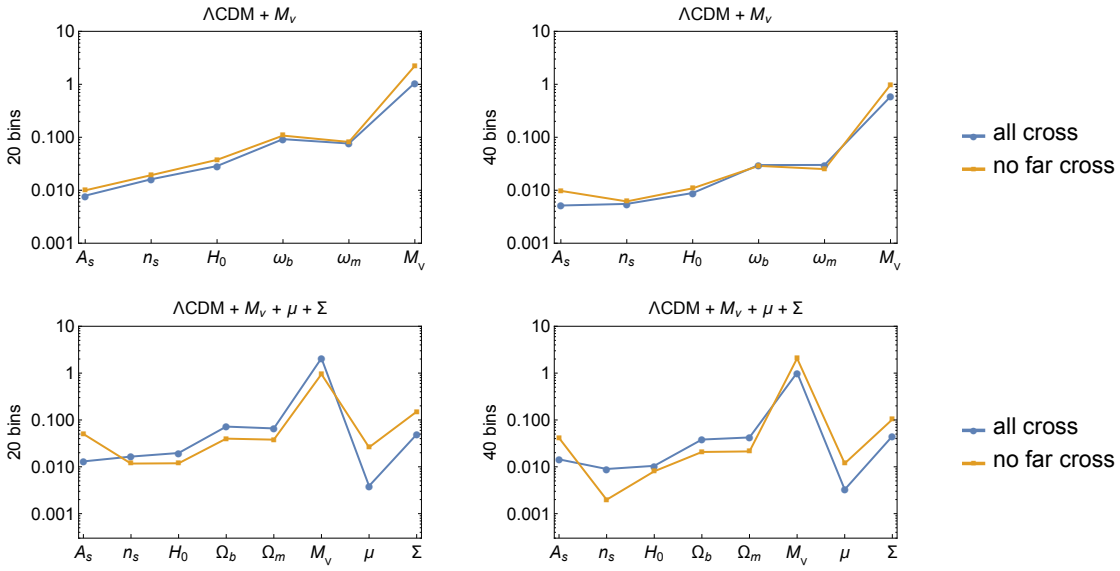


Figure 5.9: Shifts in the best-fit values of the parameters for the two extensions of the ΛCDM model for SKA. We compare the results obtained by including all the cross-correlations (blue line) and those obtained by neglecting the far-bin cross-correlations (orange line). The shifts are given in units of the fiducial values for both the configurations.

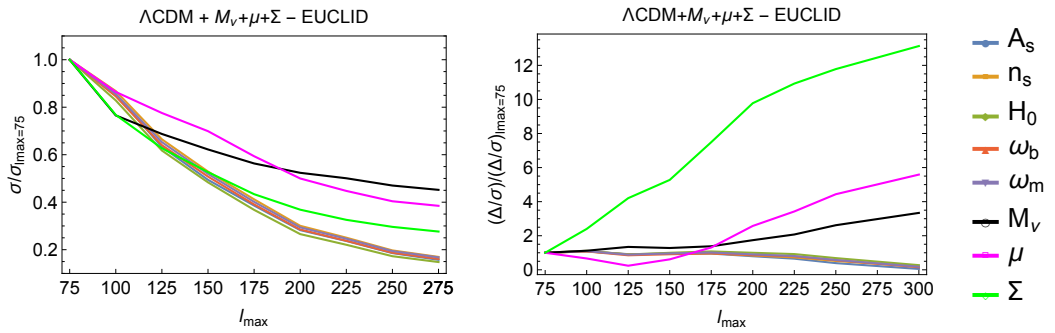


Figure 5.10: Results for Euclid in the 10 bins configuration. Left panel: errors for the cosmological and modified gravity parameters for different values of the maximum multipole. Right panel: shift in the best-fit values of the cosmological and modified gravity parameters in units of the standard error for different values of the maximum multipole. Both are normalized to $\ell_{\rm max} = 75$. Here the errors are calculated with the Newtonian theoretical power spectra and lensing in the covariance.

not compensate the bias from larger scales and the result is that the values of Δ/σ increases, as show in the right panel of figure 5.10.

5.5 Summary

In this chapter we have investigated the impact of neglecting lensing magnification in future galaxy clustering surveys, motivated by the deeper redshift range probed by them. We consider three cosmological models, namely the standard ΛCDM model

and two extensions: the first including massive neutrinos and the second including deviations to General Relativity. We focus both on photometric and spectroscopic redshift surveys, showing their different behaviour regard lensing magnification.

In particular, photometric surveys due to the poor redshift resolution measure mainly transversal modes which are more affected by lensing magnification. This leads to a larger bias induced by neglecting lensing magnification, in particular for the terms beyond ΛCDM. We show that the cosmological parameters which exhibit a larger bias are the ones more correlated with lensing magnification. This is clear for the parameters (μ and Σ) which parametrise deviations from General Relativity, where a crucial information is carried by the lensing potential.

To fully consider the impact on spectroscopic surveys we would need a number of redshift bins such that the width of them is comparable with the non-linear scale. While this imposes strong numerical limitations, we show the behaviour by increasing the number of redshift bins, hence by approaching a spectroscopic resolution. By doing that we including more and more radial modes which are unaffected by lensing (being a pure transversal effect) and therefore the relative information carried by lensing potential is less relevant. Interestingly this does not apply completely to the Σ parameter, whose information is mainly carried by the lensing potential. As well known redshift space distortion measures the growth of structures, and it can be used to constrain the underling theory of gravity. In our analysis we show that when we start approaching a good enough redshift resolution and we begin to probe (small) scale radial modes, the constraining power on the μ parameter is dominated by redshift space distortion compared to lensing magnification. This different behaviour between Σ and μ reflects the modification of eqs. (5.12) and (5.13).

We have also shown that, while lensing magnification dominates the cross-correlation between far away redshift bins, the information encoded in such correlations is very subdominant with respect to closer redshift correlations and by including or not far cross-correlations does not change the results.

We have therefore shown that it is important to include lensing magnification in data analysis, in order to use next generation of clustering surveys to test for deviations from General Relativity. In particular, our results show that the estimation of the modified gravity parameters is biased for photometric surveys whereas the shift in the best-fit values stay below 1σ for spectroscopic surveys.

Some similar issues on the same topic are addressed in (Lorenz et al. 2017) which appeared while the work presented in this chapter was in the final stage of preparation. Wherever possible we verified that our results agree with theirs. We emphatise that our results address the issues for spectroscopic surveys as well.

We show here the values that we obtain for the correlation with the lensing parameter, the ratio of the errors for the cosmological and modified gravity parameters - errors with lensing to errors without lensing - and the shift in their best-fit value expressed in unit of the sigma obtained from a Newtonian analysis. We report our results for the 10 bins configuration for Euclid and for the 40 bins for SKA.

	EUCLID - 10 bins			SKA - 40 bins		
	$\rho_{\epsilon_{\rm L}}$	$\sigma_{\rm lens}/\sigma_{\rm w/o\ lens}$	Δ/σ	$\rho_{\epsilon_{\rm L}}$	$\sigma_{\rm lens}/\sigma_{\rm w/o\ lens}$	Δ/σ
A_s	0.07617	1.000	-0.6094	0.005150	0.9990	0.3705
n_s	0.44019	0.9118	-1.753	-0.005522	0.9966	-0.5332
H_0	-0.2958	0.9632	1.206	0.008866	1.001	0.3834
ω_b	-0.2991	0.9619	1.183	0.02988	1.000	0.4077
ω_m	-0.3711	0.9389	1.480	0.03003	0.9977	0.6918
M_ν	0.3923	0.9440	-1.996	0.5854	0.9965	0.6359

Table 5.1: Summary of results for the extended ΛCDM model (ΛCDM $+M_\nu$). Here we report the results for SKA obtained by considering all the cross-correlations.

	EUCLID - 10 bins			SKA - 40 bins		
	$\rho_{\epsilon_{\rm L}}$	$\sigma_{\rm lens}/\sigma_{\rm w/o\ lens}$	Δ/σ	$\rho_{\epsilon_{\rm L}}$	$\sigma_{\rm lens}/\sigma_{\rm w/o\ lens}$	Δ/σ
A_s	-0.1445	0.9806	-0.1079	-0.7145	0.6403	0.4535
n_s	0.3504	0.9033	-0.1467	0.2279	0.9709	-0.4784
H_0	-0.2219	0.9647	0.2922	-0.06334	0.9990	0.3500
ω_b	-0.2030	0.9661	0.1822	-0.05601	0.9990	0.3634
ω_m	-0.2420	0.9511	0.2259	-0.04806	0.9983	0.5603
M_ν	0.1209	0.9838	-2.112	-0.1797	0.9637	0.2793
μ	-0.3055	0.8646	-1.996	-0.3998	0.8852	0.04112
Σ	-0.7867	0.5042	-1.878	-0.9299	0.3312	0.2919

Table 5.2: Summary of results for the extended ΛCDM model (ΛCDM $+M_\nu + \mu + \Sigma$). Here we report the results for SKA obtained by considering all the cross-correlations.

Chapter 6

OPTIMAL GALAXY SURVEY FOR DETECTING THE DIPOLE IN THE CROSS-CORRELATION WITH 21 CM INTENSITY MAPPING

6.1 Introduction

The state-of-the-art analysis of the Cosmic Microwave Background (CMB) anisotropies (Ade et al. 2016a) and of the Large Scale Structure (LSS) of the universe (Alam et al. 2017b) offers robust observational evidence in favor of the standard cosmological model, known as ΛCDM. However, the ΛCDM model presents some theoretical troubles, above all the cosmological constant problem (Weinberg 1989) and the nature of the dark components in the matter sector (Zwicky 1937; Clowe et al. 2006). These issues keep raising an increasing interest in the scientific community and strongly necessitate further investigation of the foundational principles from which the standard model was built (Bull et al. 2016). A pivotal role in testing the laws of gravity at cosmological scales will be played by the forthcoming LSS surveys aiming to map with stunning precision the large scale structure of the universe. The huge amount of incoming data requires an equally powerful advance from the theoretical and modelling side. One of the issues related to the correct modelling of the LSS observables is the fact that we do not observe directly the dark matter density field, but we observe some biased tracers at a given observed redshift and direction in the sky. A fully relativistic computation of what is really measured in a galaxy survey was performed at linear order (Yoo et al. 2009; Yoo 2010; Bonvin & Durrer 2011; Challinor & Lewis 2011; Schmidt & Jeong 2012; Bertacca et al. 2012; Raccanelli et al. 2016b, 2014) and at second order in perturbation theory (Yoo & Zaldarriaga 2014; Bertacca et al. 2014; Bertacca 2015; Di Dio et al. 2014a, 2016a; Nielsen & Durrer 2017; Umeh et al. 2017; Di Dio 2017; Jolicoeur

et al. 2017; Bertacca et al. 2017). This formalism has been further extended to other observables such as the lensing convergence (Bonvin 2008), the HI brightness temperature (Hall et al. 2013) and the Lyman-α transmitted flux (Iršič et al. 2016). Since we measure galaxy positions by collecting photons which have travelled in a clumpy universe, the LSS observables are also affected by other contributions, which were commonly neglected, beyond the standard local overdensity and Redshift Space Distortion (RSD) (Kaiser 1987). These terms, which include gravitational lensing convergence, gravitational redshift, Doppler and integrated Sachs-Wolfe effect, may bias the estimation of some cosmological parameters (Camera et al. 2015b,a; Raccanelli et al. 2016c; Montanari & Durrer 2015; Di Dio et al. 2016b; Cardona et al. 2016), in particular for deviations from ΛCDM (Villa et al. 2018). More interestingly, a measurement of these effects offers the opportunity to extract further cosmological information, allowing to test the Equivalence Principle or the theory of gravity at the largest observable scales. Even if all these corrective terms are subdominant on sub-horizon scales with respect to the local density and RSD, it has been shown that it is possible to isolate some of them, by correlating two different tracers[1] (McDonald 2009; Yoo et al. 2012; Bonvin et al. 2014; Alonso & Ferreira 2015; Fonseca et al. 2015; Iršič et al. 2016; Bonvin et al. 2016; Gaztanaga et al. 2017; Dai et al. 2016; Borzyszkowski et al. 2017; Abramo & Bertacca 2017). This can be achieved using the fact that some terms carry an odd symmetry with respect to the line of sight, which can be exploited with different tracers. In fact, in the same way as RSD introduces an anisotropy that sources other even multipoles in the correlation function of a LSS tracer apart from the monopole, gravitational redshift and Doppler effects break the symmetry with respect to the exchange of two tracers along the line of sight. Therefore, they induce non-vanishing odd multipoles in the cross-correlation function or an imaginary part in the Fourier power spectrum. Measuring the relativistic effects is a new opportunity for testing the consistency of general relativity (Lombriser et al. 2013; Bonvin 2014) and may offer an alternative method to measure the peculiar velocity field of the sources (Hall & Bonvin 2017; Bonvin et al. 2017). The first measurement of the dipole for the cross-correlation of two galaxies populations (Gaztanaga et al. 2017) was a measurement of the dipole induced by the so-called large-angle effect. Nevertheless, at the lowest order in the distant observer approximation, the large-angle effect can be written as a combination of the monopole and the quadrupole and therefore it does not provide new information, at least in the regime where this approximation is reliable. In this

[1] It has also been claimed (Raccanelli et al. 2016a) that futuristic surveys may be able to detect them through a single tracer analysis.

work we will investigate the possibility to detect the relativistic Doppler corrections by cross-correlating galaxies and 21 cm Intensity Mapping (IM). The prospect of measuring the relativistic dipole by cross-correlating these two tracers was addressed in (Hall & Bonvin 2017), where a signal-to-noise analysis is presented for specific IM experiments and galaxy survey. Here we present a survey-independent analysis, with the aim of understanding which types of galaxies are more suitable for detecting the relativistic dipole and how the signal-to-noise can be optimized by properly choosing the luminosity threshold of the galaxy catalogue. Since the Doppler corrections are more relevant at low redshift, this is the regime we are focused on.

Cross-correlation observational studies between galaxies and IM at low redshift ($z < 1$) have already been performed in recent years with the goal of detecting the diffuse neutral hydrogen (HI) (Chang et al. 2010; Masui et al. 2013) and also at high redshift exploiting the cross-correlation between IM and Lyman-break galaxies (Villaescusa-Navarro et al. 2015) and the Lyman-α forest (Carucci et al. 2017) in the post-reionization era. Furthermore, higher redshift investigations of the cross correlations are also of primary importance during or around HI reionization (e.g. (Lidz et al. 2009)).

The chapter is structured as follows. In section 6.2 we revise the relativistic formalism for galaxy surveys and IM experiments. Furthermore, we summarize the formalism that will be used in the following sections to compute the multipoles of the cross-correlation for the two tracers and its covariance, and we discuss two possible contaminations to the relativistic dipole: the lensing dipole and the wide-angle effect. More details on the formalism and the approximation we adopt in this work can be found in the appendices B.1 and B.2. In section 6.3 we use the halo model to describe in a fairly general way the HI and galaxy distribution properties, and we show how the biases of the two tracers can be related to their shot-noise. In section 6.4 we present a signal-to-noise analysis for the relativistic dipole. In particular, we investigate the behavior of the signal-to-noise as a function of the galaxy bias and magnification bias, for two HI models. In section 6.5 we model luminosity-threshold galaxy catalogues with an halo occupation distribution model and we studied the signal-to-noise for the relativistic Doppler dipole as a function of the limiting magnitude. Finally, in section 6.6 we sum up the results of our work and we draw the conclusions.

We stress that the rationale of this work is to provide a first quantitative investigation of the optimal strategy to detect relativistic effects by exploiting the cross-correlation signal between IM and galaxies, in doing that we will learn that the low redshift regime is important to have a high value of the signal-to-noise-ratio

of the relativistic effects. This work is also motivated by the fact that there are indeed wide area low redshift surveys planned or under way that could provide the necessary data for the galaxy populations (e.g. the planned EMU (Norris et al. 2011)) or the WISE data set (Bilicki et al. 2016)) to be interfaced with IM data provided by radio telescopes like LOFAR, Murchison Wide-field Array, GMRT, the Ooty Radio Telescope, CHIME, ASKAP, MeerKAT and SKA[2].

Throughout all the chapter we assume a spatially flat ΛCDM cosmology with parameters $h = 0.67556$, $\Omega_{\rm cdm}h^2 = 0.12038$, $\Omega_{\rm b}h^2 = 0.022032$. The primordial amplitude and spectral index are $A_{\rm s} = 2.215 \times 10^{-9}$ and $n_{\rm s} = 0.9619$, respectively. The matter power spectrum was computed with the Cosmic Linear Anisotropy Solving System (CLASS) code (Lesgourgues 2011; Blas et al. 2011), with pivot scale $k = 0.05\,{\rm Mpc}^{-1}$.

6.2 Cross-correlation odd multipoles

In this section we will report the expression for the observable quantities in a galaxy redshift survey and in a 21 cm intensity mapping experiment and the corresponding cross-correlation. We consider a perturbed Friedmann-Lemaître-Robertson-Walker (FLRW) metric and we work in longitudinal gauge[3]

$$ds^2 = a(\eta)^2 \left(-(1+2\Psi)\, d\eta^2 + (1-2\Phi)\, d\mathbf{x}^2 \right) , \tag{6.1}$$

where η denotes the conformal time, $a(\eta)$ is the scale factor, and the metric perturbations, Ψ and Φ, are the Bardeen potentials. We also remark that the equations summarized in this section do not assume General Relativity (GR).

6.2.1 Galaxy number counts

In a galaxy clustering experiment we measure the number of galaxies $N(\mathbf{n}, z)$ in terms of an angular direction $\mathbf{n}$,[4] and a measured redshift z. We can then define the galaxy number counts as

$$\Delta_{\rm gal}(\mathbf{n}, z) = \frac{N(\mathbf{n}, z) - \langle N \rangle (z)}{\langle N \rangle (z)}, \tag{6.2}$$

[2] http://www.lofar.org, http://www.mwatelescope.org, http://gmrt.ncra.tifr.res.in, http://rac.ncra.tifr.res.in/ort.html, http://chime.phas.ubc.ca/, https://www.atnf.csiro.au/projects/askap/index.html, http://www.ska.ac.za/gallery/meerkat/,https://www.skatelescope.org.

[3] Since we consider only observable quantities, the gauge choice will not affect any result.

[4] $\mathbf{n}$ denotes here the unit vector (direction) in which the photons propagate, while the angular position in the celestial sphere is $-\mathbf{n}$.

where $\langle..\rangle$ denotes the angular average at fixed observed redshift z. The galaxy number counts were computed to first order in perturbation theory (Yoo et al. 2009; Yoo 2010; Bonvin & Durrer 2011; Challinor & Lewis 2011), by accounting that galaxies are a biased tracer of the underline dark matter field, and they can schematically be expressed as (Di Dio et al. 2013)

$$\begin{aligned}\Delta_{\rm gal}(\mathbf{n},z,m^*) = {} & b_{\rm gal}(z,m^*)D + \frac{1}{\mathcal{H}(z)}\partial_r(\mathbf{V}\cdot\mathbf{n}) \\ & + (5s(m^*,z)-2)\int_0^{r(z)}\frac{r(z)-r}{2r(z)r}\Delta_\Omega(\Phi+\Psi)dr \\ & + \left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2}+\frac{2-5s(m^*,z)}{r\mathcal{H}}+5s(m^*,z)-f_{\rm evo}^{\rm gal}(m^*,z)\right)(\mathbf{V}\cdot\mathbf{n}) \\ & + (f_{\rm evo}^{\rm gal}-3)\mathcal{H}V + (5s-2)\Phi + \Psi + \frac{1}{\mathcal{H}}\dot{\Phi} + \frac{2-5s}{r(z)}\int_0^{r(z)}dr(\Phi+\Psi) \\ & + \left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2}+\frac{2-5s}{r(z)\mathcal{H}}+5s-f_{\rm evo}^{\rm gal}\right)\left(\Psi+\int_0^{r(z)}dr(\dot{\Phi}+\dot{\Psi})\right), \end{aligned} \tag{6.3}$$

The bias factors in the expression above, $s(m^*,z)$ and $f_{\rm evo}(m^*,z)$, are the magnification and the evolution biases of the galaxy catalogue, respectively. Modelling the magnification and the evolution biases of a galaxy catalogue requires a prior knowledge on the luminosity function of the targeted galaxy population.

6.2.2 21 cm brightness temperature fluctuation

Intensity mapping (Peterson et al. 2009) is a novel technique which aims to map the Large Scale Structure of the universe by measuring the collective emission of many galaxies without resolving individual sources. 21 cm IM experiments target the emission line of neutral atomic hydrogen. The observable quantity is the flux density, i.e. the integral of the specific intensity over the solid angle of the telescope beam, which can be related to the HI brightness temperature in the Rayleigh-Jeans regime (see the appendix in (Bull et al. 2015) for a more detailed discussion). The observed fluctuation in the 21 cm brightness temperature has been computed in linear theory in Ref. (Hall et al. 2013), including all the relativistic corrections. The full expression is mathematically equivalent to the expression in (6.3), with the magnification bias value set to $s = 2/5$, such that the lensing contribution vanishes. The fact that the observable in an intensity mapping survey is not affected by gravitational lensing to linear order is due to surface brightness conservation. Indeed the change in the

solid angle $d\Omega$ is exactly compensated by the change in the observed flux. The full expression for the observed fluctuation in the 21 cm brightness temperature is

$$\Delta_{21\mathrm{cm}}(\mathbf{n},z) = b_{\mathrm{HI}}(z)D + \frac{1}{\mathcal{H}(z)}\partial_r(\mathbf{V}\cdot\mathbf{n}) + \left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2} + 2 - f^{\mathrm{HI}}_{\mathrm{evo}}(z)\right)(\mathbf{V}\cdot\mathbf{n}) + (f^{\mathrm{HI}}_{\mathrm{evo}} - 3)\mathcal{H}V$$
$$+ \Psi + \frac{1}{\mathcal{H}}\dot{\Phi} + \left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2} + 2 - f^{\mathrm{HI}}_{\mathrm{evo}}\right)\left(\Psi + \int_0^{r(z)} dr(\dot{\Phi}+\dot{\Psi})\right), \tag{6.4}$$

where $b_{\mathrm{HI}}(z)$ and $f^{\mathrm{HI}}_{\mathrm{evo}}(z)$ are the bias and the evolution bias of the neutral hydrogen, respectively.

The evolution bias for the HI can be defined similarly to the galaxy evolution bias in (3.34). Taking into account that we observe all the HI emissions from a patch of the sky, it depends on the redshift evolution of the HI comoving density $\bar{\rho}_{\mathrm{HI}}$ (see the appendix in (Alonso & Ferreira 2015))

$$f^{\mathrm{HI}}_{\mathrm{evo}} = \frac{\partial \ln \bar{\rho}_{\mathrm{HI}}(z)}{\mathcal{H}\partial\eta} = -(1+z)\frac{\partial \ln \bar{\rho}_{\mathrm{HI}}(z)}{\partial z}. \tag{6.5}$$

6.2.3 21 cm - Galaxies cross-correlation

Cross-correlation studies are promising techniques to study relativistic effects. In fact, in the past years it has been pointed out that relativistic effects source odd multiples of the correlation function or Fourier space power spectrum, when two different tracers are cross-correlated (McDonald 2009; Yoo et al. 2012; Bonvin et al. 2014; Alonso & Ferreira 2015; Fonseca et al. 2015; Iršič et al. 2016; Bonvin et al. 2016; Gaztanaga et al. 2017; Borzyszkowski et al. 2017).

The two-point cross-correlation between the HI temperature and galaxy number count fluctuations, in terms of the observed coordinates, is defined as

$$\xi^{\mathrm{HI,gal}}(z_1,z_2,\theta) = \langle \Delta_{21\mathrm{cm}}(\mathbf{n}_1,z_1)\Delta_{\mathrm{gal}}(\mathbf{n}_2,z_2)\rangle\,, \qquad \cos\theta \equiv \mathbf{n_1}\cdot\mathbf{n_2}, \tag{6.6}$$

where the $\langle ...\rangle$ denotes the ensemble average, replaced in observation by the average over observed directions at a fixed observed redshift. In figure 6.1 we represent a scheme of the observed coordinates for the system under investigation.

The observed coordinates z_1, z_2 and θ can be converted, by assuming a cosmology, into comoving distances. We denote with r_1 and r_2 the comoving distance at z_1 and z_2, respectively, while $-\mathbf{n}_1$ and $-\mathbf{n}_2$ are the unit vectors pointing in the direction of the two tracers (HI and galaxies, respectively). Due to angular statistical isotropy, the cross-correlation can be the written in terms of three coordinates. In this work, we will adopt the following coordinate system (see figure 6.1 for

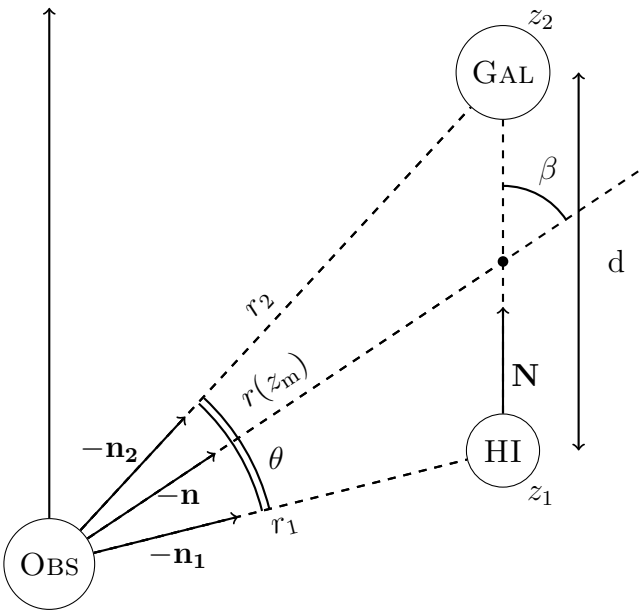


Figure 6.1: Illustration of the position of the two tracers under investigation with respect to the observer, in terms of the observed coordinates $\mathbf{z}_1$, $\mathbf{z}_2$ and θ and the coordinate system adopted in this work.

a schematic representation): the comoving separation r at the mean redshift $z_m = (z_1 + z_2)/2$, the separation between the two sources d and the angle β (or $\mu = \cos\beta$), where $\cos\beta = -\mathbf{n} \cdot \mathbf{N}$, while $-\mathbf{n}$ and $\mathbf{N}$ are the unit vectors pointing in the direction of the mean redshift and the distance between the two tracers, respectively, as defined in figure 6.1.

The cross-correlation function can be then expanded into Legendre polynomials L_ℓ,

$$\xi^{\mathrm{HI,gal}}(r, d, \mu) = \sum_\ell \xi_\ell(r, d) L_\ell(\mu), \tag{6.7}$$

and the coefficients of this expansion, the multipoles of the correlation function, are defined as

$$\xi_\ell(r, d) = \frac{2\ell + 1}{2} \int_{-1}^{1} \xi^{\mathrm{HI,gal}}(r, d, \mu) L_\ell(\mu) d\mu. \tag{6.8}$$

This definition does not include an optimal weight based on the galaxy number density, that is generally included in order to reduce the shot-noise (Ross et al. 2012). We do not include this effect for the following reasons. First, the weight is survey dependent, therefore a correct modelling of this effect will require to make some assumptions on the survey specifics. Furthermore, this weight is expected to reduce the noise, therefore our analysis without any assumption on the weight can be considered conservative.

In a similar way we can define[5] the Fourier cross power spectrum $P^{\rm HI,gal}(z_{\rm m}, \mathbf{k})$ as

$$\langle \Delta_{\rm 21cm}(\mathbf{k}_1, z_1)\Delta_{\rm gal}(\mathbf{k}_2, z_2)\rangle = (2\pi)^3 P^{\rm HI,gal}(z_{\rm m}, \mathbf{k}_1)\delta_D(\mathbf{k}_1 + \mathbf{k}_2), \tag{6.9}$$

where $\Delta_{\rm 21cm}(\mathbf{k}_1, z_1)$ and $\Delta_{\rm gal}(\mathbf{k}_2, z_2)$ denotes the Fourier transform of $\Delta_{\rm 21cm}(\mathbf{n}_1, z_1)$ and $\Delta_{\rm gal}(\mathbf{n}_2, z_2)$, respectively, and δ_D is the Dirac delta. The multipoles of the power spectrum can be computed similarly to the correlation function multipoles, and they are proportional to the linear matter power spectrum at the mean redshift, $P(k, z_{\rm m})$,

$$\langle D(\mathbf{k}_1, z_{\rm m})D(\mathbf{k}_2, z_{\rm m})\rangle = (2\pi)^3 P(k_1, z_{\rm m})\delta_D(\mathbf{k}_1 + \mathbf{k}_2). \tag{6.10}$$

In the distant observer limit, i.e. $d \ll r$, the angular position $-\mathbf{n}$ is assumed to be fixed for the two observed sources $\mathbf{n}_1 = \mathbf{n}_2 = \mathbf{n}$ and the full-sky correlation function can be simplified. Indeed, the full expression of the correlation function (Eq. (B.1) in appendix B.1) can be written as power series expansion in d/r, and by taking only the lowest order (i.e. assuming $d \ll r$) the multipoles of the correlation function can be expressed in terms of multipoles of the power spectrum (Hall & Bonvin 2017),

$$\xi_\ell(d, r(z_{\rm m})) = (-i)^\ell \int \frac{k^2 dk}{2\pi^2} P_\ell(k, z_{\rm m}) j_\ell(k\, d), \tag{6.11}$$

where

$$P_\ell(k, z_{\rm m}) = \frac{2\ell + 1}{2} \int_{-1}^{1} P^{\rm HI,gal}(k, z_{\rm m}, \mu) L_\ell(\mu) d\mu \tag{6.12}$$

are the coefficients of the expansion of the angle dependent power spectrum in Legendre polynomials, at fixed redshift z, and j_ℓ are the spherical Bessel function of order ℓ. Following the same strategy adopted in (Hall & Bonvin 2017), we consider a local expansion of the multipoles of the power spectrum in power of $(\mathcal{H}/k)$ and we include the leading terms with respect to the expansion parameter $(\mathcal{H}/k)$. The even multipoles are dominated by the Newtonian contribution (i.e. the first line in (6.3) and the first two terms in (6.4)), which are simply proportional to the matter power spectrum, whilst the odd multipoles are suppressed by a factor $(\mathcal{H}/k)$, which is provided by the correlation of Doppler contribution with density and redshift space distortions (i.e. the third line in Eq. (6.3) and the third term in Eq. (6.4)).

[5] We neglect the redshift evolution in the following definition. As shown in Ref. (Bonvin et al. 2014) the redshift evolution corrections are subdominant compared to the wide-angle dipole contamination of the standard terms. We also do not consider the integrated terms here. While time-delay and ISW effects are negligible, lensing magnification may contaminate the measurement of the Doppler dipole. We therefore study the contamination of magnification lensing in Sec. 6.2.4.

For the cross-correlation of galaxies and 21 cm brightness temperature, the leading contributions to the power spectrum multipoles are

$$
\begin{aligned}
P_0(k) =& \left[b_{\rm HI} b_{gal} + \frac{f}{3}(b_{\rm HI} + b_{\rm gal}) + \frac{f^2}{5} \right] P(k)\,, \\
P_1(k) =& (-i) \left[\Big(b_{\rm gal} C_{\rm HI} - b_{\rm HI} C_{\rm gal} \Big) f + \frac{3}{5} \Big(C_{\rm HI} - C_{\rm gal} \Big) f^2 \right] \frac{\mathcal{H}}{k} P(k)\,, \\
P_2(k) =& \left[\frac{2}{3} f (b_{\rm HI} + b_{\rm gal}) + \frac{4}{7} f^2 \right] P(k)\,, \\
P_3(k) =& \, i \frac{2}{5} \Big(C_{\rm gal} - C_{\rm HI} \Big) f^2 \frac{\mathcal{H}}{k} P(k)\,, \\
P_4(k) =& \frac{8}{35} f^2 P(k),
\end{aligned}
\tag{6.13}
$$

where $f = d\ln D / d\ln a$ is the growth factor. In ΛCDM the growth factor is given by $f(z) = \Omega_m(z)^{4/7}$ (Lahav et al. 1991; Carroll et al. 1992). In Eq. (6.13) the mean redshift is assumed to be fixed and, for the sake of simplicity, it is omitted from the notation. The coefficients $C_{\rm gal}$ and $C_{\rm HI}$ are defined as

$$
\begin{aligned}
C_{\rm gal} =& \left(\frac{\mathcal{H}'}{\mathcal{H}^2} + \frac{2 - 5s}{r\mathcal{H}} + 5s - f^{\rm gal}_{\rm evo} \right), \\
C_{\rm HI} =& \left(\frac{\mathcal{H}'}{\mathcal{H}^2} + 2 - f^{\rm HI}_{\rm evo}(z) \right).
\end{aligned}
\tag{6.14}
$$

In figure 6.2 we compare the cross-correlation dipole (top panels) and monopole (bottom panels) at two different mean redshifts $z_{\rm m} = 0.15$ (left panel) and $z_{\rm m} = 0.7$ (right panel). In this plot, we set the values of the clustering biases to be $b_{\rm gal} = 1$ and $b_{\rm HI} = 0.6$. The evolution bias for the HI is assumed to be $f^{\rm HI}_{\rm evo} = -1.5$, while we denote with different colors different values of the magnification bias s and with different line-styles two different values of the galaxy evolution bias. Interestingly, we remark that the sign of the dipole depends strongly on the magnification bias factor and, therefore, it can not be omitted in the analysis.

We see that both the dipole and the monopole signals decrease at larger redshift and that the monopole is significantly larger, in amplitude, than the dipole. Furthermore, at $z_{\rm m} = 0.7$ the terms depending on the galaxy evolution bias become dominant in the dipole. Since modelling the evolution properties of a galaxy population is not an easy task, the cosmological information we can extrapolate from the dipole at large redshift can be contaminated and limited from a prior knowledge about the evolution bias. The decrease of the dipole at larger

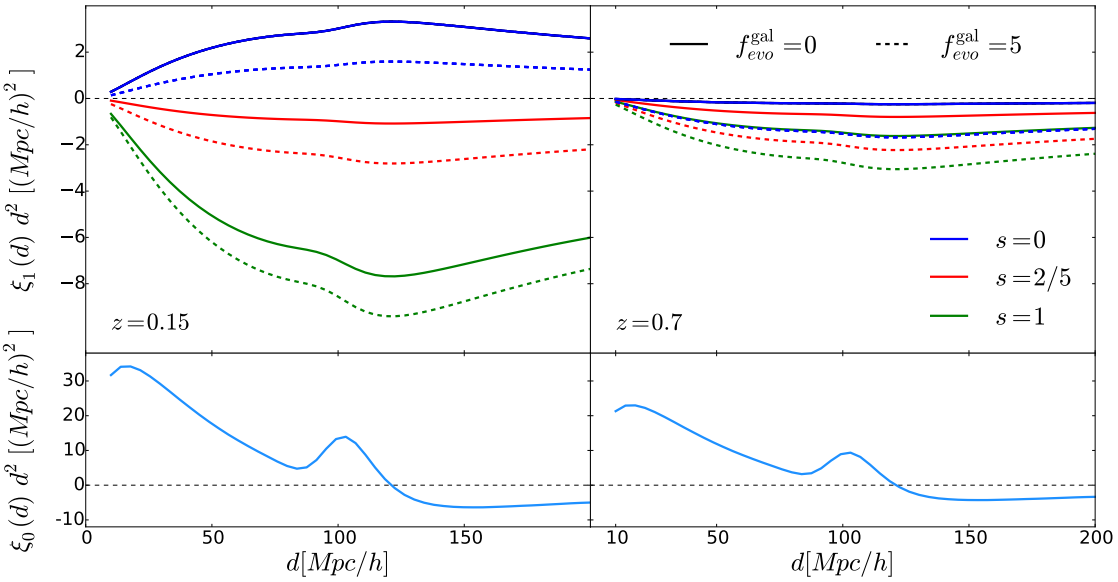


Figure 6.2: Cross-correlation dipole (top panels) and monopole (bottom panels), computed from Eq. (6.11), at $z = 0.15$ (left panel) and $z = 0.7$ (right panel). In the top panels different colors denotes different values for the magnification bias of the galaxy catalogue, whilst different line-styles refer to two different values of the galaxy evolution bias. The clustering biases are set to the values $b_{\rm gal} = 1$ and $b_{\rm HI} = 0.6$, the evolution bias for the HI is set to be $f_{\rm evo}^{\rm HI} = -1.5$.

redshift depends on two elements: the time evolution of the linear matter power spectrum (which affects the monopole as well) and the terms in the coefficients (6.14) proportional to $(r\mathcal{H})^{-1}$. Therefore, in order to detect relativistic effects we expect an ideal galaxy catalogue in the low redshift regime, and for this reason we will set the mean redshift of observation for our analysis to be $z_{\rm m} = 0.15$.

6.2.4 Contaminations to the relativistic dipole

The dipole of the cross-correlation that we discussed in the previous section is sourced by the Doppler corrections to the galaxies number counts and to the observed brightness temperature of the 21 cm emission. This is usually considered the main contribution to the dipole of the cross-correlation between two tracers, in fact the Doppler corrections depend on the projection of the peculiar velocity along the line of sight and therefore they are intrinsically anisotropic.

Nevertheless, a measurement of the dipole would be contaminated by other sources of anisotropy (Bonvin et al. 2014). In this section we will discuss two possible contaminations: the dipole induced by gravitational lensing and the wide-angle effects. The latter have been extensively studied in Refs. (Szalay et al. 1998; Szapudi 2004; Papai & Szapudi 2008; Matsubara 2000; Bharadwaj 1999; Raccanelli et al. 2010;

Samushia et al. 2012; Bertacca et al. 2012; Raccanelli et al. 2014, 2013). Beside these two contaminants, there are further corrections induced by the redshift evolution of the bias and the growth factors. They are generally subdominant with respect to the wide-angle correction (Bonvin et al. 2014), therefore they will be neglected.

The gravitational lensing asymmetry (Bonvin et al. 2014, 2017) comes from the cross-correlation of the HI density and the gravitational lensing term in the galaxy number counts

$$\xi^{\rm lens}(r,d,\beta) \equiv \left\langle (b_{\rm HI}(z_1)D\left(\mathbf{n}_1,z_1\right)) \left(\frac{5s-2}{2}\int_0^{r_2}\frac{r_2-r'}{r_2 r'}\Delta_\Omega(\Phi+\Psi)\left(\mathbf{n}_2,z'\right)dr'\right)\right\rangle, \tag{6.15}$$

and it emerges from the fact that galaxies behind an HI overdensity with respect to the observer will be lensed, while the HI temperature fluctuations are not lensed, to linear order, by the galaxies in front. The lensing correlation function defined above has been computed in (Bonvin et al. 2014) and further studied in (Bonvin et al. 2017). In the Limber approximation and to the lower order in d/r, it reads

$$\xi^{\rm lens}(r,d,\beta) = (1+z_{\rm m})\frac{3\Omega_m\pi}{4}b_{\rm HI}(5s-2)\,d\,\mathcal{H}_0\,\cos\left(\beta\right)\Theta(r_2-r_1)\mu_{\rm lens}(\beta), \tag{6.16}$$

where Θ is the Heaviside function, D_1 is the linear growth factor and the function $\mu_{\rm lens}$ is

$$\mu_{\rm lens}(\beta) = \int_0^\infty \frac{k_\perp dk_\perp}{2\pi^2}\mathcal{H}_0 P(k_\perp)J_0(k_\perp\, d\,\sin\left(\beta\right)), \tag{6.17}$$

being J_0 is the order-0 Bessel function. The lensing dipole can therefore be computed similarly to the relativistic dipole

$$\xi_1^{\rm lens}(r,d) = \frac{3}{2}\int_{-1}^{1}\xi^{\rm lens}(r,d,\mu)L_1(\mu)d\mu. \tag{6.18}$$

The lensing dipole is proportional to the radial distance between the two tracers and the redshift. Therefore, we expect it to be relevant in the high redshift regime, in particular because the Doppler dipole decreases with redshift. In figure 6.3 we show the amplitude of the lensing dipole compared to the dipole sourced by the Doppler terms, at redshift $z_{\rm m} = 0.15$. We see that the lensing dipole is always few orders of magnitude smaller then the Doppler dipole, in this redshift regime. Therefore, in the rest of the chapter it will be neglected.

The second correction we will discuss here is the wide-angle effect. In the limit $d \ll r$, we can expand the full-sky correlation function in power of d/r. Therefore the Taylor expansion of all the functions of r_1 and r_2 around r in the leading terms, namely density and redshift space distortions, induces a non-vanishing dipole

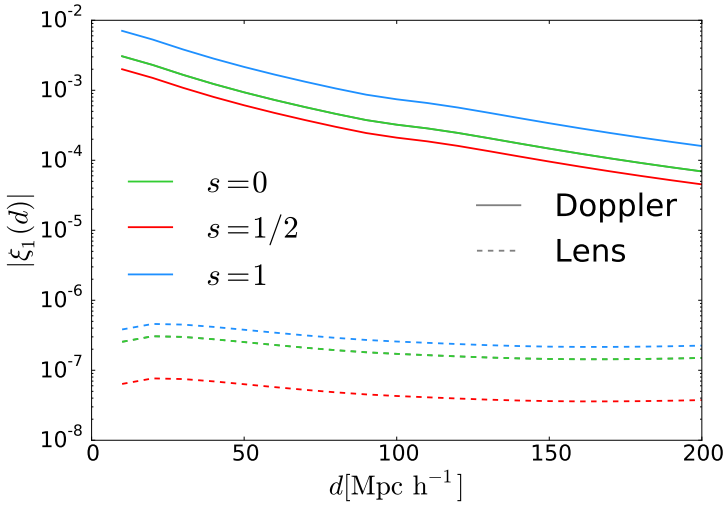


Figure 6.3: Comparison between the amplitude of lensing dipole (dashed lines) and the Doppler dipole (continuous lines) at $z_{\rm m} = 0.15$. The clustering bias of the two tracers are fixed to the values $b_{\rm HI} = 0.6$ and $b_{\rm gal} = 1$. Different colors denote different values of the magnification bias. The HI evolution bias is $f_{\rm HI}^{\rm evo} = -1.5$, while we fix the galaxy evolution bias to be zero.

suppressed by d/r. Considering that the largest contributions to the relativistic dipole at low redshift is of the order $d/r \sim 1/(kr) \sim \mathcal{H}/k$, we need to account for the wide-angle contamination in the dipole. The leak from the monopole to the dipole due to the wide-angle contribution, at fixed redshift and at the lowest order in d/r, is described by (Bonvin et al. 2014; Hall & Bonvin 2017)

$$\xi_1^{\rm WA}(d) = \frac{2f}{5}(b_{\rm gal} - b_{\rm HI})\frac{d}{r}\int \frac{k^2 dk}{2\pi^2} P(k) j_2(k\,d). \tag{6.19}$$

This correction can be written as a combination of the quadrupole of the auto-correlation of two tracers. Therefore, we correct the dipole estimator $\hat{\xi}_1$ for the bias due to the wide-angle effect (Hall & Bonvin 2017)

$$\hat{\xi}_1(d,r) \to \hat{\xi}_1(d,r) - \frac{3}{10}(\hat{\xi}_2^{gal} - \hat{\xi}_2^{\rm HI})\frac{d}{r}. \tag{6.20}$$

In figure 6.4 we show how the signal changes when the wide-angle correction in Eq. (6.20) is applied to the estimator. The magenta dotted line represent the wide-angle contribution to the dipole, computed from Eq. (6.19). Its magnitude is comparable to the one of the relativistic dipole, thus it is clearly a not negligible contribution. Furthermore we see that correcting for the wide-angle effect does not necessarily reduces the signal: if its sign agrees with the one of the dipole, the signal is boosted. This leads also to an extra contribution to the covariance of the

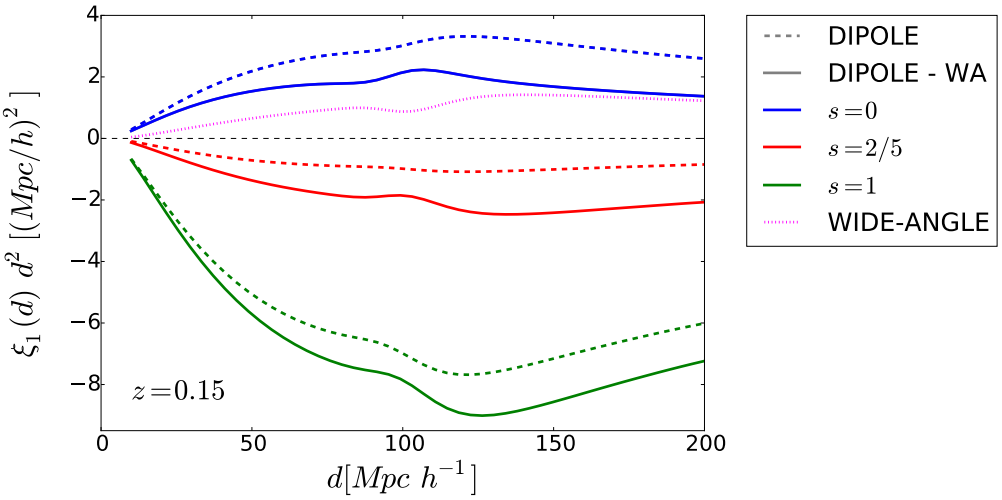


Figure 6.4: Cross-correlation dipole with (continuous line) and without (dashed line) the wide-angle correction in (6.20). Different colors denotes different values of the galaxy magnification bias. The clustering biases are set to be $b_{\rm gal} = 1$ and $b_{\rm HI} = 0.6$. The evolution bias is set $f_{\rm HI}^{\rm evo} = -1.5$ for the neutral hydrogen, while for galaxies it is set to be zero. The magenta dotted line represents the wide-angle correction in (6.19).

estimator (Hall & Bonvin 2017). In the next section we will show how to compute the full covariance of the estimator, the explicit contribution to the covariance of the wide-angle correction can be found in appendix B.2.

6.2.5 Covariance for the cross-correlation dipole

The full covariance matrix for the multipoles of the 2-point correlation function (2PCF) is presented in (Grieb et al. 2016) for the single tracer case and in (Hall & Bonvin 2017) for the multiple tracers case. Here we will apply the generic expression in (Hall & Bonvin 2017) (Eq. 17) to the dipole.

$$
\begin{aligned}
\mathrm{COV}(d_1,d_2) = &-\frac{9}{V}\int\frac{k^2dk}{2\pi^2}j_1(k\,d_1)j_1(k\,d_2)\left(\frac{1}{5}P_1^2(k)+\frac{8}{35}P_1(k)P_3(k)+\frac{23}{315}P_3^2(k)\right)\\
&+\frac{9}{V}\int\frac{k^2dk}{2\pi^2}j_1(k\,d_1)j_1(k\,d_2)\left[\frac{1}{3}\left(N_{\rm HI}P_0^{\rm gal}(k)+\frac{1}{n_{\rm gal}}P_0^{\rm HI}(k)\right)\right.\\
&\left.+\frac{2}{15}\left(N_{\rm HI}P_2^{\rm gal}(k)+\frac{1}{n_{\rm gal}}P_2^{\rm HI}(k)\right)\right]\\
&+\delta_{d_1,d_2}\frac{3}{4\pi V d_1 d_2 L_p}N_{\rm HI}\frac{1}{n_{\rm gal}}\\
&+\frac{9}{V}P_{\rm N}\int\frac{k^2dk}{2\pi^2}j_1(k\,d_1)j_1(k\,d_2)\left(\frac{1}{3}P_0^{\rm gal}(k)+\frac{2}{15}P_2^{\rm gal}(k)\right)\\
&+\delta_{d_1,d_2}\frac{3}{4\pi V d_1 d_2 L_p}P_{\rm N}\frac{1}{n_{\rm gal}},
\end{aligned}
\tag{6.21}
$$

where V is the overlapping volume of the galaxy and the intensity mapping surveys; $P_\ell^{\rm gal}$ and $P_\ell^{\rm HI}$ denotes the galaxy and the HI power spectrum multipoles, respectively; $n_{\rm gal}$ is the comoving number density of galaxies; $N_{\rm HI}$ and $P_{\rm N}$ are the shot-noise and the interferometer noise for the HI, respectively; $L_{\rm p}$ denotes the resolution of the IM survey and $\delta_{x,y}$ is the Kronecker delta.

The first line represents the purely cosmic variance contribution to the covariance, the second and the third lines are the cosmic variance - Poisson noise terms, the fourth line is the purely Poissonian contribution, while the last two lines are the interferometer noise - cosmic variance term and the interferometer noise - galaxy Poisson noise term, respectively. The terms, which do not involve the integral of the power spectrum, have been integrated by using the orthogonality relation of the spherical Bessel functions

$$
\int_0^\infty dk k^2 j_1(kd_1)j_1(kd_2) = \frac{\pi}{2d_1d_2}\delta_D(d_1-d_2),
\tag{6.22}
$$

where in the discrete limit (Hall & Bonvin 2017) $\delta_D(d_1-d_2)\to\delta_{d_1,d_2}/L_p$. The other integrals have been solved numerically, and a smooth cutoff is applied in all the integrals to model the finite resolution of the interferometer. To be more precise, all the integrands are multiplied by a top-hat filter in Fourier space $W^4(kR)$, defined as

$$
W(kR) = \frac{3[\sin(kR)-kR\cos(kR)]}{(kR)^3},
\tag{6.23}
$$

where the scale R is set to be the size of the pixel for the IM, $L_{\rm p}$.

Figure 6.5 shows the different terms contributing to the diagonal covariance entries, at redshift $z_{\rm m}=0.15$. The bias values for the two tracers are the same

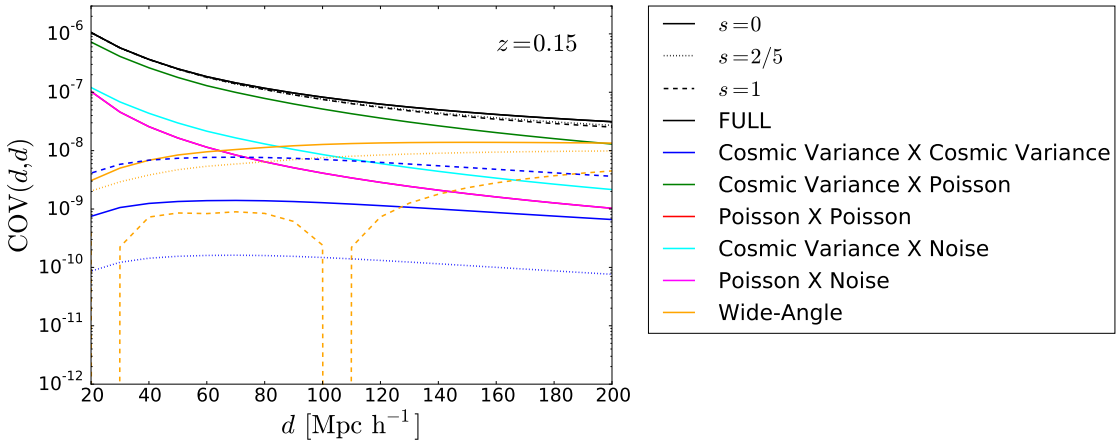


Figure 6.5: Diagonal entries of the covariance, computed from (6.21). Different colors denote different terms contributing to the covariance. The full covariance is the black lines. Different line-styles represent different values of the magnification bias, which affects only the cosmic variance X cosmic variance contribution (blue line).

adopted in figure 6.4. Furthermore, we set the shot-noise and the interferometer noise for the HI to $N_{\rm HI} = 100\,({\rm Mpc/h})^3$ and $P_{\rm N} = 100\,({\rm Mpc/h})^3$, respectively. The comoving galaxy number density is assumed to be $n_{\rm gal} = 10^{-3}\,({\rm h/Mpc})^3$ and the volume of the survey is computed assuming that the sky coverage of the cross-correlation is $f_{\rm sky} = 0.2$ and that the redshift bin is $z \in [0.05, 0.25]$. The size of the pixel is chosen to be $L_{\rm p} = 2\,{\rm Mpc/h}$. At small scales, the dominant components of the covariance are the terms in which the cosmic variance is cross-correlated with the noise. On large scales, the covariance introduced by the wide-angle correction becomes more important and on scales $\sim 190\,{\rm Mpc}/h$ or larger is the dominant contributor (see the appendix B.2 for details on how to estimate the covariance introduced by the wide-angle correction). The magnification bias enters only in the computation of the purely cosmic variance term, but since this term result to be subdominant it does not affect significantly the full covariance (at least its diagonal components).

Both the interferometer noise $P_{\rm N}$ and the size of the pixel $L_{\rm p}$ depend on the specifics of the IM survey (see (Bull et al. 2015; Hall & Bonvin 2017)). In order to stay as general as possible, we choose fiducial values ($P_{\rm N} = 100\,({\rm Mpc}/h)^3$ and $L_{\rm p} = 2\,{\rm Mpc}/h$, respectively). They are approximately the values of a survey similar to the Canadian Hydrogen Intensity Mapping Experiment (CHIME) (Bandura et al. 2014), extrapolated to low redshifts ($z \sim 0.1 - 0.2$).

6.3 Halo Model approach

As shown in section 6.2, the signal and the covariance for the dipole depend on many different parameters (clustering and evolution biases, galaxy magnification bias, shot-noise of the two tracers, sky fraction of the cross-correlation and redshift range). In this section we aim to model the properties of the two tracers in order to find a relation between these parameters, which are not all independent. This will allow us to find the optimal specification to look for relativistic effects detection in LSS. Moreover, by determining the relation between all these different parameters we can parameterize better their uncertainties if we need to marginalize over them. In order to model a generic galaxy population and the neutral hydrogen distribution we will adopt the so-called halo model (see (Cooray & Sheth 2002) for a general review). The halo model was first proposed for modelling the galaxies properties (Seljak 2000; Peacock & Smith 2000; Scoccimarro et al. 2001), but it has been more recently successfully applied to neutral hydrogen (Villaescusa-Navarro et al. 2014; Castorina & Villaescusa-Navarro 2017; Padmanabhan & Refregier 2017; Padmanabhan et al. 2017). With this approach we are able to model the relation between the tracers clustering bias and their shot-noises.

6.3.1 HI model

We adopt for the neutral hydrogen the model based on (Castorina & Villaescusa-Navarro 2017). As supported by numerical simulations (Villaescusa-Navarro et al. 2014), we can safely assume that the contribution of neutral hydrogen outside the dark matter halos is negligible. Hence, the neutral hydrogen comoving density $\bar{\rho}_{\rm HI}$ at a given redshift z can be computed as

$$\bar{\rho}_{\rm HI}(z) = \int_0^\infty n(M,z) M_{\rm HI}(M,z) dM, \tag{6.24}$$

where $n(M,z)$ is the halo mass function at redshift z, i.e. the comoving number density of halos with masses in the range between M and $M+dM$, $M_{\rm HI}(M,z)$ is the average HI mass in a halo of mass M at redshift z. The halo mass function can be expressed as

$$n(M,z) = -\frac{\bar{\rho}_m}{M^2} f(\sigma) \frac{d\ln\sigma}{d\ln M}, \tag{6.25}$$

where $\bar{\rho}_m = \Omega_m(z)\,\rho_c^0$, ρ_c^0 is the critical density at $z=0$, σ is the root mean square of the variance of the linear density field, smoothed on the scale $R(M)$, the radius enclosing an amount of mass equal to M

$$\sigma^2(R,z) = \frac{1}{2\pi^2}\int_0^\infty k^2 P(k,z) W^2(kR) dk, \qquad R = \left(\frac{3M}{4\pi\bar{\rho}_m}\right)^{\frac{1}{3}}. \tag{6.26}$$

The smoothing function W is the Fourier transform of a top-hat filter (the same functional form of (6.23)). The function $f(\sigma)$ is generally calibrated from N-body simulation. In this work we use a Tinker mass function (Tinker et al. 2008). Within this framework, the shot-noise and the HI bias can be written as (Bull et al. 2015; Castorina & Villaescusa-Navarro 2017)

$$N_{\rm HI}(z) = \left(\frac{1}{\bar{\rho}_{\rm HI}(z)}\right)^2 \int_0^\infty n(M,z) M_{\rm HI}^2(z) dM, \tag{6.27}$$

$$b_{\rm HI}(z) = \frac{1}{\bar{\rho}_{\rm HI}(z)} \int_0^\infty n(M,z) b(M,z) M_{\rm HI}(z) dM, \tag{6.28}$$

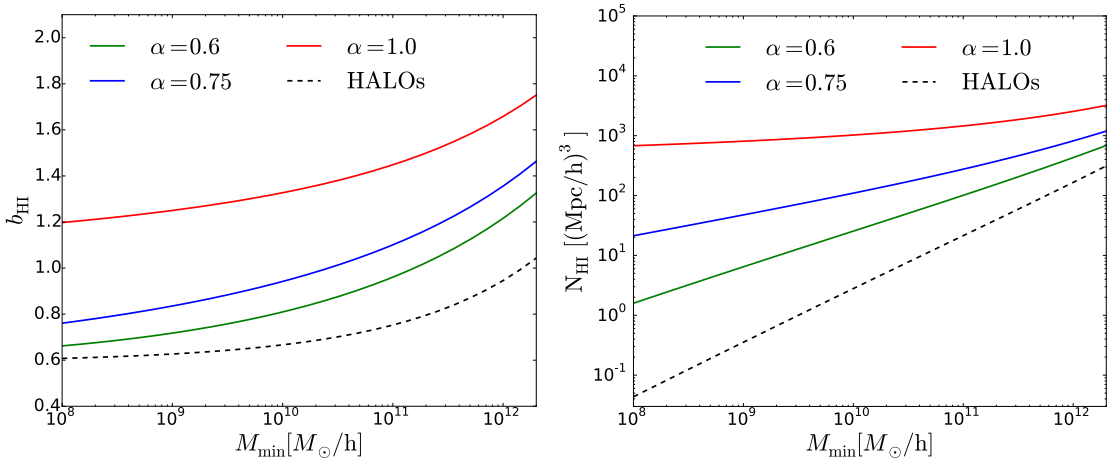
where $b(M,z)$ is the halo bias, calibrated on the N-body simulation from (Tinker et al. 2010).

We model the average HI mass within an halo of mass M as redshift independent (Castorina & Villaescusa-Navarro 2017)

$$M_{\rm HI}(M,z) = {\rm C}\ (1-Y_{\rm p})\frac{\Omega_{\rm b}}{\Omega_{\rm m}} \exp\Big[-(M_{\rm min}/M)\Big] M^\alpha, \tag{6.29}$$

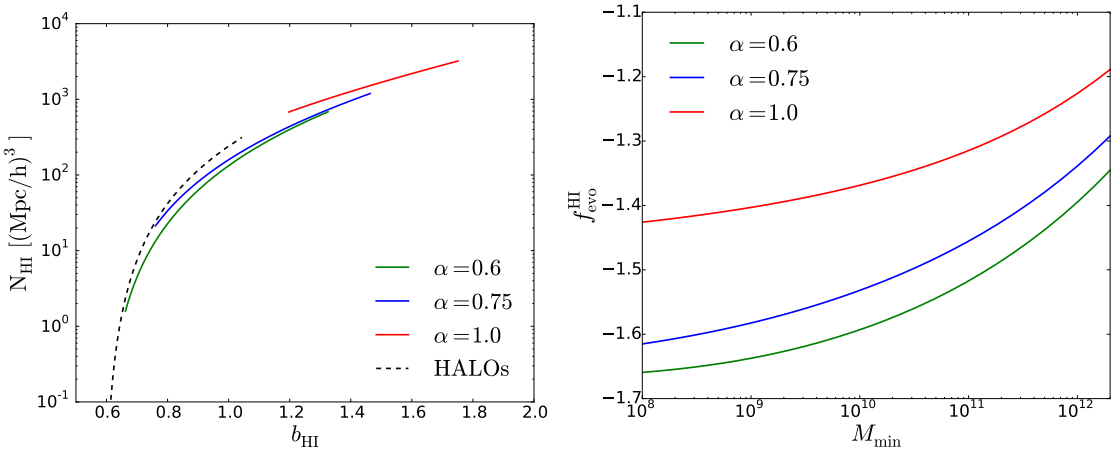
where $Y_{\rm p} = 0.24$ is the Helium fraction (note that $(1-Y_{\rm p})\frac{\Omega_{\rm b}}{\Omega_{\rm m}}$ is the HI mass fraction), $M_{\rm min}$, α and C are the free parameters of the model. Eq. (6.29) entails that the mass of cosmic hydrogen within a halo scales as a power law of the total virial mass of the halo, at large halo masses. The efficiency of this scaling is regulated by the exponent α (larger values of α correspond to larger amount of hydrogen within a halo of fixed mass). At low halo masses, we expect instead an exponential suppression, due to different physical processes, such as photoionization from the UV background or galactic winds. The parameter $M_{\rm min}$ regulates the range of halo masses for which this suppression is effective (larger values of $M_{\rm min}$ imply a larger range of masses for which the suppression is relevant). The parameter C is an overall normalization constant, that needs to be fixed by matching the theoretical abundance of neutral hydrogen, predicted by (6.24), with what is measured by HI galaxy survey at a given redshift. As we can notice from Eqs. (6.27) and (6.28), both the shot-noise and the HI bias do not depend on this normalization factor, therefore we can fix C = 1 without loss of generality.

In figure 6.6 we show the HI bias and HI shot-noise, respectively, as a function of the cutoff parameter $M_{\rm min}$, for different values of α. We can see that if the neutral hydrogen is concentrated in halos with large mass (this corresponds to larger values for $M_{\rm min}$ and α), both the shot-noise and the HI bias have larger values with respect to the case in which we find considerable amount of neutral



(a) HI bias **(b)** HI shot-noise

Figure 6.6: HI bias (left panel) and HI shot-noise (right panel) as a function of the cutoff parameter $M_{\rm min}$, at redshift $z = 0.15$. Different colors denote different values of the exponent α in (6.29). The black dashed lines represent the halo bias and the shot-noise of the halo population.



(a) Shot noise VS bias - HI **(b)** Evolution bias - HI

Figure 6.7: Left panel: Shot-noise of the HI as a function of the HI bias. Right panel: Evolution bias as a function of the cutoff parameter $M_{\rm min}$. The notation is the same as figure 6.6.

hydrogen also in small halos. The halo bias (in black dashed lines) is computed from Eq. (15) in (Marulli et al. 2011)

$$b_{\rm halos}(z) = \frac{\int_0^\infty n(M,z)b(M,z)dM}{\int_0^\infty n(M,z)dM}, \tag{6.30}$$

while the shot-noise for the halo population is computed as

$$N_{\rm halos} = \frac{1}{\int_0^\infty n(M,z)dM}. \tag{6.31}$$

From figure 6.6 we can find a relation between the bias of HI, $b_{\rm HI}$, and the shot-noise, $N_{\rm HI}$. In figure (6.7a) we show how the HI bias and the HI shot-noise are related to each others, for different values of the parameter α, by considering $M_{\rm min}$ as the parametrisation of the different curves. As expected, in all models higher biased HI models correspond to higher shot-noise. Furthermore, models with an higher efficiency in accreting neutral hydrogen within halos (larger values of α) correspond to higher values of bias and shot-noise. According to hydrodynamics simulations (Villaescusa-Navarro et al. 2015, 2016) and semi-analytic models (Kim et al. 2017), values of $\alpha < 0.9$ are more realistic.

In figure 6.7b we show the dependence of the evolution bias, computed from (6.5), for different models. The evolution bias for the HI does not vary significantly within the parameter space, because it is described by the redshift evolution of the halo mass function. Its negative value indicates that the density of HI drops at low redshift.

We note that the model outlined in this section is more focussed on the high redshift regime and needs to be better tested and improved in the low redshift Universe in order to capture the complex astrophysical effects (Evoli et al. 2011; Mancuso et al. 2017) that could impact on the HI distribution inside galaxies.

6.3.2 Galaxy model

As discussed in the previous section for neutral hydrogen, we can use the halo model to relate the bias parameters and the shot-noise of galaxy distributions. In this framework, we assume that all the galaxies are found within dark matter halos. The relevant quantity we need to model is their comoving number density, which can be computed from the halo mass function as

$$n_{\rm g}(z) = \int_0^\infty n(M,z)N_{\rm av}(M,z)dM, \tag{6.32}$$

where $N_{\rm av}$ is the average number of galaxies for an halo of mass M. The galaxy bias can be modeled as

$$b_{\rm g}(z) = \frac{1}{n_{\rm g}(z)}\int_0^\infty n(M,z)b(M,z)N_{\rm av}(M,z)dM, \tag{6.33}$$

while the shot-noise is simply given by the inverse of the number density, i.e.

$$N_{\rm gal} \equiv \frac{1}{n_{\rm g}} = \frac{1}{\int_0^\infty n(M,z) N_{\rm av}(M,z)\, dM}. \tag{6.34}$$

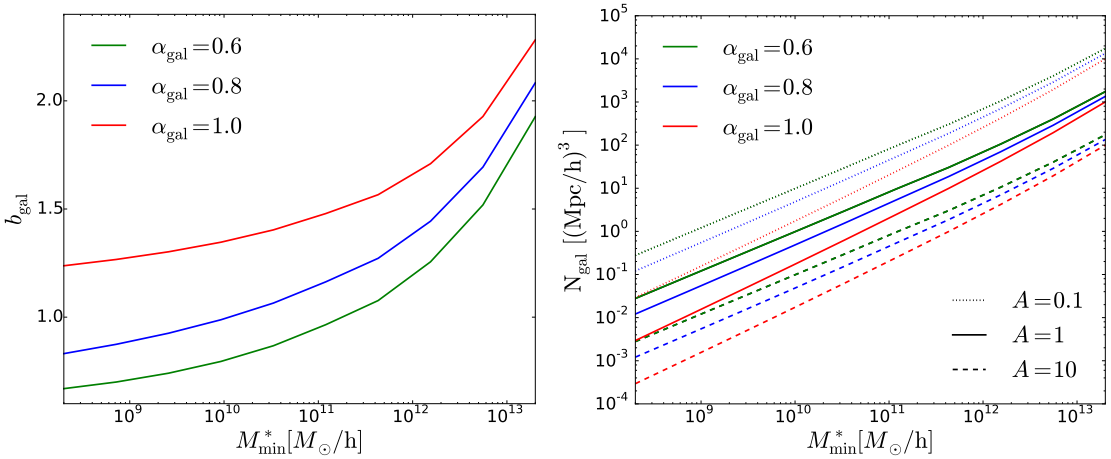
We model the average number of galaxies within a halo of mass M as

$$N_{\rm av}(M) = \begin{cases} 0 & \text{if } M \leq M^*_{\rm min} \\ A\left(\frac{M}{M^*_{\rm min}}\right)^{\alpha_{\rm gal}} & \text{if } M > M^*_{\rm min} \end{cases} \tag{6.35}$$

This model is similar to the model employed in (Cooray & Sheth 2002) to model red and blue galaxies, which provides a good fit to the number of subhalos expected from numerical simulations (White et al. 2001). The parameter $M^*_{\rm min}$ represents a threshold mass, below which a halo cannot host a galaxy, because their potential wells are shallower with respect to more massive halos. Therefore, some physical processes such as supernova feedback, can be efficient enough to expel a huge percent of baryons from the halo and therefore suppress the star formation within the halo itself. For halos with masses larger than this threshold, we assume that the average number of galaxies per halo increases with the halo mass, following a power law $N_{\rm av}(M) \propto M^{\alpha_{\rm gal}}$, so that $\alpha_{\rm gal}$ represents its slope. The parameter A represents a normalization constant. From Eq. (6.33) we see that, as we found for the HI, the galaxy bias is not affected by the value of the normalization, while the shot-noise, Eq. (6.34), strongly depends on A. In the literature more flexible and physically motivated models have been proposed (see e.g. (Zheng et al. 2005; Yang et al. 2012)). Nevertheless, the simple model described above captures the features that we need for the purpose of our analysis with the minimal number of parameters.

In figure (6.8) we show the galaxy bias (left panel) and galaxy shot-noise (right panel) as a function of the threshold mass $M^*_{\rm min}$. The bias increases exponentially with the threshold mass, while the shot-noise increase as a power law whose slope depends the slope of the model (6.35).

In figure (6.9) we summarize the information of figure (6.8) by displaying the relation between galaxy bias and shot-noise, for different models. Similarly to HI, highly biased galaxy populations (which are more massive) show a larger shot-noise, compared to the one with lower bias.



(a) Galaxy bias **(b)** Galaxy shot-noise

Figure 6.8: Galaxy bias (left panel) and galaxy shot-noise (right panel) as a function of the cutoff parameter $M^*_{\rm min}$, at redshift $z = 0.15$. Different colors denote different values of the exponent $\alpha_{\rm gal}$ in (6.29). In the right panel, different line-styles denote different values of the normalization constant A.

6.4 Signal-to-noise analysis

In this section we will present a signal-to-noise analysis for the dipole of the HI-galaxies cross-correlation. The signal-to-noise for the dipole is defined as (Hall & Bonvin 2017)

$$\left(\frac{S}{N}\right)^2 = \sum_{d_1,d_2=d_{\rm min}+L_{\rm p}/2}^{d_{\rm max}-L_{\rm p}/2} \xi_1(d_1){\rm COV}^{-1}(d_1,d_2)\xi_1(d_2)\,. \tag{6.36}$$

The minimum distance $d_{\rm min}$ is chosen to be the non-linear scale. We set $d_{\rm min} = 30\,{\rm Mpc}/h$, which correspond to $k_{\rm NL} \approx 0.2\,h/{\rm Mpc}$. We also study the dependence of the signal-to-noise ratio on both $d_{\rm min}$ and $d_{\rm max}$. The maximum distance is set to be the distance at which the dipole estimator, computed to lowest order in the d/r expansion from (6.20), coincides with the full-sky quantity up to 3%. Since we are interested in relatively low redshift measurements, we provide a comparison between the wide-angle correction to linear order in d/r and the exact full-sky contributions. The details of this comparison are shown in the appendix B.1. Our results show that for $d \leq 200\,{\rm Mpc}/h$ the discrepancy between the two quantities is smaller than the 3% threshold, therefore we set $d_{\rm max} = 200\,{\rm Mpc}/h$. Although more numerically expensive, it is possible to consider the full-sky expression if further accuracy is required. Nevertheless, as shown in figure 6.10, the main information is encoded on scales where the expansion to linear order in d/r is accurate enough.

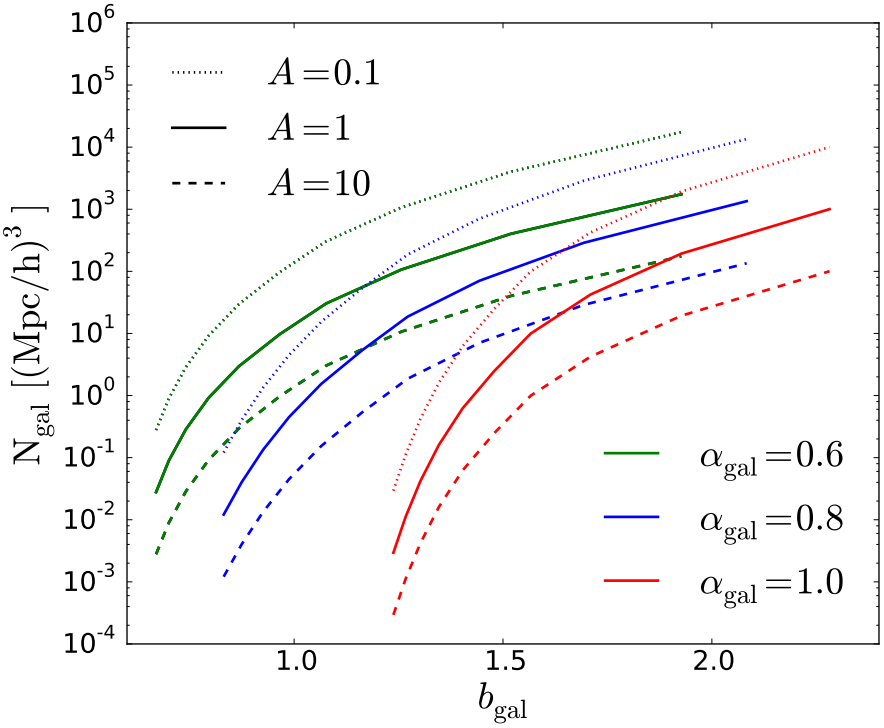


Figure 6.9: Shot-noise of the galaxies as a function of the galaxy bias, for different values of the slope $\alpha_{\rm gal}$ and different values of the normalization constant A.

The cross-correlation volume V is computed as

$$V = \frac{4}{3}\pi\Big(r^3(z_{\rm max}) - r^3(z_{\rm min})\Big) f_{\rm sky}, \tag{6.37}$$

where $r(z)$ is the comoving distance at redshift z, $z_{\rm max}$ and $z_{\rm min}$ are the maximum and minimum observed redshift, respectively, and $f_{\rm sky}$ is the fraction of the sky available for the observation. We consider a redshift range between $z_{\rm min} = 0.05$ and $z_{\rm max} = 0.25$ and a fraction of sky $f_{\rm sky} = 0.2$, which is the sky coverage of a survey similar to the Baryon Oscillation Spectroscopic Survey (BOSS) (Alam et al. 2017b). We assumed the interferometer employed for observing the neutral hydrogen in intensity mapping can resolve pixels of size $L_{\rm p} = 2\,{\rm Mpc}/h$.

We will consider two models for the neutral hydrogen, based on the formalism presented in section 6.3.1. The first model (conservative model) is similar to the one employed in (Villaescusa-Navarro et al. 2017), with bias close to unity at low redshift and relatively high shot-noise. The second model (optimistic model) is similar to the one adopted for the forecast in (Bull et al. 2015), with smaller bias and lower shot-noise. The specific values of the parameters of the two models are highlighted in Table 6.1.

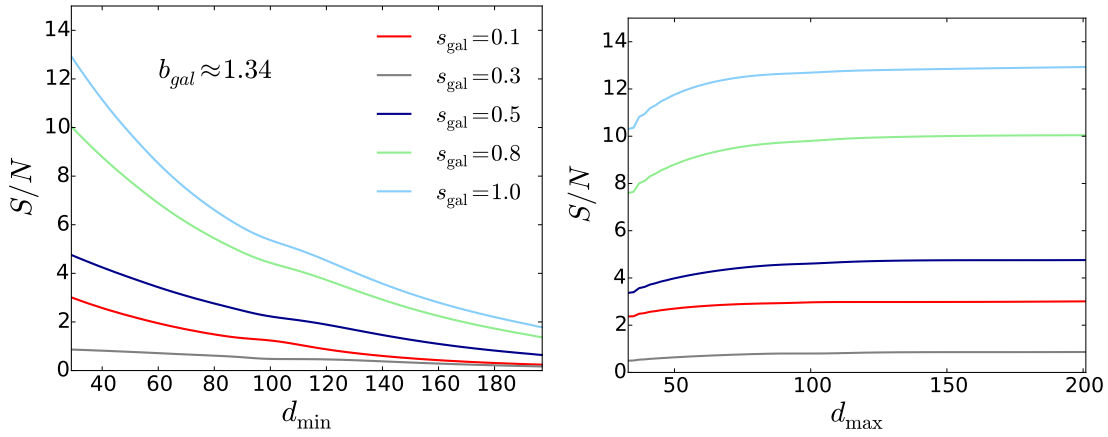


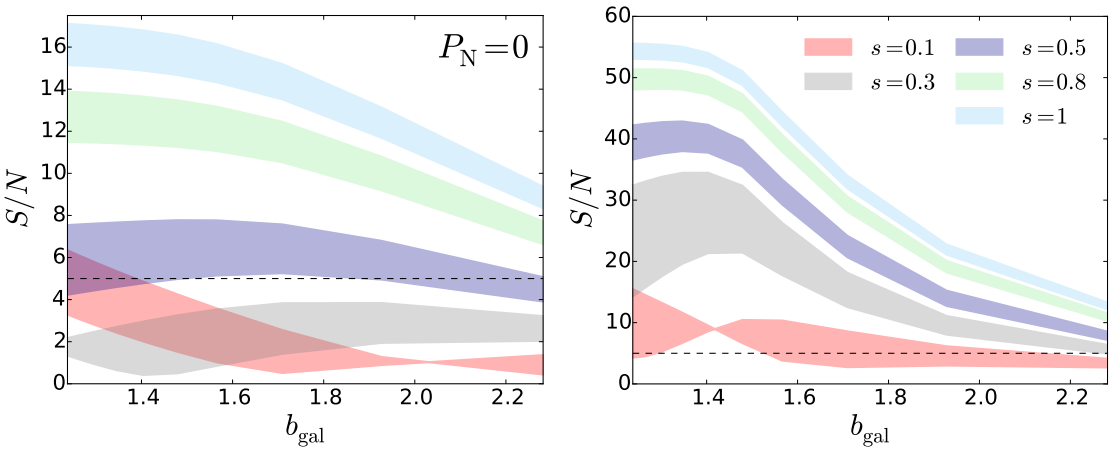
Figure 6.10: Signal-to-noise as a function of $d_{\rm min}$ and $d_{\rm max}$. The HI model is the model 1 in Table 6.1. The galaxy bias is fixed to be $b_{\rm gal} \approx 1.34$ here, and the interferometer noise is set to be $P_{\rm N} = 100\,({\rm Mpc}/h)^3$. Different colors denotes different values for the magnification bias.

	MODEL 1 (CONSERVATIVE)	MODEL 2 (OPTIMISTIC)
α	0.75	0.6
$M_{\rm min}$	$\approx 1.7 \times 10^{10} M_\odot/{\rm h}$	$10^8 M_\odot/{\rm h}$
$b_{\rm HI}$	≈ 0.99	≈ 0.67
$N_{\rm HI}$	$\approx 143{\rm Mpc}^3/{\rm h}^3$	$\approx 2{\rm Mpc}^3/{\rm h}^3$
$f_{\rm evo}^{\rm HI}$	≈ -1.51	≈ -1.66

Table 6.1: Values of the parameters for the two HI models we consider for our analysis. Model 1 has larger bias and larger shot-noise with respect to Model 2.

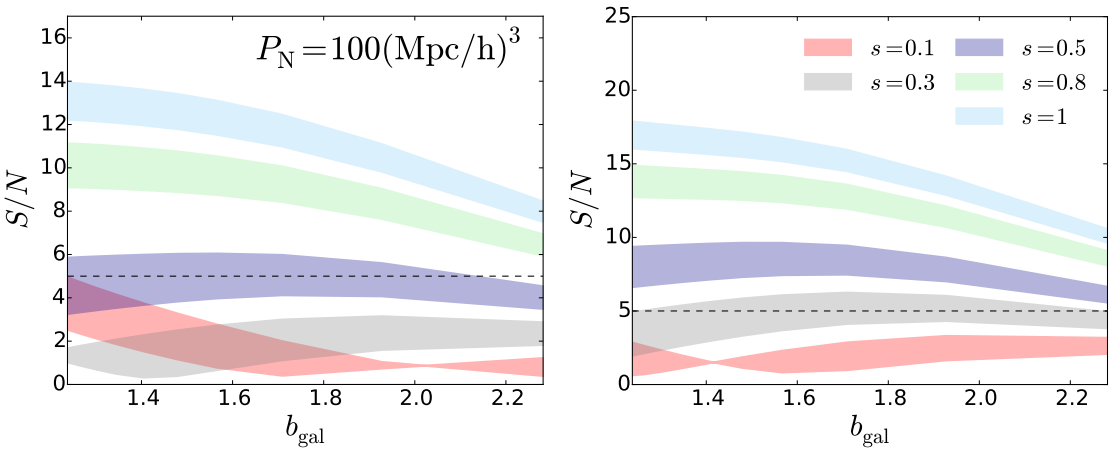
Concerning the galaxy model, the parameter A is a normalization constant and does not qualitatively affect the relations between shot-noise and bias. Therefore, in the rest of the chapter we will fix its value to be $A = 1$, which correspond to the minimum value of the average number of galaxies per halos whose masses are above the threshold value $M^*_{\rm min}$. Furthermore, we fix the value of $\alpha_{\rm gal} = 1$, which corresponds to the most biased galaxies.

Figure 6.11 shows the dipole signal-to-noise at $z = 0.15$, as a function of the galaxy bias, for the two models described above. The signal-to-noise is computed for different values of the magnification bias, which is treated here as a free parameter. We find that the signal-to-noise is generally optimized for the largest value of the magnification bias (which in this case is $s_{\rm gal} = 1$). Let us remark that the signal-to-noise ratio does not grow monotonically with the magnification bias, but it has a minimum between the lines $s = 0.1$ and $s = 0.3$ of Fig. (6.11). Indeed as shown in Fig. (6.4), the largest pre-factor of the dipole in flat-sky is $(2 - 5s)/(r\mathcal{H})$. Then, considering wide-angle effects and the impact of the other bias factors (in



(a) MODEL 1 (CONSERVATIVE) **(b)** MODEL 2 (OPTIMISTIC)

Figure 6.11: Signal-to-noise for the cross-correlation dipole as a function of the galaxy bias. The slope of the model for the galaxy population is fixed at the value $\alpha_{gal} = 1$. Different colors denote different values of the magnification bias. The corresponding shaded region represents the signal-to-noise for $f_{evo}^{gal} \in [-2, 2]$. The interferometer noise is neglected here.



(a) MODEL 1 (CONSERVATIVE) **(b)** MODEL 2 (OPTIMISTIC)

Figure 6.12: Signal-to-noise for the cross-correlation dipole as a function of the galaxy bias, with an interferometer noise $P_N = 100\,(\mathrm{Mpc/h})^3$. The same conventions of figure 6.11 are adopted. The black dashed line highlight the $S/N = 5$ threshold.

particular the large negative value of the evolution bias of neutral hydrogen), the dipole vanishes for lower values of magnification bias.

In the optimistic model for the HI (figure 6.11b), the signal-to-noise is > 5 for all the values of the magnification bias. We see that it increases with the bias up to $b_{gal} \approx 1.4$, then it declines for more biased galaxies. This is due to the fact that the dipole signal increases for larger difference between the biases of the two tracers, but highly biased tracers are more massive and we observe fewer of them.

Therefore for the most biased galaxies the growth of the shot-noise dominates over growth of the signal and the S/N results to be suppressed.

Similar comments are valid for the more conservative HI model. In this case we need $s_{\rm gal} \geq 0.5$ in order to have a $S/N > 5$. The maximum value of the signal-to-noise in the two models ranges between $(S/N)_{\rm max} \approx 17$, for the more conservative model, to $(S/N)_{\rm max} \approx 53$ in the optimistic model.

In figure 6.12 we show how the interferometer noise affects the signal-to-noise analysis for the two models. We assume the fiducial value $P_{\rm N} = 100\,({\rm Mpc}/h)^3$, which corresponds to the noise of a CHIME-like survey (Bandura et al. 2014) observing at the mean redshift considered in our work, computed from Refs. (Hall & Bonvin 2017; Bull et al. 2015). For both models the signal-to-noise is suppressed and the effect results to be more prominent for the optimistic model. We see that the interferometer noise suppresses the maximum value of the signal-to-noise ratio roughly from 17 to 14 for the conservative model, while the maximum S/N decreases from 53 to 18 in the optimistic model. Indeed, in the optimistic model the HI shot-noise results to be much smaller than the reference interferometer noise, while in the conservative model the two quantities are comparable. Interestingly, we observe that, even when the interferometer noise is included in the analysis, the signal-to-noise results to be > 5 for $s_{\rm gal} \geq 0.5$ in both models we considered for the HI.

6.5 HOD approach for modelling Luminosity-threshold galaxy catalogues

The galaxy model described in section 6.3.2, and applied in the signal-to-noise analysis in section 6.4, is intuitive and easy to implement. Nevertheless, its limitations are many. In fact, galaxy surveys can observe sources with luminosity larger that a threshold value. Within the framework described in 6.3.2 , it is not possible to model the luminosity function of a sample of galaxies, and therefore both magnification and evolution bias were treated as free parameters. In this section we will partially trade the generality of this framework with a more realistic model for a galaxy catalogue.

In order to model a luminosity-threshold galaxy catalogue, we will assume the model based on the Halo Occupation Distribution (HOD) described in (Zheng et al. 2007; Zehavi et al. 2011). The average number of galaxies, within a halo of mass M and with apparent magnitude below a certain threshold value m^* is modelled

as a sum of two contributions: the contributions from the central galaxies and the one from the satellite galaxies (Kravtsov et al. 2004; Zheng et al. 2005)

$$N_{\rm AV}(<m^*, M) = N_{\rm cent}(<m^*, M) + N_{\rm sat}(<m^*, M). \tag{6.38}$$

The central galaxies are modeled as a step function

$$N_{\rm cent}(<m^*, M) = \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log M - \log M_{\rm min}(m^*)}{\sigma(m^*)}\right)\right], \tag{6.39}$$

while the satellite galaxies are a product of the same step function and a power law

$$N_{\rm sat}(<m^*, M) = N_{\rm cent}(<m^*, M) \times \left(\frac{M - M_0(m^*)}{M_1(m^*)}\right)^{\alpha(m^*)}. \tag{6.40}$$

The model described above involves 5 parameters. $M_{\rm min}$ is the halo mass such that the average number of central galaxies with luminosity above the cut luminosity is $1/2$, σ regulates the efficiency at which the number of galaxies increases from the small to the large halos regime, M_0 is the cutoff mass scale for the satellite galaxies, M_1 is a normalization factor and α is the slope of the power-law that determine the number of galaxies in the highly massive halos regime. These parameters are not independent: they all depend on the threshold magnitude of the considered catalogue. In Ref. (Zehavi et al. 2011) the parameters have been computed for the SDSS galaxy catalogue by considering samples with different luminosity thresholds. In this section we will assume a galaxy sample that follows the same behavior. Nevertheless, the functional form in the equations (6.38), (6.39) and (6.40) can be in principle applied to other galaxy catalogues. Figure 6.13 (left panel) represents the average number of galaxies as a function of the halo mass, for galaxy samples for different values of the maximum absolute magnitude M_r [6].

By assuming this behaviour for $N_{\rm AV}$, we can use the same framework defined in section 6.3.2 to relate the shot-noise, the clustering bias and the magnification bias to the magnitude threshold, at fixed redshift. In fact, we can compute the comoving number density of galaxies as

$$\bar{n}_{\rm gal}(<M_{\rm r}) = \int_0^\infty n(M, z) N_{\rm AV}(<M_{\rm r}, M) dM. \tag{6.41}$$

[6]The absolute magnitude is related to the apparent magnitude m^* by $M_r = m^* - 5\log_{10}\frac{d_L(z)}{d_L(z_{\rm ref})} - K(z)$, where $d_L(z)$ is the luminosity distance at a given redshift and $K(z)$ is the k-correction, which corrects the measured magnitude into the one that would be measured in the source's rest frame.

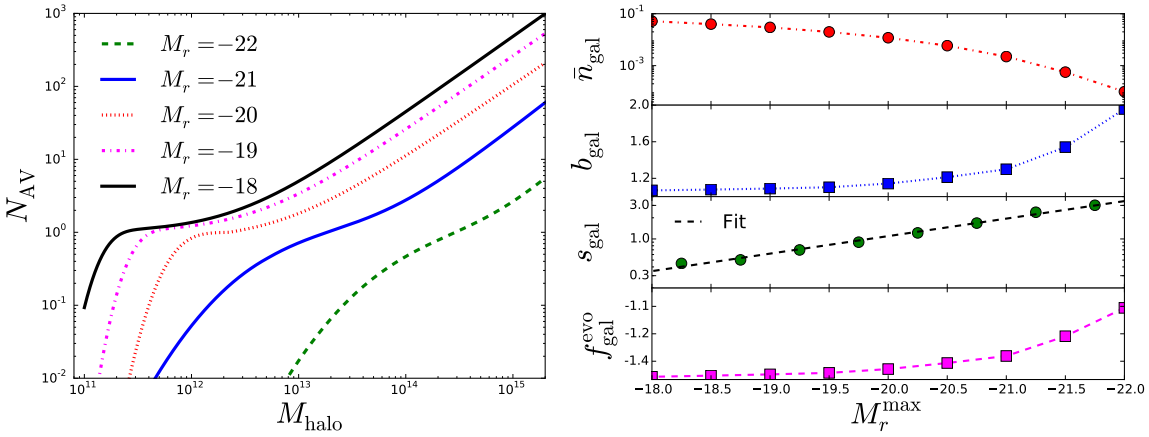


Figure 6.13: Left panel: Average number of galaxies within a halo as a function of the halo mass. Different lines denote galaxies samples with different values of the maximum absolute magnitude M_r. Right panel: Comoving number density (red), clustering bias (blue), magnification bias (green) and evolution bias (magenta) as a function of the maximum absolute magnitude M_r of the selected galaxy sample.

The shot-noise is the inverse of $\bar{n}_{\rm gal}(< M_{\rm r})$, while the galaxy and the magnification biases are

$$b_{\rm g}(M_{\rm r}) = \frac{1}{n_{\rm g}(z)} \int_0^\infty n(M,z) b(M,z) N_{\rm AV}(< m^*, M) dM, \tag{6.42}$$

$$s_{\rm gal}(M_{\rm r}) = \frac{\partial \log_{10}(\bar{n}_{\rm gal}(< M_{\rm r}))}{\partial M_{\rm r}}. \tag{6.43}$$

Note that since we fixed the redshift, the derivative with respect to the apparent magnitude coincides with the one respect the absolute magnitude $M_{\rm r}$. The evolution bias can be directly computed from (3.35). In figure 6.13 (right panel) we plot $\bar{n}_{\rm gal}(< M_{\rm r})$, $b_{\rm g}$ and $s_{\rm gal}$ as a function of the maximum magnitude of the sample. We see that samples with a smaller value of the magnitude threshold correspond to higher value of the biases and higher shot-noise (lower comoving density). Furthermore, we observe that the magnification bias increases for smaller values of the magnitude cut $M_{\rm r}$, reaching values up to ≈ 3. The evolution bias is also larger for smaller values of $M_{\rm r}$, but its value are relatively small in all the range of magnitude threshold.

Once we have the behaviour of the biases and of the shot-noise in terms of the magnitude cut, we can compute the correspondent signal to-noise for the dipole from (6.36). The result is plotted in figure 6.14, where different colors and line-styles refer to different values of the interferometer noise. We assumed the model 1 (the conservative one) in Table 6.1 for the HI, while all the galaxies properties shown in figure 6.13. We found that the signal-to-noise is optimized for a certain value

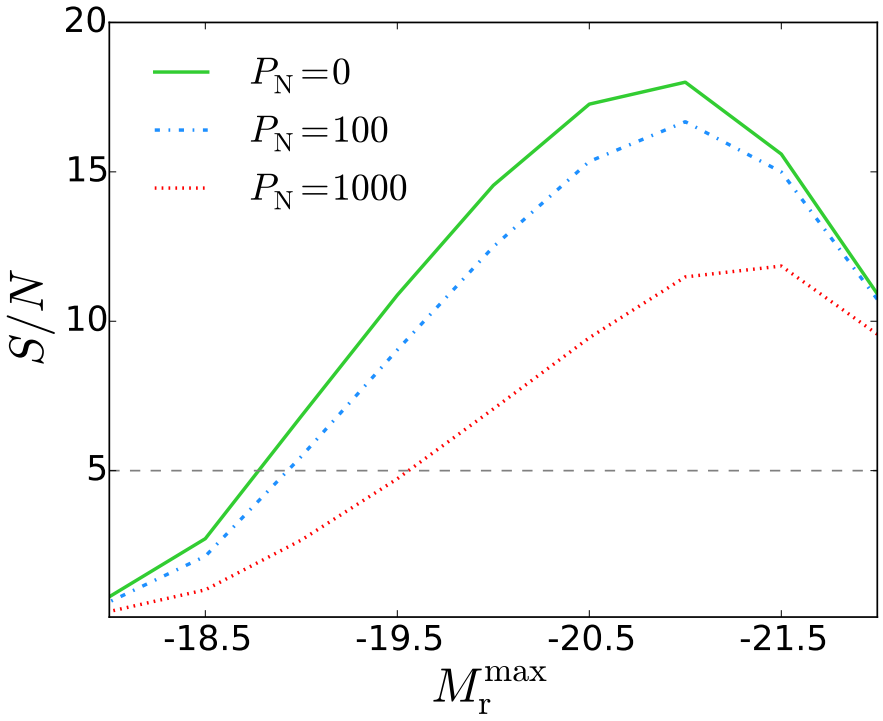


Figure 6.14: Signal-to-noise as a function of the magnitude threshold of the galaxy survey. The galaxies are described by the HOD model discussed in this section, while the model for the HI is the more conservative model described in the previous section (see Table 6.1). Different colors and line-styles denote different values of the interferometer noise (the unit is Mpc^3/h^3). The dashed horizontal line denotes a detection threshold of $S/N = 5$.

of the limiting magnitude, which is not its maximum values (which corresponds to the minimum shot-noise), because the magnification bias, and therefore the dipole signal, has larger values if the maximum magnitude threshold is set to be smaller. Up to $M_\mathrm{r} \approx -21$ the signal increases faster than the noise, while for $M_\mathrm{r} < -21$ the shot-noise dominates over the signal growth rate. This result implies that, even if for a galaxy catalogue the expected signal-to-noise for the dipole is below the detection threshold (in figure 6.14 we set $S/N \geq 5$ for a possible detection), it is possible to properly choose a smaller limiting magnitude for the sample and reject the galaxies with magnitude above this value, in order to amplify the S/N above the detection threshold.

6.6 Summary

In this chapter we studied the relativistic dipole in the cross-correlation function of HI intensity mapping and galaxies. We model the HI and galaxy parameters

(clustering bias, magnification and evolution bias and shot-noise) that affects the signal in the general framework of the halo model.

We present a signal-to-noise analysis for the relativistic dipole sourced by the Doppler effect of the cross-correlation between the galaxy number counts and the HI brightness temperature. Our analysis aims to study the properties of a galaxy population that optimize the signal-to-noise ratio. We consider two HI models and in both cases we find that the signal-to-noise is suppressed for highly biased tracers, in despite that the dipole is proportional to the bias difference of the tracers. Therefore galaxies with lower bias are preferred for detecting the Doppler dipole in the cross-correlation of HI intensity mapping and galaxies. Furthermore the signal appears to be considerably larger for higher values of the magnification bias, therefore galaxy surveys with steeper luminosity function are favored: we find that in order to have a $S/N > 5$ we need roughly $s \geq 0.5$.

Finally, we perform a similar signal-to-noise analysis to a luminosity threshold galaxy catalogue. We describe the number of galaxies with magnitude below a given threshold using a model based on the Halo Occupation Distribution. We study the relation between the parameters of the galaxy population and the limiting magnitude. In particular, we investigate the signal-to-noise of the dipole as a function the magnitude threshold. Our results, see figure 6.14, indicate that the maximum value for the limiting magnitude of the survey does not correspond to the higher signal-to-noise ratio. Indeed, we show how to select an optimal value for the magnitude threshold to maximize the signal-to-noise. This analysis depends on the specific Halo Occupation Distribution model that we employed, which was built to fit the SDSS data. Nevertheless, analogue methods can be applied to other galaxy catalogues and can be useful for selecting the optimal galaxy sample to measure the relativistic dipole.

We conclude with some words of caution. First of all both the modelling of the HI and the galaxy populations are based on suitable, physically motivated extensions of the halo model that are expected to be reliable. However, the modelling of HI inside galaxies in the low redshift Universe is crucial and it is not fully explored in this work how observational results of the HI content of galaxies impact on the halo model parameters. In this respect we notice that attempts of modelling the IM signal by incorporating these physical effects into the simple parametric model have been made (Villaescusa-Navarro et al. 2016). Secondly, here we neglect any modelling of the foreground signal of the IM that we know dominates by several orders of magnitude. Foreground removal techniques are of primary importance in order to fully exploit the cross correlation signal and preliminary results are

encouraging (e.g. (Alonso et al. 2015)) also for the IM-galaxy cross-correlation (Villaescusa-Navarro et al. 2015). The relatively high values of the signal-to-noise ratio of the effect under study is however suggesting that it could still be detected once these caveats are properly modelled.

Part III

LSS in the non-linear regime

Chapter 7

THE GENERATION OF VORTICITY IN COSMOLOGICAL LARGE SCALE STRUCTURE

In part II we focused on the linear regime of the large scale structures. In the present part we will focus on a highly non-linear phenomenon: we will investigate the generation of vorticity in large scale structure.

7.1 Introduction and motivation

It is well known that most galaxies rotate and that the angular velocities of neighboring galaxies are correlated. It has recently been argued that observations of radio galaxies indicate that these correlations extend also to substantially larger scales, up to 10-20 Mpc (Taylor & Jagannathan 2016). While we only observe baryons, in most physical dark matter models one expects that similar correlations exist between the respective host dark matter halos. Within linear perturbation theory, cold dark matter (CDM) is usually modelled as a pressureless perfect fluid. However, since vorticity cannot be generated in a pressureless perfect fluid that is initially free of vorticity, and since within standard cosmology the initial vorticity is negligible, this vortical motion must be modelled with theories going beyond the perfect fluid description of the cosmological matter.

The evolution of small (linear) perturbations in the early Universe is such that the fluid rotational velocity Ω decays like $1/a$ in a matter dominated Universe (Durrer 2008), where a denotes the cosmic scale factor. Even though the amplitude of Ω remains constant in a radiation dominated Universe, it must be very small initially since it is of the order of $\Omega \simeq (k/\mathcal{H})^2\sigma^{(V)}$, where $\sigma^{(V)}$ is the helicity-1 (vortical) contribution to the shear, which can be at most of the order of 10^{-4} after inflation and the factor $(k/\mathcal{H})$ is very small after inflation for all cosmologically relevant scales. A rough order of magnitude for the value of the Hubble parameter after

reheating is $\mathcal{H}_{\rm RH} \sim 10^{10}(T_{\rm RH}/1\ {\rm TeV}){\rm Mpc}^{-1}$. For a cosmological scale of 1 Mpc this yields $(k/\mathcal{H}_{\rm RH})^2 \lesssim 10^{-20}$, hence this possible perturbative vorticity can safely be neglected. This is under the assumption that vorticity is generated by some non-standard inflationary model. Most inflationary models actually do not generate helicity-1 modes to begin with, so that $\sigma^{(V)} = 0$ after inflation.

Furthermore, according to Helmholtz's theorem, a pressureless perfect fluid that is initially irrotational remains irrotational in the absence of external rotational forces also beyond perturbation theory (Helmholtz 1858)[1]. This theorem is also non-perturbatively valid in General Relativity (Lu et al. 2009). In order to generate vorticity, a non-vanishing velocity dispersion is required. Assuming velocity dispersion to be present in dark matter at some given level, the induced vorticity has been determined perturbatively in (Cusin et al. 2017). In this approach, velocity dispersion is an intrinsic property of the dark matter fluid, while the velocity dispersion driven by the shell-crossing of cold dark matter particles cannot be encoded in this framework. Furthermore, it has the drawback to require the introduction of a free parameter for the velocity dispersion which regulates the amplitude of the vorticity spectrum. Several studies of shell-crossing in the context of perturbation theory have been recently presented (see (Rampf & Frisch 2017; McDonald & Vlah 2018; Pietroni 2018; Saga et al. 2018) and references therein). However, the application of these techniques to the full 3D problem is still lacking.

The alternative to a perturbative approach is to numerically investigate the generation of vorticity. Here we pursue this approach. The dark matter phase space distribution obeys the Vlasov-Poisson system (Bernardeau et al. 2002). Therefore, a full numerical study would require solving the full system for the six-dimensional phase space. From a computational point of view, this is a very challenging task. There have been several attempts in the literature to solve the Vlasov-Poisson system for the full phase space distribution function (see (Sousbie & Colombi 2016; Widrow & Kaiser 1993; Kopp et al. 2017; Mocz et al. 2018) and references therein). Nevertheless, nowadays the most common approach is to simplify the problem by resorting to the N-body method, which samples the phase space distribution function at some discrete locations corresponding to the particle positions and velocities. N-body simulations are an extremely powerful tool to study the large-scale structure of the Universe in the non-linear regime, capturing very complex astrophysical phenomena (Efstathiou et al. 1985; Frenk et al. 2000; Springel et al. 2005, 2006;

[1]Helmholtz's original work assumed incompressibility for the fluid. However, his law was generalized beyond this approximation by other authors. The validity of the theorem for a barotropic fluid was proved by Hankel (see (Villone & Rampf 2017; Frisch et al. 2017) for an english translation of his work).

Boylan-Kolchin et al. 2009; Shandarin et al. 2012; Springel 2010; Springel et al. 2018) and including gravitational effects beyond the Newtonian approximation (Adamek et al. 2014, 2016a,b, 2017). In cosmological N-body simulations, structure formation leads to shell (orbit) crossing, i.e. particles at (nearly) the same position can have very different velocities. This induces significant velocity dispersion which in turn sources vorticity. The fact that N-body simulations have a finite mass and spatial resolution induces additional (numerical) velocity dispersion and it is crucial to carefully check whether we can distinguish the latter numerical artifact from the former physical phenomenon.

A first numerical study of the vorticity field with N-body simulations was presented in (Pueblas & Scoccimarro 2009), where the vorticity power spectrum has been computed in simulations with different mass resolutions, thus raising the issue of the resolution dependence of the vorticity spectrum in N-body simulations. The vorticity power spectrum was computed from N-body simulations also in (Thomas et al. 2015) and in (Hahn et al. 2015), where it was also studied in the context of the warm dark matter scenario.

The N-body simulations presented in this chapter are performed with the publicly available relativistic code *gevolution* (Adamek et al. 2016a,b)[2]. Here the relativistic aspect of the code is less important than the fact that we work with a fixed grid, which also determines our spatial resolution. In the previous studies the spatial resolution was instead provided by the softening length, a numerical artifact used in the computation of the forces to prevent divergences. Even though the results of the previous studies qualitatively agree, (Pueblas & Scoccimarro 2009) and (Thomas et al. 2015) find e.g. a different time evolution for the vorticity at relatively high redshift ($z > 1$). Apart from the different N-body code, in this chapter we use a different method to compute the vorticity and we compare our results with these previous studies. The goal of this work is to improve our understanding of the vorticity generation process in N-body simulations and its spectral properties, which is crucial in order to properly model the effect of vector perturbations in redshift-space clustering analysis (Zhu et al. 2017b; Bonvin et al. 2018) and which may be useful for understanding the intrinsic alignments of galaxies (Troxel & Ishak 2014).

The remainder of this chapter is organized as follows. In section 7.2 we detail our method, define the relevant quantities and we describe the numerical calculations and the tests we perform. In section 7.3 we present our results and compare them with previous studies. In section 7.4 we sum up and give a conclusion. In two

[2] https://github.com/gevolution-code.

appendices we study resolution effects and compare our method with the Delaunay tessellation method used in (Pueblas & Scoccimarro 2009).

Notation: We use the metric in the Poisson gauge $ds^2 = a^2[-(1+2\Psi)dt^2 - 2B_i dx^i dt + (1-2\Phi)\delta_{ij}dx^i dx^j]$, hence t denotes conformal time. Here $B^i_{,i} = 0$, hence B_i denotes a vector perturbation, 'frame dragging' and we neglect tensor perturbations (gravitational waves). While 4d indices are denoted by greek letters and raised and lowered with the metric $g_{\mu\nu}$, spatial, 3d indices are denoted by latin letters. They are raised and lowered with the flat 3d metric δ_{ij} and hence 3d index positions are irrelevant. The conformal Hubble parameter is denoted by $\mathcal{H} = \dot{a}/a = aH$, where an overdot denotes the derivative with respect to conformal time t. Spatial indices are denoted by Latin letters, and spacetime indices by Greek letters. Spatial vectors are denoted in bold face. In the figures, the velocity divergence and vorticity are normalized in such a way that the linear velocity divergence power spectrum matches the linear matter density power spectrum.

7.2 Method

In this section we will describe the numerical implementation of the vorticity field in the N-body code *gevolution* [3].

The velocity field can be split into a gradient and the rotational part such that

$$\mathbf{v} = -\nabla v + \mathbf{v}_R \equiv \mathbf{v}_G + \mathbf{v}_R \ , \text{ with} \qquad \nabla \cdot \mathbf{v}_R = 0\,. \tag{7.1}$$

We define the divergence θ and the vorticity $\boldsymbol{\omega}$ by

$$\theta = \nabla \cdot \mathbf{v} = -\Delta v\,, \tag{7.2}$$

$$\boldsymbol{\omega} = \nabla \wedge \mathbf{v} = \nabla \wedge \mathbf{v}_R. \tag{7.3}$$

The implementation of the vorticity is performed in two steps: the computation of the velocity field in real space that is detailed in section 7.2.1, and the projection of the velocity in its divergence and rotational part which is described in section 7.2.2.

The linear power spectrum displayed in our plot for comparison and the linear transfer functions needed for the initial conditions of our simulations are computed with the Cosmic Linear Anisotropy Solving System (CLASS) code (Lesgourgues 2011; Blas et al. 2011). We assume a spatially flat ΛCDM cosmology compatible with the Planck 2015 results (Ade et al. 2016a): $h = 0.67556$, $\Omega_{\rm cdm}h^2 = 0.12038$, $\Omega_{\rm b}h^2 = 0.022032$. The primordial amplitude and spectral index are $A_{\rm s} = 2.215\times10^{-9}$

[3] https://github.com/gevolution-code

and $n_\mathrm{s} = 0.9619$, respectively, at the pivot scale $k_* = 0.05\,\mathrm{Mpc}^{-1}$. Note that the initial velocity field in linear perturbation theory is a pure gradient, and the particles are assigned the velocities only after their initial displacement $x_i \to x_i' = x_i + \delta x_i(\mathbf{x})$, that is, the velocity field is evaluated at the displaced positions, ensuring that initially we have $\boldsymbol{\omega} = 0$. Due to finite numerical precision and CIC projection effects, however, some spurious vorticity is still generated initially. The latter is resolution dependent, and in section 7.3.3 we study the convergence of the obtained vorticity power spectrum with respect to the resolution of the simulation.

7.2.1 Computation of the velocity field

In *gevolution*, particles are evolved using the geodesic equation and their energy momentum tensor is then calculated by a particle-to-mesh projection, see (Adamek et al. 2016a,b) for details. The stress-energy tensor components are related to the density and the momentum density as follows:

$$T^0_0(\mathbf{x}) = -\rho(\mathbf{x}), \qquad T^i_0(\mathbf{x}) = -\rho(\mathbf{x}) v^i(\mathbf{x}), \qquad T^i_j(\mathbf{x}) = \rho(\mathbf{x})\, v^i(\mathbf{x}) v_j(\mathbf{x}). \tag{7.4}$$

We define the velocity field simply by

$$v^i(\mathbf{x}) = \frac{T^i_0(\mathbf{x})}{T^0_0(\mathbf{x})}. \tag{7.5}$$

The particle-to-mesh projection for both the T^0_0 and T^i_0 components adopts the Cloud-In-Cell (CIC) method (Hockney & Eastwood 1988). Details about the implementation of the CIC method in *gevolution* can be found in appendix B of Adamek et al. (2016b) or in the monograph on N-body simulations (Hockney & Eastwood 1988). We use the CIC projection also for the momentum density, in order to obtain a velocity that is effectively a mass-weighted average in each cell. In our simulations, vector metric perturbations are included in the equations of motion, while in the computation of the velocity field we neglect them.

The definition (7.5) is singular if the cells surrounding the grid point $\mathbf{x}$ are empty and thus $\rho(\mathbf{x}) = 0$, i.e. inside voids. This is of course a consequence of the finite mass resolution of our N-body simulations and it raises the well-known issue of characterizing the velocity field from a discrete, and not uniformly distributed, sample of points. In the literature there are several ways to approach this problem. In (Bertschinger et al. 1990) a smoothing is employed to extract the velocity field from a galaxy redshift survey[4]. In this approach, the discrete velocity field

[4]The application to a galaxy catalogue presents the further complication that the galaxies' velocity vectors are known only in the line-of-sight direction.

is convolved with a kernel with non-compact support. Therefore, the smoothed density field is never identically zero. Another approach which has been successfully applied to N-body simulations, consists of building a tessellation from the irregular distribution of particles (Bernardeau & van de Weygaert 1996). In particular, the Delaunay tessellation has been used to extract the velocity and vorticity field from N-body simulations in (Pueblas & Scoccimarro 2009; Thomas et al. 2015). A refined method, which employs a phase-space interpolation technique, was presented in (Abel et al. 2012; Hahn et al. 2015). This method appears to perform better than the tessellation method, especially on small scales. However, both the phase space interpolation and the tessellation results are difficult to incorporate in a pre-built N-body code, and they usually require the post-processing of snapshot of a simulation. We tested four alternative methods that allow us to compute the velocity field on the fly and, therefore, reduce the computational cost and the memory requirement.

The first three methods estimate the velocity field from Eq. (7.5) in the domain of the simulation where the density is not zero, while in the voids we invoke the following three mechanisms to assign velocities:

1. We set $\mathbf{v} = 0$ to the grid points belonging to an empty cell. This is the simplest prescription, however it neglects the expected outflows of the particles from low to high density regions. We call this the *zero* method.

2. In the initial conditions all grid points have a finite mass density, $\rho(\mathbf{x}, t_{\rm in}) \neq 0$. When $\rho(\mathbf{x}, t_i) = 0$ in some grid point $\mathbf{x}$, we set $v(x, t_i) = v(x, t_{i-1})$. According to this prescription, the velocity field is frozen when a void region forms in the simulation. We call this the *past* method.

3. When $\rho(\mathbf{x}, t_i) = 0$ in some grid point $\mathbf{x}$, we set $v(\mathbf{x}, t_i) = \dot{D}_1(t_i)/\dot{D}_1(t_{i-1})v(\mathbf{x}, t_{i-1})$, i.e. we evolve the velocity field with its linear growth factor. Here D_1 is the linear density growth factor and the continuity equation implies $\theta = -\dot{\delta}$. The logarithmic growth rate is defined as

$$f = \mathcal{H}^{-1}\dot{D}_1/D_1 \quad \text{hence } \theta = -f\mathcal{H}\delta \tag{7.6}$$

within linear perturbation theory. We call this method the '*rescaled* method'.

4. The fourth method consists of a Gaussian smoothing for both the momentum density and the density field. The velocity field is then estimated as the ratio of the two smoothed fields. This method requires the introduction of a typical scale which sets the size of the smoothing. We also note that a Gaussian

smoothing of the fields in real space is not practical, as this would completely destroy the parallelizability of the code, which relies on the fact that each processor needs to know the particle positions only in a small region of the grid (and its *halo* (Adamek et al. 2016b)). We therefore smooth in Fourier space, where the real space convolution becomes a simple multiplication. This requires two additional Fourier transforms of T_0^0 and T_0^i, but only at the time steps at which we want to determine the velocity power spectrum.

In section 7.3 we shall show that the first method somewhat overestimates the vorticity power spectrum, while the three other methods agree to a good approximation, provided the smoothing scale is chosen sufficiently small, i.e. of the order of 1/10 of a grid spacing or less.

We also compute the trace of the velocity dispersion tensor. The velocity dispersion σ^2 is computed similarly to the velocity field, i.e. we define

$$\sigma^2 = -\frac{T_i^i}{T_0^0} - v^i v_i, \tag{7.7}$$

where T_i^i, T_0^0 and v^i are computed on the vertices of our grid cells with the CIC method (Hockney & Eastwood 1988). The definition of σ^2, similarly to the velocity field, requires a prescription for estimating the squared velocity field $v^2 \equiv T_i^i/T_0^0$ in the voids. We adopt for this computation the *rescaled* method, which seems most physical, in particular we will rescale the squared velocity field with the square of the rescaling factor that we adopt for the velocity

$$v^2(\mathbf{x}, t_i) = [\dot{D}_1(t_i)/\dot{D}_1(t_{i-1})]^2 v^2(\mathbf{x}, t_{i-1}). \tag{7.8}$$

In the next section we describe the decomposition of the velocity field in a divergence and a rotational part and the computation of the divergence and vorticity spectra.

7.2.2 Decomposition of the velocity field

Since the N-body particles in *gevolution* are projected onto a regular Cartesian grid, the Helmholtz decomposition of the velocity field is easily performed in Fourier space.

In Fourier space, the divergence θ and the vorticity $\boldsymbol{\omega}$ are

$$\theta(\mathbf{k}) = k^2 v(k), \tag{7.9}$$

$$\boldsymbol{\omega}(\mathbf{k}) = i\mathbf{k} \wedge \mathbf{v}_R(\mathbf{k}). \tag{7.10}$$

As usual, we denote a function and its Fourier transform by the same symbol. The Fourier transform is only indicated by the argument $\mathbf{k}$. In the code, the velocity field is decomposed into its divergence and rotational part

$$\theta = i(k_j v^j), \qquad v_R^i = P^{ij} v_j, \tag{7.11}$$

where i, j are index running over the Fourier space coordinates and P^{ij} is the transverse projector, $P^{ij} = \delta^{ij} - k^i k^j / k^2$.

In the numerical code, we have to replace the vector $\mathbf{k}$ by the correspondent discrete wave vector,

$$\mathbf{k}_{u,v,w} \equiv \frac{1}{\Delta_{\mathrm{grid}}} \left(\sin \frac{2\pi u}{N}, \sin \frac{2\pi v}{N}, \sin \frac{2\pi w}{N} \right), \tag{7.12}$$

where Δ_{grid} is the grid spacing, N is the number of grid points in each direction and u, v, w are integers running from 1 to N.

In a spatially homogeneous Universe the power spectra of the velocity, divergence and vorticity are of the form

$$\left\langle v^i(\mathbf{k}) v^{*\,j}(\mathbf{k}') \right\rangle = \delta(\mathbf{k} - \mathbf{k}')(2\pi)^3 \left(\delta^{ij} P_1(k) + \frac{k^i k^j}{k^2} P_2(k) \right), \tag{7.13}$$

$$\left\langle v_{\mathrm{R}}^i(\mathbf{k}) v_{\mathrm{R}}^{*\,j}(\mathbf{k}') \right\rangle = \delta(\mathbf{k} - \mathbf{k}')(2\pi)^3 \frac{1}{2} \left(\delta^{ij} - \frac{k^i k^j}{k^2} \right) P_R(k), \tag{7.14}$$

$$\left\langle v_{\mathrm{G}}^i(\mathbf{k}) v_{\mathrm{G}}^{*\,j}(\mathbf{k}') \right\rangle = \delta(\mathbf{k} - \mathbf{k}')(2\pi)^3 \frac{k^i k^j}{k^2} P_G(k), \tag{7.15}$$

$$\left\langle v^i(\mathbf{k}) v_i^*(\mathbf{k}') \right\rangle = \delta(\mathbf{k} - \mathbf{k}')(2\pi)^3 P_v(k), \tag{7.16}$$

$$\left\langle \theta(\mathbf{k}) \theta^*(\mathbf{k}') \right\rangle = \delta(\mathbf{k} - \mathbf{k}')(2\pi)^3 P_\theta(k), \tag{7.17}$$

$$\left\langle \omega^i(\mathbf{k}) \omega^{*\,j}(\mathbf{k}') \right\rangle = \delta(\mathbf{k} - \mathbf{k}')(2\pi)^3 \frac{1}{2} \left(\delta^{ij} - \frac{k^i k^j}{k^2} \right) P_\omega(k). \tag{7.18}$$

The Dirac delta is a consequence of statistical homogeneity. The expression $\delta^{ij} P_1(k) + k^i k^j / k^2 P_2(k)$ of (7.13) is the most general isotropic vector power spectrum, and the prefactors of P_R in (7.14) and of P_ω in (7.18) ensure $k_i v_R^i = 0$ and $k_i \omega^i = 0$, respectively. The prefactor of P_G in (7.15) makes use of the fact that $\mathbf{v}_G$ is a gradient field. Of course these spectra are not independent but enjoy the following

relations which are easily verified:

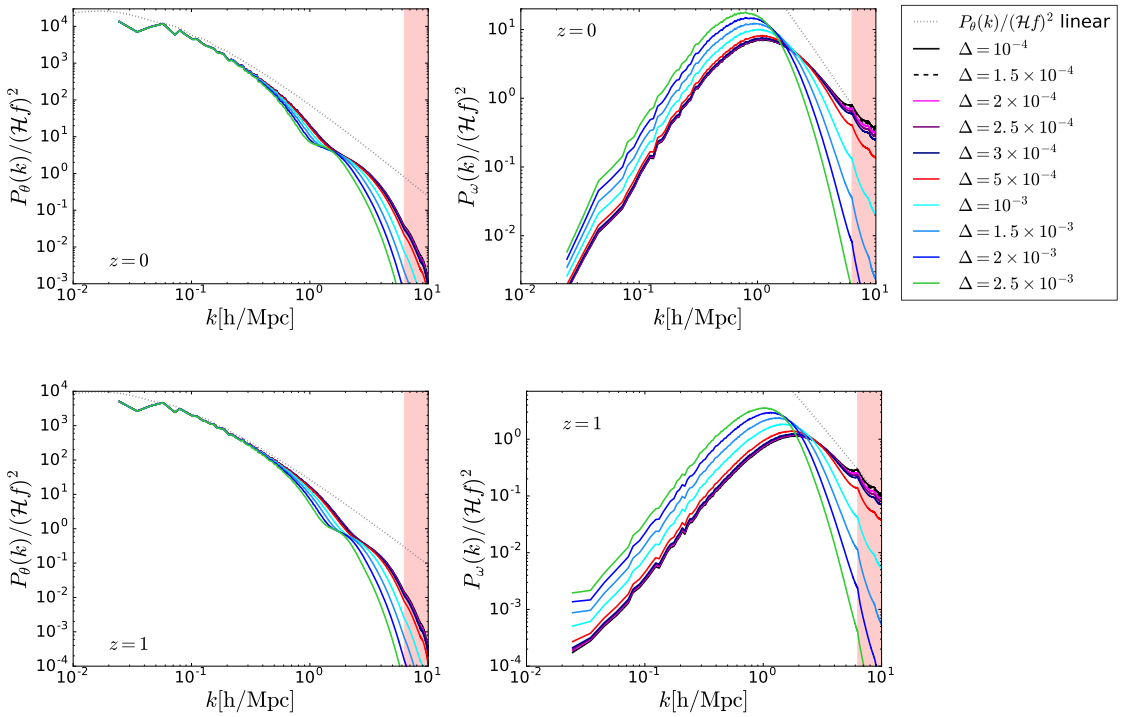
$$P_R = 2P_1, \tag{7.19}$$
$$P_G = P_2 + P_1, \tag{7.20}$$
$$P_v = 3P_1 + P_2 = P_R + P_G, \tag{7.21}$$
$$P_\theta = k^2 P_G, \tag{7.22}$$
$$P_\omega = k^2 P_R, \tag{7.23}$$
$$k^2 P_v = P_\theta + P_\omega \,. \tag{7.24}$$

In the code, we implement the power spectrum of the rotational velocity $P_R(k)$. However, in the next section we will show our results in terms of the vorticity power spectrum P_ω, which is computed from (7.23).

7.2.3 Method comparison

We tested the four methods described in section 7.2.2 by running a simulation for each method with the same parameters: box size $L = 256\,\mathrm{Mpc}/h$, which corresponds to a volume $V = L^3$; number of gridpoints, $N^3_{\mathrm{grid}} = 512^3$; number of particles $N^3_{\mathrm{part}} = 1024^3$. We will refer to these simulation settings as *low-resolution*. The initial conditions of these simulations have been generated from the same seed.

In figure 7.1 we show the power spectra of the divergence (left panels) and the vorticity (right panels), computed from the smoothed velocity field. The smoothing is performed with a Gaussian filter, therefore this method has an extra degree of freedom that we have to test, i.e. the standard deviation of the Gaussian which sets the smoothing scale, which we denote as Δ. We run 10 simulations for different values of the smoothing scale, starting from $\Delta \approx 2 \times 10^{-3}$, which is approximately the size of one grid cell, down to $\Delta \approx 10^{-4}$, i.e. a smoothing size of 1/20 the resolution of this set of simulations. Top and bottom panels refer to different redshift, $z = 0$ and $z = 1$ respectively. The divergence spectra do not depend on Δ on large scales, while we observe a substantial suppression of power for $k \gtrsim 0.3\,h/\mathrm{Mpc}$ when increasing the smoothing scale from 10^{-4} to 2.5×10^{-3}. On the other hand, the vorticity power spectrum depends on the size of the smoothing also on large scales. In particular, there is a transfer of power from small to large scales, when we increase the value of Δ. This is due to the fact that a relatively large smoothing scale (i.e. the size of a grid cell) introduces a spurious vorticity in our computation, since the velocity field at a given grid point is receiving a non-negligible contribution from particles located in the other regions of the simulation domain. Since the vorticity is a purely non-linear effect, and in our cold dark matter

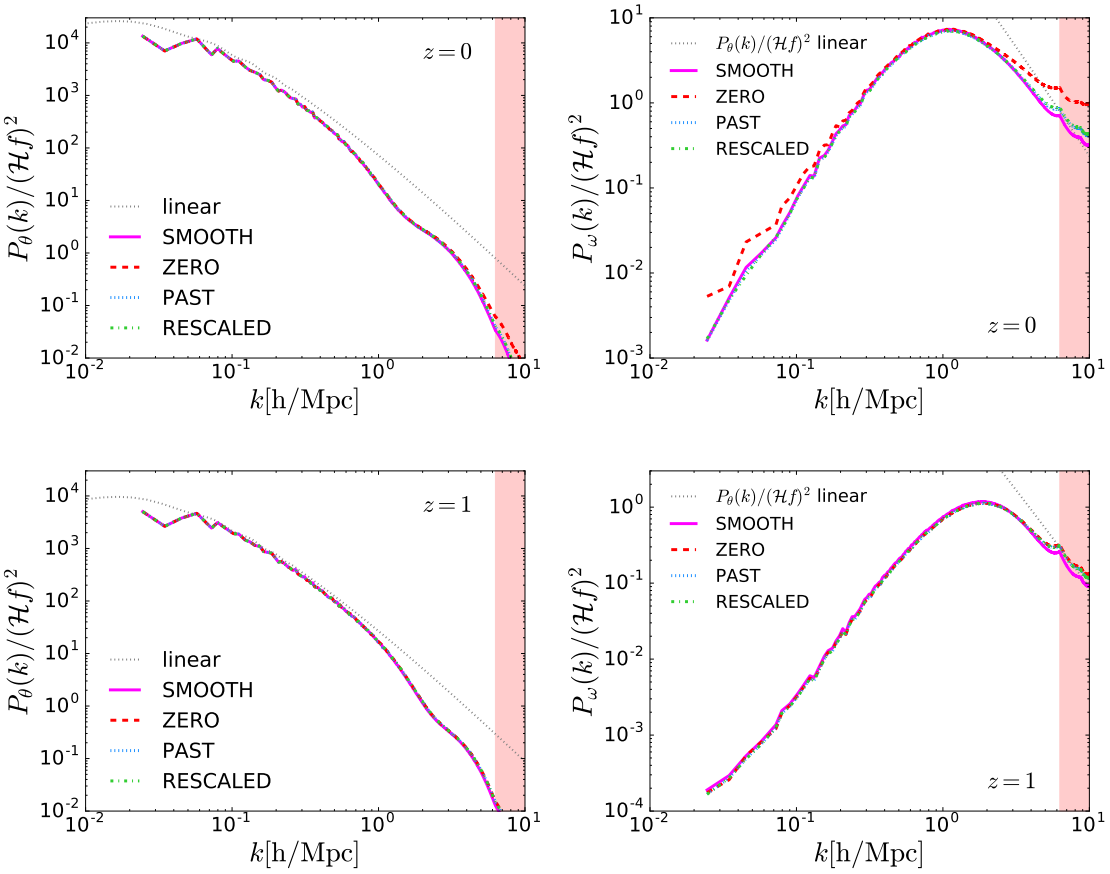


(a) Power spectrum - θ **(b)** Power spectrum - ω

Figure 7.1: Power spectrum of the divergence θ (left panel) and the vorticity ω (right panel) of the velocity field at $z = 1$ (top panels) and $z = 0$ (bottom panels). The power spectra are computed by applying a Gaussian smoothing, with standard deviation Δ (in units of the size of the box), to both the momentum density and density field. Different colors refer to different smoothing scales. A smoothing of the size of a grid cell corresponds to $\Delta = 1/512 \approx 2 \times 10^{-3}$. The fields are normalized in such a way that, on linear scales, P_θ matches the matter power spectrum. The red shaded region represents modes beyond the Nyquist frequency.

simulations it is generated by orbit crossing of the particles, the smoothing of small scales introduces additional velocity dispersion which sources vorticity, see (Cusin et al. 2017). However, this spurious effect is small for a sufficiently small smoothing scale, and from figure 7.1 we infer that the vorticity power spectrum converges if we choose a smoothing scale of 1/10 the size of a grid cell in our simulation. For this value the spurious vorticity induced by the smoothing is negligible. Therefore, from now on, we will refer to the method *smooth* to be the smoothing with a Gaussian kernel and a standard deviation of 1/10 the size of a grid cell.

In figure 7.2 we compare the power spectra P_θ and P_ω obtained with the four different methods to define the velocity field, at $z = 0$ (top panels) and $z = 1$ (bottom panels). At redshift $z = 1$ all methods give consistent results. At $z = 0$ the differences between the four prescriptions are more prominent. However, all methods, apart from the one assigning $\mathbf{v} = 0$ to the empty points (red line), do agree up to the Nyquist frequency $k_{\rm Nyq} = \pi N_{\rm grid}/L$ (red shaded zone).



(a) Power spectrum - θ **(b)** Power spectrum - ω

Figure 7.2: Power spectrum of the divergence θ (left panel) and the vorticity ω (right panel) field at $z = 0$ (top panels) and $z = 1$ (bottom panels). Different colors refer to the different methods for the velocity field computation. The method 'SMOOTH' refers to a Gaussian smoothing with width $\Delta = 2 \times 10^{-4}$ Mpc/h. The spectra are normalized by dividing by $(f\mathcal{H})^2$ so that, on linear scales, P_θ matches the matter density power spectrum, see Eq. 7.6. The results shown are for $L = 256$ Mpc/h, $N_{\rm grid} = 512$ and $N_{\rm part} = 1024$. The red shaded region represents the modes beyond the Nyquist frequency.

To summarize, we show that the *past*, *rescaled* and *smooth* methods give the same results on all the scales relevant for our simulations. The *rescaled* method is the physically best motivated prescription, and it has the advantage that it does not introduce any extra degrees of freedom that need to be tested for different parameter settings. Therefore, unless stated otherwise, the simulations presented in the next section employ this prescription.

7.3 Results

In this section we collect the main results of this chapter. The simulations discussed in this section are summarized in table 7.1. All the simulations that will be discussed

have the same physical box size of 256 Mpc/h. We tested for finite-volume effects by running a simulation with a larger box size and same spatial resolution $L/N_{\rm grid}$ as the low-resolution simulations. The results are shown in appendix C.1. We found that the divergence and vorticity power spectra coincide in the two cases. Therefore, the finite-volume effects for $L = 256$ Mpc/h are negligible. In section 7.3.1 we show the snapshots of our low-resolution simulations for the divergence, the vorticity and the velocity dispersion. In section 7.3.2 we investigate the large-scale slope of the vorticity power spectrum, while in section 7.3.3 we show the result of the convergence test and we study the redshift dependence of the amplitude of the vorticity power spectrum. Finally, in section 7.3.4 we investigate the correlations of the velocity divergence with the density field and the velocity dispersion, and the correlations between the rotational part of the velocity and the vector perturbation of the metric $\mathbf{B}$. In appendix C.2 we compare our method for estimating the vorticity field with the Delaunay tessellation method (DTFE). We show that the DTFE method has a faster convergence than our method with respect to the grid resolution. However, we argue that the difference between the two methods does not have an impact on our final results which are obtained by higher resolution simulations than the ones used in the test.

	$N_{\rm grid}$	$N_{\rm part}$	L [Mpc/h]	# Realizations
Low resolution	512	1024	256	16
High resolution 1	1024	2048	256	1
High resolution 2	2048	4096	256	1

Table 7.1: Summary of the simulations discussed in sections 7.3.1, 7.3.2, 7.3.3, 7.3.4.

7.3.1 Snapshots

In figure 7.3 we show a slice through one of our low resolution simulations at redshift $z = 3$ (top panels), $z = 1$ (center panels) and $z = 0$ (bottom panels). We show the divergence, the vorticity and the velocity dispersion. The divergence and the vorticity are normalized by the factor $(\mathcal{H}f)$, therefore they are dimensionless and the divergence field, at the linear level, satisfies $\theta/(\mathcal{H}f) = -\delta$, where δ is the dark matter density fluctuation.

At first glance we can notice that the overall structures agree very well. The divergence field is negative in the most over-dense regions, which translates into an inflow of particles, while it is positive in the voids, as expected. We also observe that both the vorticity and the velocity dispersion are peaked locally in the collapsing regions, where the orbit crossing of the particles occurs. They have less large-scale

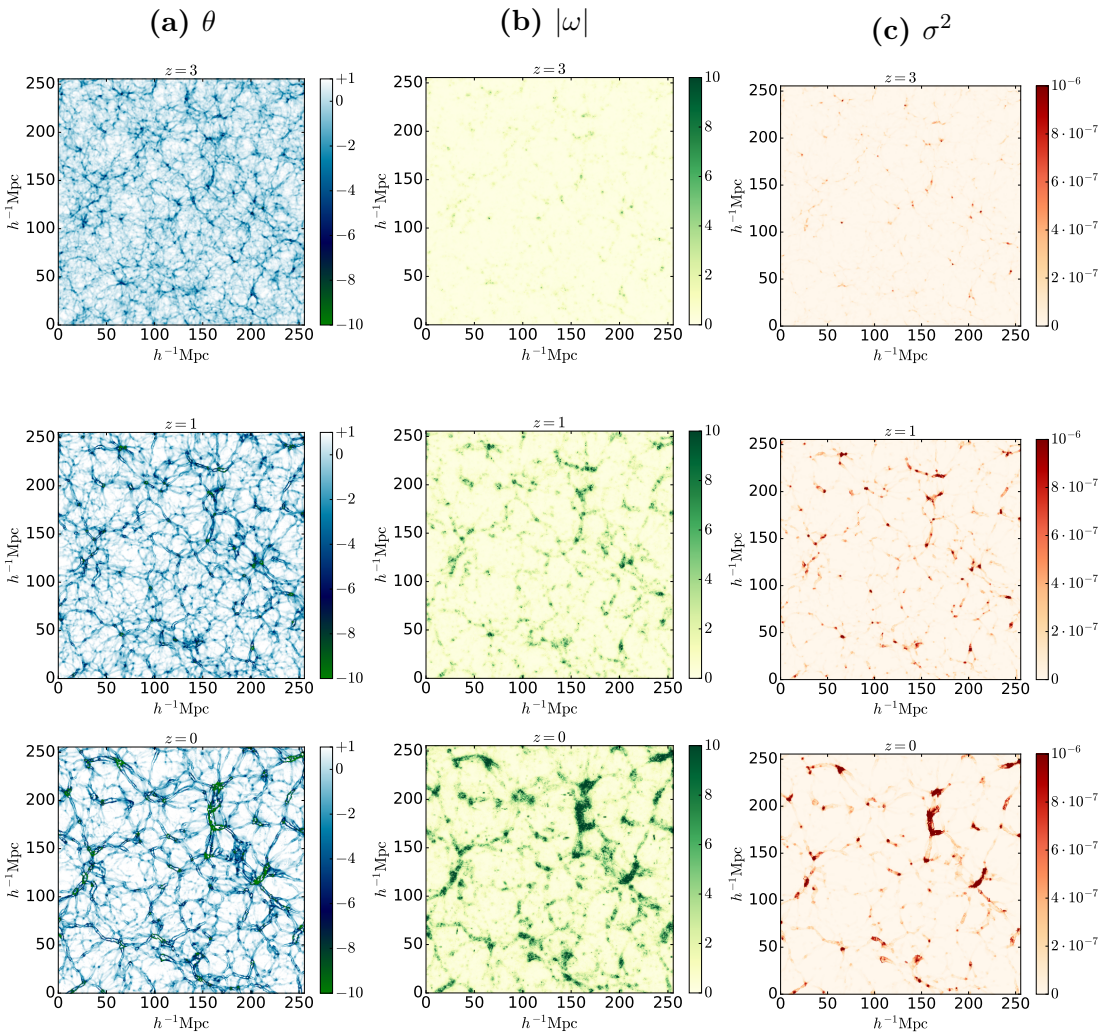


Figure 7.3: A slice through our simulation at $z = 3$ (top panels), $z = 1$ (middle panels), $z = 0$ (bottom panels) showing the divergence field (left panels), the absolute value of the vorticity (middle panels) and the velocity dispersion (right panels).

power than θ which extends over the entire grid. At $z = 3$ both the vorticity and the velocity dispersion are very close to zero in the entire domain. At $z = 3$ the mean velocity dispersion in the simulation is approximately $\bar{\sigma}^2 \approx 5 \times 10^{-9}$, and it increases to $\bar{\sigma}^2 \approx 10^{-8}$ at $z = 1$ and to $\bar{\sigma}^2 \approx 3 \times 10^{-8}$ at $z = 0$. Hence it seems to grow somewhat faster than $1/(z + 1)$, but not like $1/(z + 1)^2$ as a background velocity dispersion would, as argued in (Cusin et al. 2017).

7.3.2 Large-scale behavior

In order to study the large-scale behavior of the vorticity power spectrum, we run 16 realizations of the *low-resolution* simulations described in table 7.1. In fact, while the amplitude of the vorticity power spectrum is resolution dependent

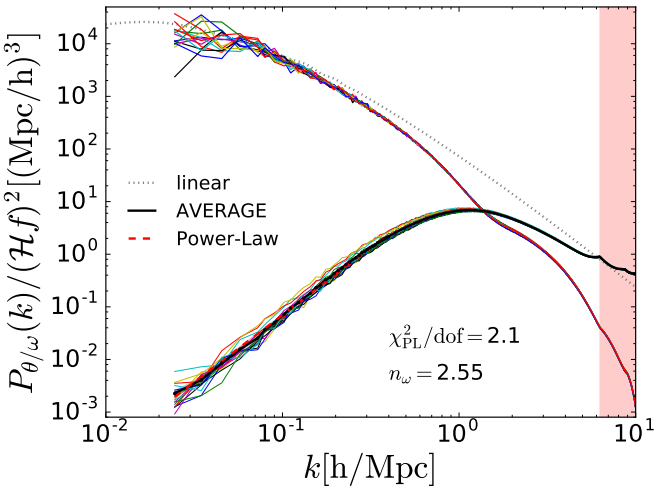


Figure 7.4: Divergence and vorticity power spectra for different realizations of the low-resolution simulation from table 7.1, at $z = 0$. Different colors refer to different initial seeds. The black continuous line represents the average over the 16 realizations, while the red dashed line represents the result of the power-law fit, with $k_{\mathrm{max}} = 0.4\,h/\mathrm{Mpc}$. The χ^2 divided by the degrees of freedom for the fit is approximately 2.

and needs a higher resolution simulation to be estimated, the spectral index is not affected by the resolution.

In figure 7.4 we show the divergence and vorticity power spectra for these simulations. Different colors denote different initial seeds. The averaged power spectrum is represented by the black continuous line. For $k \lesssim 0.4\,h/\mathrm{Mpc}$ the vorticity power spectrum is well modelled by a power-law with amplitude A_ω and spectral-index n_ω,

$$P_\omega(k) = A_\omega k^{n_\omega}. \tag{7.25}$$

We have estimated the spectral index by fitting the model (7.25) to the average power spectrum from our simulations. The errors on the average power spectrum are computed as the standard deviation divided by the square root of the number of realizations.

The red dashed line in figure 7.4 represents the result of our fit, and the best fit value for the spectral index is

$$n_\omega = 2.55 \pm 0.02, \tag{7.26}$$

where the uncertainty quoted above is the statistical error. The best-fit value of the spectral index slightly depends on the $k-$range selected for the fit, with fluctuations of the order of 10% in the range $k_{\mathrm{max}} \in [0.3, 0.5]$ h/Mpc.

The values of the spectral index that we find is fairly consistent with previous results obtained from N-body simulations (Pueblas & Scoccimarro 2009; Hahn et al. 2015; Thomas et al. 2015), where the simulations were performed with a different N-body code and the vorticity field was reconstructed with different techniques. However, from a theoretical point of view, there is no unequivocal prediction for the vorticity power spectrum.

In (Cusin et al. 2017), using a perturbative approach, the vorticity power spectrum, in the large-scale limit has a power-law behavior with spectral index $n_\omega = 4$. However, in that analysis, vorticity is generated by the intrinsic dark matter velocity dispersion, while in our N-body approach, CDM has initially no velocity dispersion and it is the orbit crossing of particles that induces both, velocity dispersion and vorticity.

In (Scoccimarro 2001; Pueblas & Scoccimarro 2009) an expression for the vorticity power spectrum, generated by orbit-crossing, is computed analytically. Assuming that the vorticity that develops after shell-crossing is proportional to the curl of the mass-weighted velocity field, the vorticity power spectrum is expected to have a quadratic scale dependence in the low-k limit, i.e. $n_\omega = 2$. This is also expected from simple causality: the initially vanishing vorticity can be correlated only in causally connected points. Therefore, the vorticity correlation function has compact support, hence its Fourier transform, the power spectrum is analytic for small $\mathbf{k}$. The non-analytic pre-factor then requires that $P_\omega \propto k^{2n},\ n \in \mathbb{N}$ (see e.g. (Durrer & Caprini 2003) for more details). Without some special 'conservation law' we therefore expect $P_\omega \propto k^2$ on large scales. One might argue that the deviation from $n_\omega = 2$ comes from numerical contributions to the vorticity. But we have checked that our results are stable under increase of resolution (see Appendix B) and therefore are confident that the result at intermediate scales can be trusted. We rather suppose that we might miss some vorticity on the very largest simulated scales which would be induced from even larger scales not contained in the simulation. This would render the power spectrum too steep.

Results for the vorticity spectrum have also been obtained within the framework of the Effective Field Theory of Large Scale Structures (EFT of LSS) (Carrasco et al. 2012, 2014). The second of these references predicts $n_\omega = 2.8$ for $k \in [0.2, 0.6]$ h/Mpc, while the slope of the power spectrum becomes steeper on larger scales, i.e. $n_\omega = 3.6$ for $k \in [0.1, 0.2]$ h/Mpc.

It is also interesting to note that at $k \simeq 1$ h/Mpc where the vorticity power spectrum peaks, the divergence power spectrum has a dip and drops significantly below the linear perturbation theory value (dotted line in fig. 7.4). Hence when

non-linear structures form, part of the power in the divergence is converted into vorticity. We interpret this as the effect of angular momentum conservation which prevents further infall and forces the particles to rotate around a given structure.

7.3.3 Resolution and redshift dependence

In this subsection, we describe some important tests we performed to ensure convergence, and consequently the reliability, of the results obtained from our simulations. We remark that, since in *gevolution* the fields are computed on a regular Cartesian grid, it is crucial to test the convergence of our results with respect to two quantities: the number of particles in each grid cell and the total number of particles (or total number of cells) in the box. These two parameters are tested in sections 7.3.3 and 7.3.3, respectively.

On the choice of $N_{\mathbf{part}}/N_{\mathbf{grid}}$

As *gevolution* is a code based on the particle-mesh (PM) scheme, the initial conditions are such that there are $R^3 \equiv (N_{\rm part}/N_{\rm grid})^3$ particles in a single grid cell, and our results may depend on it, hence we need to determine what is a good value for R. We ran several simulations for values $R = 1, 2, 4$, with a fixed box size $L = 256$ Mpc$/h$, and fixed $N_{\rm grid} = 512$.

The results are displayed in figure 7.5. As we can immediately see, the vorticity power spectrum is significantly higher for the case $R = 1$ than for the others. On the other hand, the values $R = 2$ and $R = 4$ give nearly identical results, except at the smallest scales close to the Nyquist frequency. As the $R = 4$ simulations would be much more computationally demanding, in almost all of our simulation runs we have used the value $R = 2$, corresponding to 8 particles per grid cell on average.

Dependence on spatial resolution

In order to make sure the results are reliable, we have tested the dependence of the obtained power spectra on the spatial resolution of the simulation. In figure 7.6 we show the velocity divergence (left panels) and vorticity (right panels) power spectra for the three simulations in table 7.1, which have number of particles per side $N_{\rm part} = 1024, 2048, 4096$ with a fixed particle-to-grid ratio $N_{\rm part}/N_{\rm grid} = 2$. Top and bottom panels refer to redshifts $z = 0$ and $z = 1$, respectively.

We notice that, at large scales, the velocity divergence spectrum is not very sensitive to the spatial resolution, while the dependence increases as we go to nonlinear scales. Of course we can expect the simulations to agree only on scales which are sufficiently larger than their Nyquist frequencies.

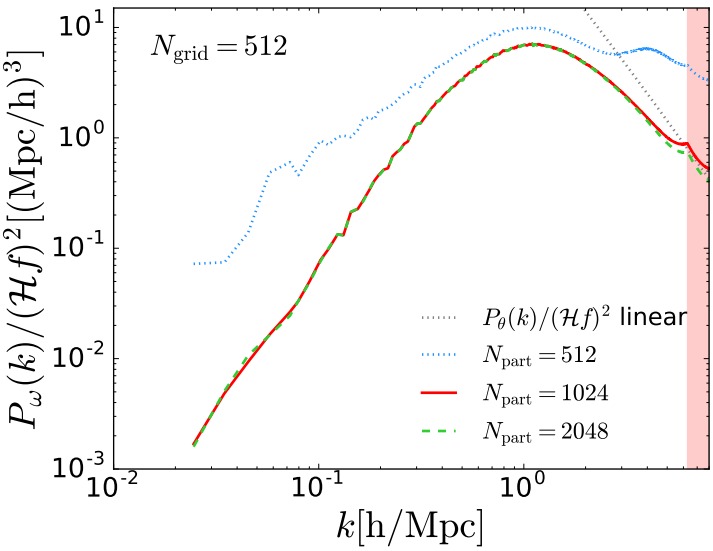


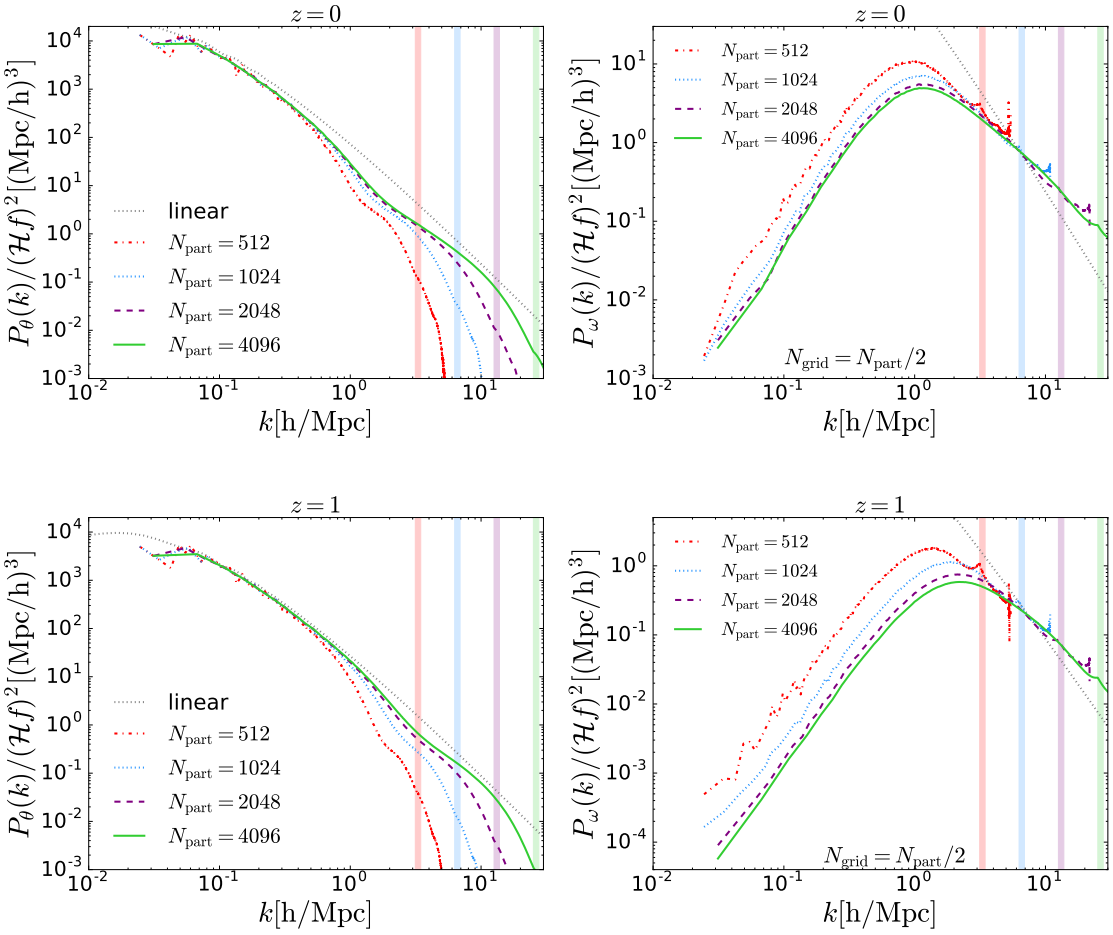
Figure 7.5: The vorticity power spectrum for $R = 1, 2, 4$ at $z = 0$, with $N_{\rm grid} = 512$ and $L = 256$ Mpc$/h$

On the other hand, the power spectrum of the vorticity, being an entirely nonlinear phenomenon, exhibits a resolution dependence at all scales, notably the amplitude is too high for low resolution simulations, and it eventually converges for a sufficient spatial resolution, while the shape of the spectrum is roughly resolution independent.

Comparing the $z = 0$ and $z = 1$ plots, we see that the convergence is seemingly worse for higher redshift. This can be attributed to the fact that shell (orbit) crossing, which is the generation mechanism for vorticity in our simulation, occurs much less at higher redshifts. Therefore, the spurious vorticity induced by numerical velocity dispersion resulting from the CIC projection of the particle velocities on the grid is more significant when there is less physical vorticity in our simulation, and we would need a higher spatial resolution to obtain reliable results at higher redshifts. Consequently, we expect the power spectrum of the vorticity to have a progressively worse convergence as we go to higher redshifts.

This is shown in figure 7.7, where we compared the two simulations *high resolution 1* and *high resolution 2* in table 7.1 at redshift $z = 0, 1, 2, 3, 5$. While for $z = 0$ and $z = 1$ the vorticity power spectra are in good agreement, for $z > 1$ the difference between the results of the two simulations is significant and this trend worsens with increasing redshift. For this reason, we focus our attention primarily on the redshift range from $z = 0$ to $z = 1$.

It is interesting to notice in figure 7.7 that the position of the peak of the vorticity power spectrum is redshift dependent, and it is moving from small scales at



(a) Power spectrum - θ **(b)** Power spectrum - ω

Figure 7.6: Velocity divergence and vorticity power spectra at $z = 0$ (*top*) and $z = 1$ (*bottom*) respectively. The highlighted bands represent the Nyquist frequency for each resolution.

high redshift to larger scales at $z = 0$. Between $z = 1$ and $z = 0$ the peak position is shifted from $k_{\mathrm{p}} \approx 2\ h/\mathrm{Mpc}$ to $k_{\mathrm{p}} \approx 1\ h/\mathrm{Mpc}$, while its amplitude grows from $P_\omega/(\mathcal{H}f)^2(k_{\mathrm{p}}, z = 1) \approx 0.6\,\mathrm{Mpc}^3/h^3$ to $P_\omega/(\mathcal{H}f)^2(k_{\mathrm{p}}, z = 0) \approx 5\,\mathrm{Mpc}^3/h^3$. This shift of the peak is roughly in the same position as the dip in the divergence power spectrum, see figure 7.6, and it indicates the scale where non-linearities become strong which moves to successively larger scales. As mentioned in the previous subsection, we interpret this as the scale at which further infall of particles is slowed down due to angular momentum conservation. Similar results have also been obtained in (Zhu et al. 2010; Zhu & Feng 2017).

Following (Pueblas & Scoccimarro 2009), we have also considered in our analysis possible aliasing effects due to the finite size of the grid and the use of discrete Fourier transforms, however they were found to be negligible for our resolution and redshift range.

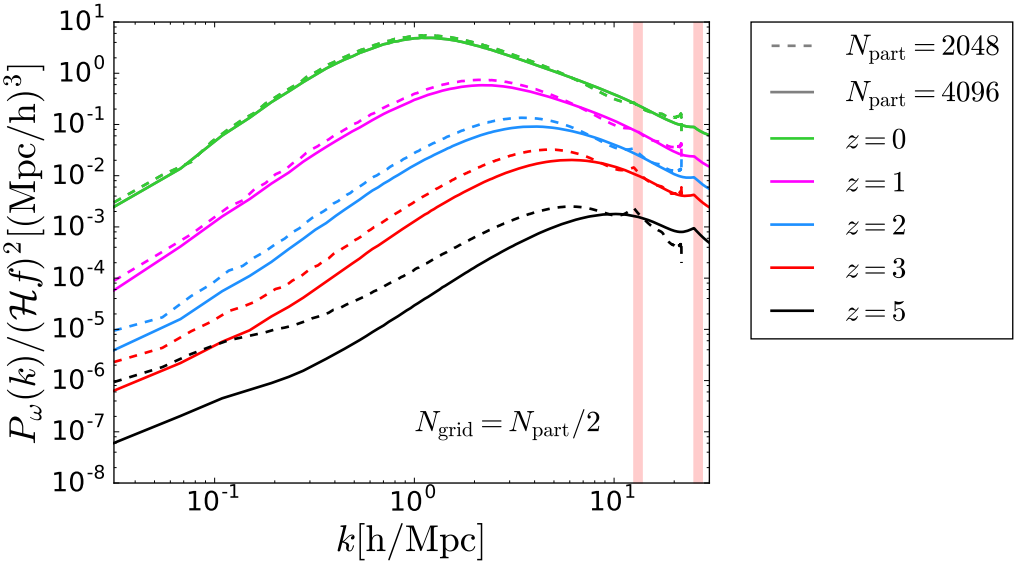


Figure 7.7: The vorticity power spectra for redshifts $z = 0, 1, 2, 3, 5$.

Redshift dependence

Focusing primarily on the range $z \in [0, 1]$, we have studied the large scale redshift dependence of the vorticity power spectrum. We have chosen three different modes in the large scale regime, $k = 0.05, 0.2, 0.4\ h/\mathrm{Mpc}$. Figure 7.8 shows the redshift and growth-factor dependence of the vorticity power spectrum at fixed scale. Following (Pueblas & Scoccimarro 2009), we have assumed that the vorticity power spectrum has a redshift dependence of the form

$$P_\omega(k, z) \propto (\mathcal{H}f)^2 D^\gamma(z). \tag{7.27}$$

We have performed a least squares fit to the data; the fits to the numerical data are displayed in black dashed colors in figure 7.8. We found values of γ in the range $[7, 7.4]$, which is mostly consistent with the result obtained in (Pueblas & Scoccimarro 2009; Thomas et al. 2015). However, we note that a) we are not able to test the fit (7.27) for higher redshift due to the convergence issued outlined in the previous section, b) we cannot obtain a good fit on all the scales of the simulations with a global parameter γ. Our findings on the redshift evolution of the vorticity spectrum confirm the results of Thomas et al. (2015), namely that the growth factor for the vorticity spectrum cannot be modelled as a simple power law in a wide redshift range.

We also notice that the time evolution of the vorticity spectrum extracted from N-body simulation results is very different from the result obtained from a

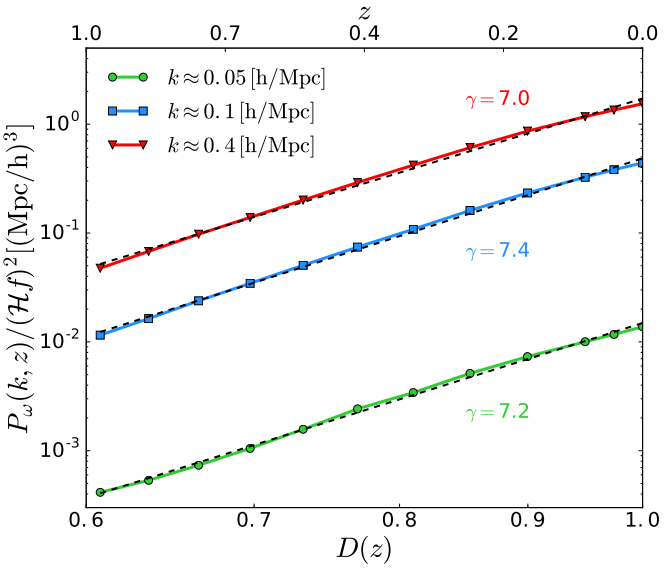


Figure 7.8: Vorticity power spectrum at fixed mode k as a function of the growth function (bottom axis) and redshift (top axis). The black dashed lines are the fit to the numerical data.

perturbative approach. In fact, the power spectrum computed in (Cusin et al. 2017) grows linearly with the growth function,

$$P_\omega^{\rm pert}(k,z) \propto D(z). \tag{7.28}$$

As already noted in the previous section, this very significant difference in the redshift evolution is probably due to the different physical mechanism that is generating vorticity in the two cases. Being a perturbative approach, the method discussed in (Cusin et al. 2017) cannot really handle shell crossing which is the dominant effect leading to vorticity production in N-body simulations, as we shall see below.

In figure 7.9 we plot the divergence and vorticity spectra together with the total velocity power spectrum $k^2 P_{\rm v} = P_\omega + P_\theta$ at $z = 0$ (left panel) and $z = 1$ (right panel). On large scales, as expected, the divergence contributes most to the total velocity spectrum. However, on small scales the divergence spectrum drops faster than the vorticity and beyond $k \approx 2\ h/\text{Mpc}$, where the velocity power spectrum exceeds the linear perturbation theory result, vorticity becomes the dominant component in shaping the total velocity power spectrum.

7.3.4 Cross-spectra

In our simulations we have also computed the cross-spectra of the velocity divergence with the density contrast and the velocity dispersion σ^2 and the cross-correlation of

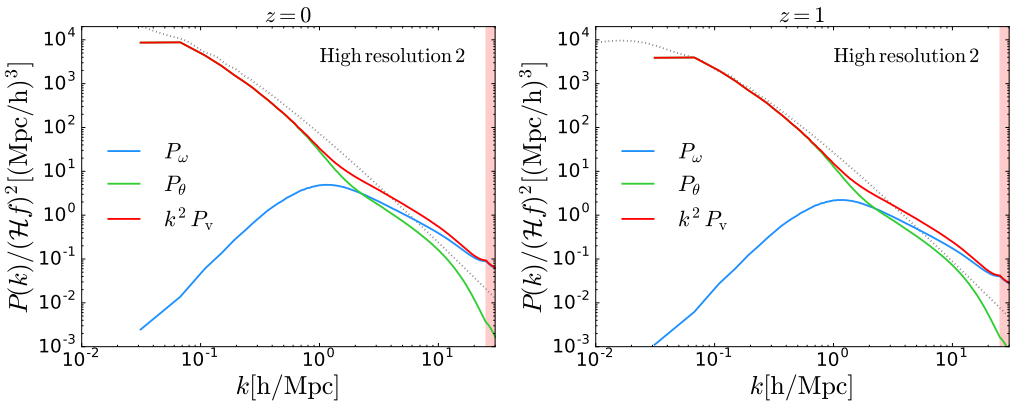


Figure 7.9: Vorticity, divergence and velocity power spectra at $z = 0$ (left panel) and $z = 1$ (right panel). Both plots refers to the simulation *high resolution 2* in table 7.1. The dotted line represents the linear divergence spectrum.

the rotational part of the velocity field with the metric vector perturbations $\mathbf{B}$ [5]. The results of this section refer to the simulation *high resolution 1* in table 7.1.

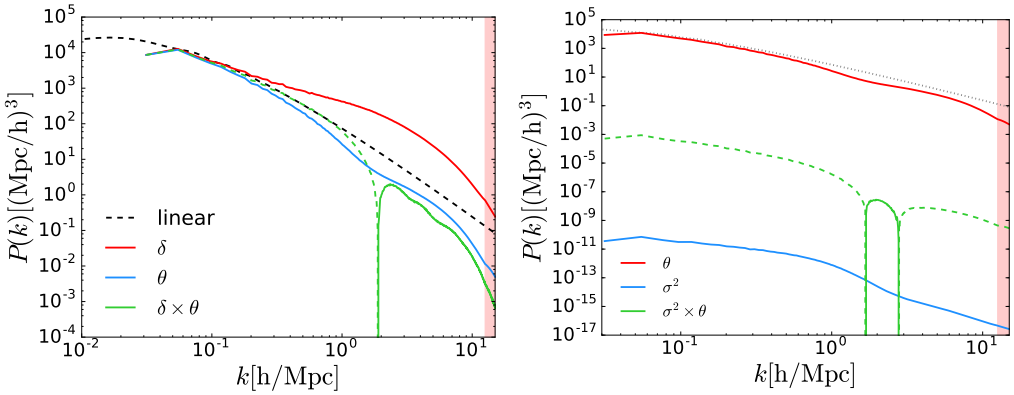


Figure 7.10: Left panel: The spectra of the density contrast (red), the velocity divergence (blue), and their cross-spectrum (green) at $z = 0$. Right panel: The spectra of the velocity divergence (red), the velocity dispersion (blue), and their cross-spectrum (green) at $z = 0$.

In figure 7.10 (left panel) we plot the power spectra of the density contrast, velocity divergence, and their cross spectra, respectively, at $z = 0$. Both the power spectrum of the divergence and the cross-spectrum are normalized in such a way that all the spectra are dimensionally consistent. From the linearized continuity equation we expect on large scales $\theta = -\mathcal{H}f\delta$. Therefore, we normalize the divergence spectrum by a factor $(\mathcal{H}f)^2$ and the cross-spectrum by a factor $\mathcal{H}f$. Consistently with this definitions, on large scale all the spectra are approximately equal. On the

[5]Note that the cross-spectrum $P_{\omega B}$ vanishes by parity conservation; under spatial inversions, ω is a pseudovector, while B is a true vector, and the relationship between the correlators in Fourier space is $\langle \omega_i B_j^* \rangle = i\epsilon_{imn} k_m \langle v_{Rn} B_j^* \rangle$, which would be odd under parity, and hence must be zero.

other hand, due to non-linear evolution at late times, the divergence spectrum drops on smaller scales while the density spectrum increases. Interestingly, the correlation spectrum remains close to the linear perturbation theory value until $k \simeq 1\,h/\mathrm{Mpc}$ where it increases until it changes sign at $k \approx 2\,h/\mathrm{Mpc}$ and then settles at a fraction of the divergence spectrum. This is an indication of the scale at which shell-crossing is relevant. In fact, after shell-crossing which leads to the formation of structures, the initially anti-correlated density and velocity divergence are positively correlated on small scales. We understand that after shell crossing a previous infall onto a filament changes sign becomes actually an outflow, leading to positive density − divergence correlations. See also (Hahn et al. 2015) for a discussion of this.

In figure 7.10 (right panel) we display the power spectra of the velocity divergence, the velocity dispersion, and their cross spectra, respectively, at $z = 0$. On large scales velocity dispersion spectrum is proportional to the θ spectrum and the two are strongly anti-correlated. This is because the fluctuation of particle velocities around their mean bulk flow is a second order effect that is most efficiently generated in collapsing regions which have negative θ. However, at small scales, the dependence becomes highly non-linear and the cross correlation between σ^2 and θ even briefly changes sign around the scale where also δ exhibits a sign change. As one can see in figure 7.11, at $z = 1$ there is no sign change in the correlation of σ^2 and θ, they remain anti-correlated on all scales.

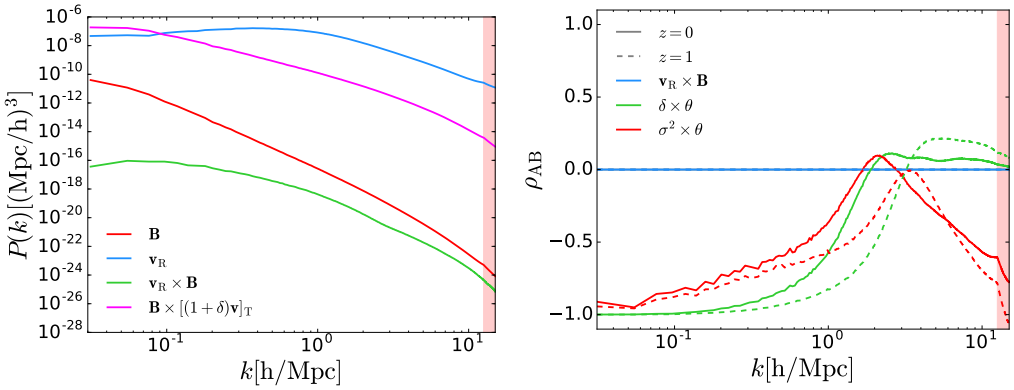


Figure 7.11: Left panel: The spectra of the metric vector perturbations (red), the rotational velocity (blue), and their cross-spectrum (green) at $z = 0$. The magenta line represents the cross-correlation of the metric perturbation with its source term (see equation (7.29)). Right panel: The correlation coefficients between the rotational velocity v_R and the metric vector perturbation B (blue), the density contrast δ and velocity divergence θ (green), the velocity dispersion σ and velocity divergence θ (red), at $z = 0$ (solid) and $z = 1$ (dashed) respectively.

In figure 7.11 (left panel) we plot the power spectra for $\mathbf{B}$ and $\mathbf{v}_R$ and their cross spectrum. The rotational part of the velocity is expected to partially source

the vector perturbation in the metric. In fact, by taking the transverse projection of the $0i$ Einstein constraint equation linearized in $\mathbf{B}$, we find

$$k^2\mathbf{B} = 6\mathcal{H}^2\Omega_{\rm m}[(1+\delta)\mathbf{v}]_{\rm T} = 6\mathcal{H}^2\Omega_m[\mathbf{v}_R + (\delta\,\mathbf{v})_{\rm T}], \tag{7.29}$$

where the underscript T denotes the transverse projection of the source term $(1+\delta)\mathbf{v}$. In figure 7.11 we compare the cross-spectrum of the metric perturbations $\mathbf{B}$ with the rotational velocity (green line) and with the full source term in equation (7.29), which is computed from equation (7.29), i.e. $P_{\mathrm{B}\,[(1+\delta)\mathrm{v}]_{\rm T}} = P_{\rm B}\,k^2/(6\mathcal{H}^2\Omega_{\rm m})$. The cross-spectrum of $\mathbf{B}$ with $\mathbf{v}_R$ is smaller by many orders of magnitude than the cross-spectrum of $\mathbf{B}$ with the full source term. Therefore, we expect that the main source of vector perturbations does not come from the rotational velocity, but from the product of the velocity with the density contrast, i.e. from $(\delta\,\mathbf{v})_{\rm T}$. This also agrees with the finding of Thomas et al. (2015), see their figure 7. There they find that $\nabla\times\mathbf{B}$ is mainly generated by $\nabla\delta\times\mathbf{v}$, and the contribution from vorticity is many orders of magnitude smaller. The B-power spectra from both simulations also agree qualitatively, see figure 4 of Thomas et al. (2015). Note that in this reference, the dimensionless power spectra are shown which differ from ours by a factor k^3.

We now define the correlation coefficient for two generic quantities A and B as

$$\rho_{\rm AB} \equiv \frac{P_{\rm AB}}{\sqrt{P_{\rm A}P_{\rm B}}}\,. \tag{7.30}$$

The correlation coefficients for the rotational velocity $\mathbf{v}_R$ and the metric vector perturbation $\mathbf{B}$, the density contrast δ and velocity divergence θ, and the velocity dispersion σ^2 and velocity divergence θ are displayed in figure 7.11 (right panel) at $z = 0$ and $z = 1$.

As discussed before, both the density contrast and the velocity dispersion are anti-correlated with θ on large scales. We also notice that the scale at which the correlation coefficients for $\delta\times\theta$ becomes positive depends on redshift. This is due to the fact that the regions that experience shell-crossing are larger at low redshift and, therefore, the velocity divergence becomes positive at larger scales. The sign change of the $\theta\times\delta$ and $\theta\times\sigma^2$ cross spectra are roughly at the same scale which is somewhat smaller than the scale of the peak of the vorticity spectrum. Below this scale density and divergence are weakly correlated and positive which indicates a slight net outflow instead of inflow around structures. On large scales, the velocity dispersion is largest where the density is highest, hence clearly also σ^2 and θ are anti-correlated. This anti-correlation nearly vanishes and for $z = 0$ even becomes slightly positive on the scales where further inflow is prohibited by angular momentum conservation.

The correlation coefficient between $\mathbf{B}$ and the rotational velocity is roughly 10^{-7} on all scales and therefore not visible on this linear plot. As discussed before, this means that the main source for the metric perturbation does not come from the rotational velocity alone but from the transverse part of $(\delta\,\mathbf{v})$.

7.4 Summary

In this chapter we numerically investigate the generation of vorticity from a set of N-body simulations. We have implemented the computation of the velocity field and its dispersion in the relativistic N-body code *gevolution*. The velocity field is estimated from the momentum and density fields, which are computed in *gevolution* using a particle-to-mesh projection. In order to validate our method, we have tested several prescriptions to deal with the empty regions of the simulations: an identically zero velocity field in voids; a Gaussian smoothing for the density and momentum field, from which the velocity is computed; a velocity field which does not evolve in voids and a velocity field which evolves in the empty regions according to the linear growth rate. All these methods give very similar results for the vorticity power spectrum. However, the main results presented in our work are obtained with the latter method, which is also physically well motivated.

The key findings can be summarized as follows:

(a) On large sub-horizon scales, the vorticity power spectrum is well modelled by a power-law with a spectral index $n_\omega \approx 2.5$. This result quantitatively agrees with the analysis presented in (Hahn et al. 2015). Causality actually requires $n_\omega = 2$ on the largest scales, but we could not see this in our simulations. In fact, we expect some change of slope around the horizon scale which, however, is not represented in our simulations because of the choice of box size.

(b) At redshift $z = 0$ the vorticity power spectrum peaks at $k_p(z = 0) \simeq 1h/\mathrm{Mpc}$. At roughly this scale also the divergence power spectrum has a dip and the density $-$ divergence correlation changes sign, a clear indication of shell crossing. The scale $k_p^{-1}(z)$ decays with increasing redshift to $k_p(z = 1) \simeq 2h/\mathrm{Mpc}$.

(c) The vorticity power spectrum in our simulations depends both on the mass resolution (number of particles per cell) and grid resolution (number of grid points). We need at least 8 particles per grid cell and a grid resolution $\Delta x = 125\ \mathrm{kpc}/h$ in order for the vorticity power spectrum to converge at $z \in [0, 1]$.

(d) The grid resolution that is needed in order to have convergence depends on redshift. At higher redshift shell-crossing is a rarer phenomenon than at $z = 0$ and it happens on smaller scales. Therefore, a finer grid resolution is required in order to resolve shell-crossing and disentangle it from numerical velocity dispersion that is always present in our simulations due to finite resolution.

(e) In the redshift range $z \in [0, 1]$, where our results are reliable, the vorticity power spectrum grows as $(\mathcal{H}f)^2 D^\gamma(z)$, with $\gamma \approx 7$. This result is consistent with (Pueblas & Scoccimarro 2009; Thomas et al. 2015).

(f) The amplitude of the vorticity spectrum at the peak position increases from $P_\omega/(\mathcal{H}f)^2(k_p) \approx 0.6\,(\mathrm{Mpc}/h)^3$ at $z = 1$ to $P_\omega/(\mathcal{H}f)^2(k_p) \approx 5\,(\mathrm{Mpc}/h)^3$ at $z = 0$. Roughly at this peak position the divergence power spectrum drops significantly below the linear perturbation theory result and the vorticity becomes the dominant contributor to the total velocity spectrum.

(g) The rotational part of the velocity is very little correlated with the vector perturbations of the metric. Therefore, we expect the metric perturbations $\mathbf{B}$ to be predominantly sourced by the combination of δ and the velocity field, $(\delta\mathbf{v})_T$ and not by the velocity field alone.

Observationally the presence of vorticity will have interesting consequences e.g. for redshift space distortions. Our results indicate that the best place to look for them is around the Mpc scale which is also the scale where we expect predictions from linear theory to become unreliable. It therefore remains a challenge to tackle this problem, and numerical simulations will probably continue to play an important role in this context.

Chapter 8

CONCLUSIONS AND OUTLOOK

Future planned LSS instruments will be able to map the distribution of galaxies and other tracers, such as the 21 cm emission from neutral hydrogen, on a wide range of scales, from the Mpc scale, where the physics is highly non-linear, up to hundreds of Mpc, where instead linear relativistic effects beyond the standard RSD may become relevant. In this thesis we have presented the results of several studies in the field of Large Scale Structure, whose general aim is to improve the current physical modelling for the LSS observables both in the linear and non-linear regime. In this section I will summarize the main results and draw possible developments of this work.

Main results

Part II of this thesis deals with the largest scales that we will test with the upcoming redshift surveys. The goal of this part is threefold. First, test the benefits that future cosmological analysis will gain by performing model-independent tests, i.e. by performing analysis based on observables quantities such as the angular power spectrum or the angular correlation function rather than the standard Fourier power spectrum, which requires a reference cosmology to be assumed to convert redshift and angular coordinates into comoving distances. Second, to test the accuracy of the standard LSS analysis, which includes density and Kaiser effect, for future clustering studies.

These two points have been investigated in chapters 4 and 5. In chapter 4 we investigate the Alcock Paczyński test applied to the baryon acoustic oscillation feature in the galaxy correlation function. The method that is employed in this chapter is model independent: the statistical analysis is based on the measurement in the BAO peak from the radial and the angular correlation function, which depends only on observables coordinates. By using a general formalism that

includes relativistic effects, we quantified the importance of the linear redshift space distortions and gravitational lensing corrections to the galaxy number density fluctuation. We have shown that redshift space distortions significantly affect the shape of the correlation function, both in the radial and transverse directions, causing different values of galaxy bias to induce offsets up to 1% in the AP test. On the other hand, we have found that the lensing correction around the BAO scale modifies the amplitude but not the shape of the correlation function and therefore does not introduce any systematic effect. Therefore, we have shown the robustness of BAO measurements with respect to lensing corrections.

In chapter 5 we study the impact of neglecting lensing magnification in galaxy clustering analysis, considering the ΛCDM model and two extensions: massive neutrinos and modifications of General Relativity. Our main focus was the biases on the constraints and on the estimation of the cosmological parameters. We performed a comprehensive investigation of these two effects for the upcoming photometric and spectroscopic galaxy surveys Euclid and SKA for different redshift binning configurations. Our results show that the information present in the lensing contribution does improve the constraints on the modified gravity parameters whereas the lensing constraining power is negligible for the ΛCDM parameters. For photometric surveys the estimation is biased for all the parameters if lensing is not taken into account. This effect is particularly significant for the modified gravity parameters. Conversely for spectroscopic surveys the bias is below 1-σ for all the parameters. Our findings show the importance of including lensing in galaxy clustering analyses for testing General Relativity and to constrain the parameters which describe its modifications, especially for future photometric surveys.

The third direction that has been taken in the first part of thesis is investigating the future perspectives of detecting the Doppler relativistic effects. In fact, their measurement will provide a test for the Euler equation and, therefore, one of the foundational principles of GR: the equivalence principle (Bonvin & Fleury 2018). In chapter 6 we focus on the relativistic dipole in the cross-correlation between the 21 cm emission in Intensity Mapping (IM) and galaxy surveys at low redshift. Neutral hydrogen (HI) and the galaxy population are modelled by means of the halo model to relate the parameters that affect the dipole signal such as the biases of the two tracers and the Poissonian noise. We studied the behavior of the signal-to-noise as a function of the galaxy and magnification biases, for two fixed models of the neutral hydrogen. In both cases we have found that the signal-to-noise does not grow by increasing the difference between the biases of the two tracers, due to the larger shot-noise yields by highly biased tracers. We also studied and provided

an optimal luminosity-threshold galaxy catalogue to enhance the signal-to-noise ratio of the relativistic dipole. Interestingly, we show that the maximum magnitude provided by the survey does not lead to the maximum signal-to-noise for detecting relativistic effects and we predict the optimal value for the limiting magnitude. Our work suggests that an optimal analysis could increase the signal-to-noise ratio up to a factor five compared to a standard one.

Part III of the thesis focused on the non-linear regime of LSS. Modelling the non-linear regime of large scale structure is a very complicated task, since perturbation theory breaks down at these scales and baryon physics becomes relevant. In this thesis we have focused on the generation of vorticity in a pure dark matter scenario and we investigate this topic with N-body simulations based on the relativistic code *gevolution*. We estimated the amplitude and the spectral index of the vorticity power spectrum and studied its redshift evolution and its relation with the vector perturbation in the metric. The vorticity power spectrum is particularly relevant for a correct modelling of the redshift space distortions, since the peculiar velocity field is generally assumed to have only a scalar degree of freedom, while our findings show that this assumption breaks down at the Mpc scales.

Future perspectives

Here we will draw some possible extensions of this work and some future perspectives in the field.

The lensing correction, which has been studied in chapter 4 and 5, has been proved to be relevant for forthcoming galaxy surveys. Nevertheless, it can be seen not just a contaminant for the observable, but as a signal itself. The lensing magnification was detected for the first time from SDSS data by cross-correlating quasars and their foreground galaxies (Scranton et al. 2005). It will be interesting to study if future measurements of the lensing magnification from galaxy surveys like Euclid and the Square Kilometre Array will be competitive or complementary with weak lensing analysis. In fact measurements of the lensing magnification suffer from different systematics and are not affected by intrinsic alignment (Montanari & Durrer 2015).

In chapter 6 we investigated the relativistic dipole, sourced by the gravitational redshift and Doppler corrections to the galaxy number count and 21 cm emission flux fluctuations at low redshift. It would be interesting to perform a similar analysis for other tracers and redshift range. In the high-redshift regime ($z \sim 2.5$) a very promising combination of tracers are the Lyman-α forest fluctuations cross-correlated with quasars (QSOs). In fact, the two tracers have a large difference in bias: the Lyman-α forest consists of a series of absorption lines and, therefore, its

bias is negative, while QSOs are characterized by a relatively high bias ($b_{\mathrm{QSOs}} \approx 4$). The state-of-the-art measurements of the Ly-α-QSOs cross-correlation currently do not model the Doppler contribution to the signal. However, the signal presents a not-negligible asymmetric component due to metal contaminations (see (du Mas des Bourboux et al. 2017)). In (Iršič et al. 2016) the relativistic contributions have been computed for the radial cross-correlation. We are currently working on the generalization of this work to the full 3-dimensional case, in order to perform realistic forecast, including the metal contaminations, for the detection of the relativistic effects with the current surveys BOSS and eBOSS and for the future survey DESI.

Our work on the large-scale vorticity has shown that on the Mpc scales, the vorticity power spectrum becomes comparable in magnitude to the power spectrum of the velocity divergence. A possible development of this work is to estimate the error that neglecting the rotational part of the velocity introduces for the standard redshift-space distortions analysis and in the future bispectrum measurements. Furthermore, in our analysis we employ dark matter only simulations. On the Mpc scales, baryonic physics comes into play and, therefore, a possible development is to study the impact of baryons on the vorticity generation. This would requires the implementation of our method for computing the vorticity spectrum in a N-body and hydrodynamical code.

Another interesting aspect for the future will be investigating the non-linear regime of the relativistic effects with the relativistic N-body code *gevolution*. This is the ground where the part II and III of this thesis meet. Gravitational redshift detection has been claimed on galaxy clusters scales in (Wojtak et al. 2011; Sadeh et al. 2015) and, more recently, a line-of-sight asymmetry has been detected in the cross-correlation of two galaxy populations in the non-linear regime on large scale structure, i.e. on scales ~ 10 Mpc (Alam et al. 2017a). The physical interpretation of the cluster measurements has been debated in the literature (Kaiser 2013; Zhao et al. 2013), and in general the theoretical predictions for the relativistic distortions in the non-linear regime are not well-established yet. Numerical simulations are playing a crucial role to understand how to physically interpret these measured effects (Zhu et al. 2017a; Breton et al. 2018) and using a relativistic code for this purpose will certainly give a valuable contribution. Finally, on a more general note, the non-linear regime of the LSS in a partial or fully relativistic framework is a promising field of research, which has not been fully explored yet. Currently, there are three main approaches to this subject:

- The most conservative approach consist of reinterpreting the Newtonian simulations by choosing a specific gauge, the *N-body gauge*, in which particles follow the Newtonian trajectories (Fidler et al. 2016). In this gauge there is a one-to-one correspondence between the Newtonian dynamics and the GR dynamics, at first order. Therefore, this method does not modify the Newtonian dynamics, already implemented in the Newtonian N-body codes, but requires to set the initial conditions in the appropriate gauge and to post-process the output of the Newtonian simulation in order to extract the relativistic observables (see also (Fidler et al. 2017, 2018)).

- The N-body code *gevolution*(Adamek et al. 2016a,b), that was employed in the part III of this thesis, provides a self-consistent way to study the structure formation, also in presence of relativistic sources. This approach, as well as the method outlined in the previous paragraph, assume the weak-field approximation. Therefore, it does not encode the fully non-linear relativistic dynamics. However, the weak-field approximation should work fairly well in the cosmological context, as long as the structures under investigation are larger than their Schwarzschild radius.

- The most challenging approach is to develop a fully-relativistic code for cosmological applications. In this direction, there have been several attempts in the literature to take advantage of the numerical relativity tools that have been developed in the past decades and successfully applied in the context of compact objects (Baumgarte & Shapiro 2010; Alcubierre 2008; Gourgoulhon 2007; Bona & Palenzuela-Luque 2005; Luciano & Zanotti 2013). This formalism has been applied in order to investigate the cosmological structure formation (Bentivegna & Bruni 2016; Giblin et al. 2017; Macpherson et al. 2018). At present, the main limit of this approach is that the stress-energy tensor is modelled as a fluid, which for the pressureless dark matter component leads to the formation of caustics and numerical divergences.

The first two methods adopt a particle description for the matter content, which is preferable for modelling collisionless dark matter. However, they do not encode the full non-linear relativistic dynamics. Furthermore, it is still not established whether the two approaches should exactly coincide. It would be certainly interesting to understand weather the full-GR approach can incorporate a particle description for the stress-energy tensor and to compare the fully relativistic dynamics to the outcome in the weak-field approximation.

Appendices

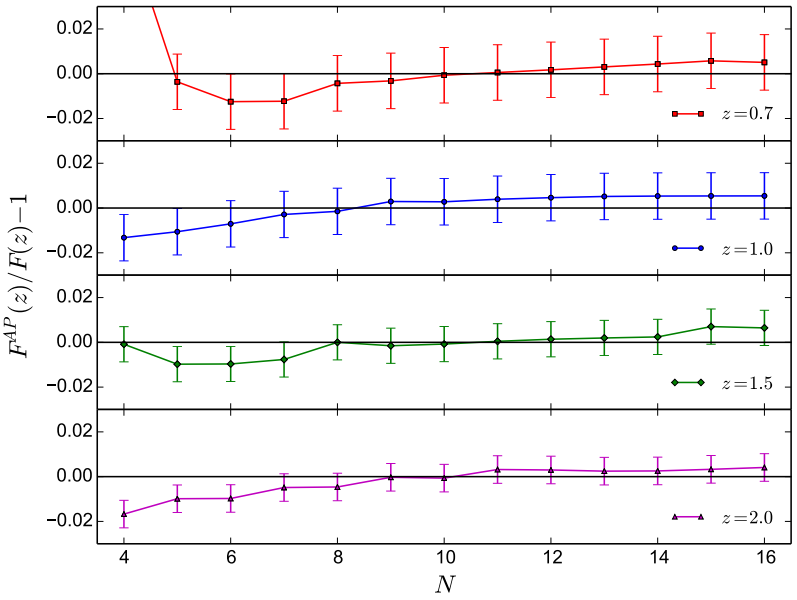


Figure A.1: AP consistency test for different degrees N of the polynomial in the parameterization of the correlation function.

Appendix A

THE ALCOCK PACZYŃSKI TEST WITH BARYON ACOUSTIC OSCILLATIONS: SYSTEMATIC EFFECTS FOR FUTURE SURVEYS

A.1 Methodology tests

Here we report the results of some tests we performed in order to find the best method to model the transversal and longitudinal correlation functions and in order to verify its reliability. We test our template model, equation (4.19), for different

degrees of the fitting polynomial. Figure A.1 shows a consistency check for values of N in the range $[4:16]$. The correlation functions used in this test are computed including the local density term and the redshift space distortions correction. We found that for N between 8 and 14 the relative difference between the fitted AP parameter and its theoretical value is smaller than 0.5% at all redshifts. We checked the accuracy of the fit comparing the sum of the squared deviations between our fit and the CLASSgal output and we found the most numerically accurate results satisfying the AP test. Therefore, we choose a polynomial of degree $N = 10$ to perform the next analysis. In order to test the goodness of this parameterization, we perform the same analysis for different cosmologies, including a cosmology with a dynamical dark energy equation of state. In all cases we find qualitatively the same behavior shown in figure A.1. We compare our parameterization also with others described in the literature. In particular, the transverse correlation function is often modeled by a *power-law + Gaussian* function (Sanchez et al. 2011)

$$\xi_{\perp}(\theta) = A + B\theta^{\gamma} + C \cdot e^{-(\theta-\theta_{BAO})^2/2\sigma^2}, \tag{A.1}$$

with A, B, γ, C, θ_{BAO} and σ as free parameters. As for the radial correlation function, we compare our model with the parameterization used in (Sanchez et al. 2013)

$$\xi_{\parallel}(\Delta z) = A + Be^{-C\Delta z} - De^{-E\Delta z} + F \cdot e^{-(\Delta z-\Delta z_{BAO})^2/2\sigma^2}, \tag{A.2}$$

where free parameters are A, B, C, D, E, F, $\Delta z_{\rm BAO}$ and σ. The result of this consistency test is shown in figure A.2. We find that both the exponential parameterization, equation (A.2), and the 10 degree polynomial, equation (4.19), fit well the radial correlation function (differences are smaller than 0.5% at all redshifts for the AP consistency check). On the other hand, we find a discrepancy between the power-law and the polynomial parameterization for the transverse correlation function. In table A.1 we compare the comoving BAO scale computed from the peak positions in the transverse and radial direction. We find that, using a *power-law+Gaussian* fit for the transverse correlation function the peak position is systematically shifted towards larger scales. This systematic introduces a 2% offset in the AP test. The polynomial and the exponential parameterization gives consistent values of the BAO scale at all redshifts. However, the polynomial fit is able to model the data over a larger range, therefore we adopt this parameterization.

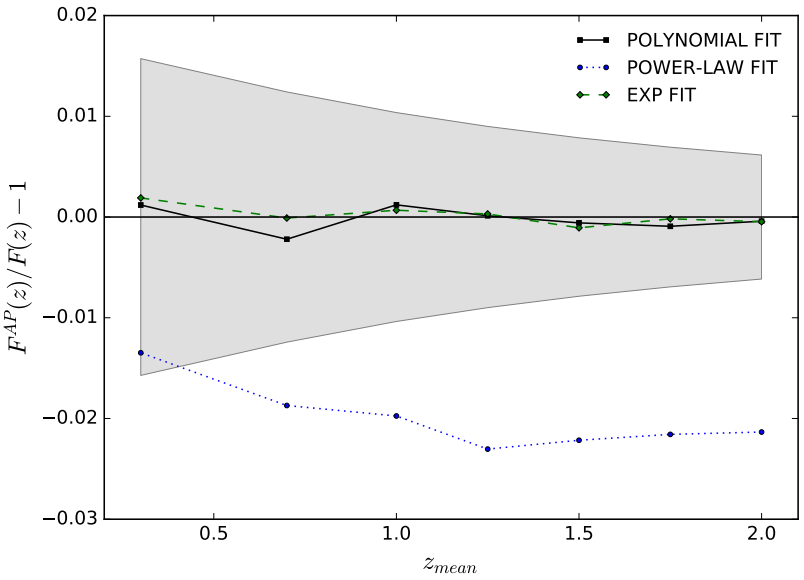


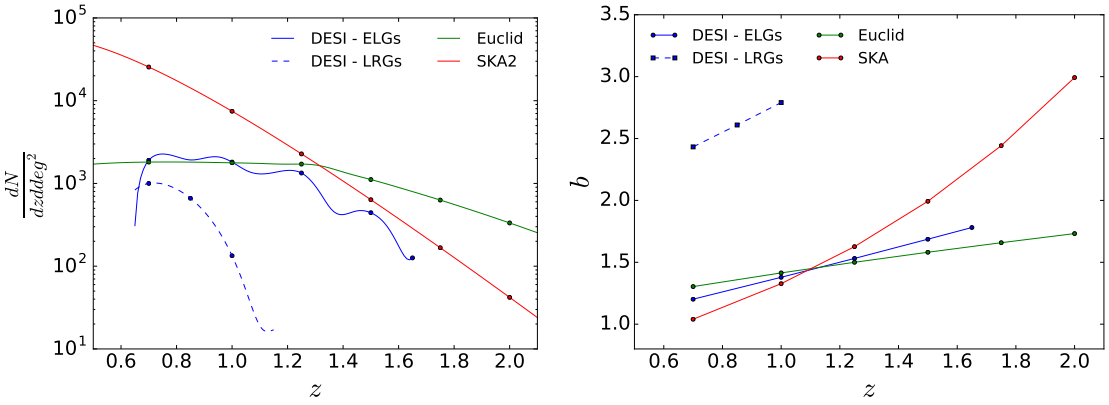
Figure A.2: AP consistency test at different redshifts, for different parameterization of the radial and transverse correlation functions. The continuous black line refers to the test performed using a polynomial fitting function for the estimation of both $\Delta z_{\rm BAO}$ and $\theta_{\rm BAO}$, the dot blue line is computed using a polynomial fit in the longitudinal direction and the *power-law+Gaussian* model in the transverse one, while the dash green line is computing using a polynomial for $\xi_{\perp}(\theta)$ and the exponential fit in (A.2) for $\xi_{\parallel}(\Delta z)$.

z	$L_{\perp}(Mpc/h)$		$L_{\parallel}(Mpc/h)$	
	Polynomial	Power-law	Polynomial	Exponential
0.3	101.7 ± 0.2	103.2 ± 0.2	101.9 ± 1.6	102.0 ± 1.6
0.7	101.9 ± 0.3	103.6 ± 0.3	101.7 ± 1.3	101.9 ± 1.3
1.0	101.8 ± 0.4	103.9 ± 0.4	101.9 ± 1.1	101.8 ± 1.1
1.25	101.8 ± 0.5	104.3 ± 0.5	101.9 ± 0.9	101.9 ± 0.9
1.5	101.9 ± 0.5	104.1 ± 0.5	101.8 ± 0.8	101.8 ± 0.8
1.75	101.9 ± 0.6	104.0 ± 0.6	101.8 ± 0.7	101.8 ± 0.7
2.0	101.9 ± 0.6	104.0 ± 0.6	101.8 ± 0.6	101.8 ± 0.6

Table A.1: The comoving BAO scale in units of Mpc/h computed using the different parameterizations given in the text for the correlation function at different redshifts.

A.2 Surveys specifications

Here we report the specifications used for the different surveys considered in section 4.5.3. The parameters of the surveys are the sky coverage $f_{\rm sky}$, the number of galaxies per unit redshift per square degree and the redshift dependence of the galaxy bias.



(a) Number of galaxies per square degree per unit redshift.

(b) Redshift evolution of the galaxy bias.

Figure A.3: Figure a and figure b show, respectively, the number of galaxies per square degree per unit redshift and the redshift dependent bias for the four cases described in this appendix: Euclid, SKA, DESI ELGs (bright Emission Line Galaxies) and DESI LRGs (Luminous Red Galaxies). The markers identify the redshifts for which we performed the AP test in section 4.5.3.

A.2.1 Euclid

Our Euclid-like survey is modeled following appendix A.3 in (Audren et al. 2013). We assume a sky fraction $f_{\rm sky} = 0.375$. The number of galaxies per unit redshift per square degree is computed from table 2 of (Geach et al. 2010), for the case of a limiting flux of 3×10^{-16}erg s^{-1} cm^{-2}. Following Ref. (Audren et al. 2013), we multiply the tabulated values by an efficiency factor 0.25 and we divide them by the factor 1.37 to get conservative prediction. The redshift dependence of the bias is modeled as in the forecasts presented in (Amendola et al. 2013)

$$b(z) = \sqrt{1+z}. \tag{A.3}$$

A.2.2 SKA

We used the technical specification for SKA reported in (Abdalla et al. 2015). Galaxy number density per unit redshift per square degree and bias evolution are given by

$$\frac{dN}{dzd\Omega} = 10^{c_1} z^{c_2} \exp\left(-c_3 z\right), \tag{A.4}$$

$$b(z) = c_4 \exp\left(c_5 z\right), \tag{A.5}$$

where we used for the coefficients c_i the best-fit values reported in table 4 of (Santos et al. 2015), for SKA2. For SKA2, the sky coverage is expected to be around 30000 deg^2, which corresponds to a sky fraction $f_{\rm sky} = 0.727$.

A.2.3 DESI

Survey specifications for DESI, for both ELGs and LRGs, are taken from the Science Technical Design Report (DESI 2015). The number of galaxies per unit redshift per square degree, for both ELGs and LRGs, are assumed to be the one reported in table 2.3 of (DESI 2015). The values at intermediate redshifts are computed by interpolation. The redshift dependence of galaxy bias is expressed in terms of the linear growth factor $D(z)$. For the ELGs we have

$$b(z) = \frac{0.84}{D(z)}, \tag{A.6}$$

while for LRGs we assume

$$b(z) = \frac{1.7}{D(z)}. \tag{A.7}$$

Following (DESI 2015), we assume for DESI a sky coverage equal to $14000\,\mathrm{deg}^2$, which corresponds to a sky fraction $f_{\mathrm{sky}} = 0.339$. In figure A.3 we compare the expected number of galaxies per unit redshift, per square degree and the bias evolution for the 3 cases discussed above. The number of galaxies per unit redshift per square degree is relevant for the computation of shot noise.

The number of galaxies per steradian n_i inside the i-th redshift bin is computed as

$$n_i = \frac{1}{4\pi} \int d\Omega \int_{z_{mean}-\Delta z/2}^{z_{mean}+\Delta z/2} \frac{dN}{dz d\Omega} dz, \tag{A.8}$$

where Δz is the width of the redshift bin.

Appendix B

OPTIMAL GALAXY SURVEY FOR DETECTING THE DIPOLE IN THE CROSS-CORRELATION WITH 21 CM INTENSITY MAPPING

B.1 Flat-sky versus full-sky dipole

In this appendix we test the validity of the flat-sky approximation by comparing the flat-sky and the full-sky relativistic dipole at redshift $z_m = 0.15$. This comparison is crucial to test in which regime we can apply the flat-sky approximation and therefore which are the maximum scales that we can model in this framework. In particular, we test the dipole expansion in terms of d/r which, to lowest order, leads to the wide-angle correction, defined in equation (6.20), that we adopted in our analysis.

The full-sky correlation function for the cross-correlation of two populations of galaxies has been studied in detail in the literature, both in the newtonian approximation (Szalay et al. 1998; Szapudi 2004; Papai & Szapudi 2008) and beyond (Raccanelli et al. 2010; Jeong et al. 2012; Samushia et al. 2012; Bertacca et al. 2012; Bonvin et al. 2014; Raccanelli et al. 2014; Yoo & Seljak 2015; Tansella et al. 2017). Here we will use the same notation as (Bonvin et al. 2014), where the full-sky relativistic correlation function is computed for two galaxy populations, including the gravitational redshift and Doppler effects, which are the quantity relevant for this work (see the appendix B in (Bonvin et al. 2014) for a detailed derivation of the relativistic part, i.e. the cross-correlation between the density plus redshift space distortion and the Doppler terms in the galaxy number counts). We will apply that expression to the cross-correlation between HI and galaxies. We will fix the values of the bias to $b_{\rm HI} = 0.6$ and $b_{\rm gal} = 1.0$, and we will compare three values of the galaxy magnification bias.

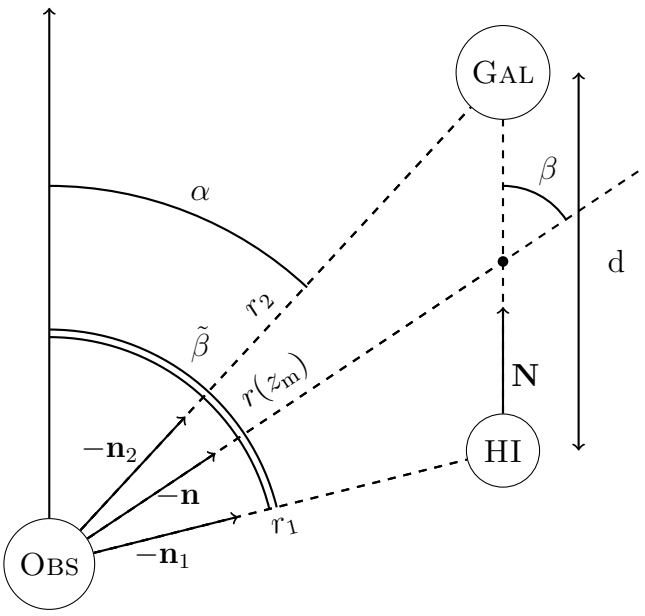


Figure B.1: Coordinate system for the full-sky relativistic dipole.

We will recast the expression for the full-sky correlation function in the way described below. The relativistic correlation function, in the coordinate system defined in section 6.2.3 (see also figure B.1), can be written in terms of some coefficients $R_i = R_i(r_1, r_2, d)$

$$\begin{aligned}\xi^{\mathrm{HI,gal}}(r,d,\mu) &= R_1\cos(\alpha) + R_2\cos(\tilde{\beta}) + R_3\cos(\alpha)\cos(2\tilde{\beta}) + R_4\cos(\tilde{\beta})\cos(2\alpha) \\ &\quad + R_5\sin(\alpha)\sin(2\tilde{\beta}) + R_6\sin(\tilde{\beta})\sin(2\alpha), \end{aligned} \tag{B.1}$$

where α is defined as $\cos(\alpha) = -\mathbf{n_2}\cdot\mathbf{N}$ and $\cos(\tilde{\beta}) = -\mathbf{n_1}\cdot\mathbf{N}$ (see figure B.1),

while the coefficients R_i are

$$\begin{aligned}
R_1 =& C_{\rm gal}(r_2)\frac{D_1(r_2)D_1(r_1)}{D_1^2(r)}f(r_2)\left[\left(b_{\rm HI}+\frac{2}{5}f(r_1)\right)\nu_1-\frac{1}{10}f(r_1)\nu_3\right],\\
R_2 =& -C_{\rm HI}(r_1)\frac{D_1(r_2)D_1(r_1)}{D_1^2(r)}f(r_1)\left[\left(b_{\rm gal}+\frac{2}{5}f(r_2)\right)\nu_1-\frac{1}{10}f(r_2)\nu_3\right],\\
R_3 =& C_{\rm gal}(r_2)\frac{D_1(r_2)D_1(r_1)}{D_1^2(r)}f(r_1)f(r_2)\frac{1}{5}\left[\nu_1-\frac{3}{2}\nu_3\right],\\
R_4 =& -C_{\rm HI}(r_1)\frac{D_1(r_2)D_1(r_1)}{D_1^2(r)}f(r_1)f(r_2)\frac{1}{5}\left[\nu_1-\frac{3}{2}\nu_3\right],\\
R_5 =& C_{\rm gal}(r_2)\frac{D_1(r_2)D_1(r_1)}{D_1^2(r)}f(r_1)f(r_2)\frac{1}{5}\left[\nu_1+\nu_3\right],\\
R_6 =& -C_{\rm HI}(r_1)\frac{D_1(r_2)D_1(r_1)}{D_1^2(r)}f(r_1)f(r_2)\frac{1}{5}\left[\nu_1+\nu_3\right],
\end{aligned} \tag{B.2}$$

where $C_{\rm gal}$ and $C_{\rm HI}$ are defined in (6.14), D_1 denotes the linear growth factor, and $\nu_{\ell=1,3}$ are defined through

$$\nu_\ell(d)=\int\frac{k^2dk}{2\pi^2}\left(\frac{\mathcal{H}}{k}\right)P(k,z_m)j_\ell(k\,d). \tag{B.3}$$

All the quantities involved in the computation of the correlation function in (B.1) can be expressed as a function of r, d and μ by using the following simple geometric relations

$$\begin{aligned}
r_1 &= \frac{1}{2}\sqrt{d^2+4r^2-4d\mu r}\,, & r_2 &= \frac{1}{2}\sqrt{d^2+4r^2+4d\mu r}\,,\\
\cos(\tilde\beta) &= \frac{-d+2r\mu}{\sqrt{4r^2+d^2-4dr\mu}}\,, & \cos(\alpha) &= \frac{d+2r\mu}{\sqrt{4r^2+d^2+4dr\mu}}\,.
\end{aligned} \tag{B.4}$$

The dipole can then be computed as

$$\xi_1^{\text{full sky}}=\frac{3}{2}\int_{-1}^{1}\xi^{\rm HI,gal}(r,d,\mu)L_1(\mu)d\mu. \tag{B.5}$$

We see the full-sky correlation function depends in a not trivial way from the angular coordinate μ. Therefore, the angular integral in (B.5) cannot be performed analytically, as in the flat-sky approximation, but it needs to be solved numerically.

A similar computation can be done for the wide-angle corrections. The full sky expression can be written in terms of some coefficients $S_i = S_i(r_1, r_2, d)$ (Bonvin et al. 2014)

$$\xi^{\rm WA}(r,d,\mu)=S_1+S_2\cos(2\tilde\beta)+S_3\cos(2\alpha)+S_4\cos(2\alpha)\cos(2\tilde\beta)+S_5\sin(2\tilde\beta)\sin(2\alpha), \tag{B.6}$$

$$
\begin{aligned}
S_1 =& \frac{D_1(r_2)D_1(r_1)}{D_1^2(r)} f(r_2)\left[\left(b_{\rm HI}b_{\rm gal} + \frac{b_{\rm HI}}{3}f(r) + \frac{b_{\rm gal}}{3}f(r) + \frac{2}{15}f^2(r)\right)\mu_0,\right.\\
&\left. - \frac{1}{3}\left(\frac{b_{\rm HI}}{2}f(r) + \frac{b_{\rm gal}}{2}f(r) + \frac{2}{7}f^2(r)\right)\mu_2 + \frac{3}{140}f^2(r)\mu_4\right],\\
S_2 =& -\frac{D_1(r_2)D_1(r_1)}{D_1^2(r)}\left[\left(\frac{b_{\rm gal}}{2}f(r) + \frac{3}{14}f^2(r)\right)\mu_2 - \frac{1}{28}f^2(r)\mu_4\right],\\
S_3 =& -\frac{D_1(r_2)D_1(r_1)}{D_1^2(r)}\left[\left(\frac{b_{\rm HI}}{2}f(r) + \frac{3}{14}f^2(r)\right)\mu_2 - \frac{1}{28}f^2(r)\mu_4\right],\\
S_4 =& \frac{D_1(r_2)D_1(r_1)}{D_1^2(r)} f(r)^2\frac{1}{5}\left[\mu_0 - \frac{1}{21}\mu_2 + \frac{19}{140}\mu_4\right],\\
S_5 =& \frac{D_1(r_2)D_1(r_1)}{D_1^2(r)} f(r)^2\frac{1}{5}\left[\mu_0 - \frac{1}{21}\mu_2 - \frac{4}{35}\mu_4\right],
\end{aligned}
\tag{B.7}
$$

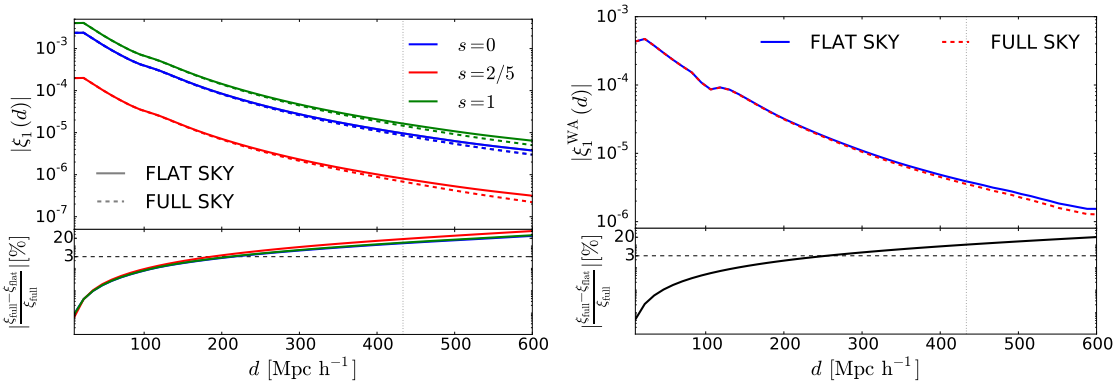
where $\mu_{\ell=0,2,4}$ are defined by

$$
\mu_\ell(d) = \int \frac{k^2 dk}{2\pi^2} P(k, z_m) j_\ell(k\, d). \tag{B.8}
$$

The dipole of the wide-angle effect, in full sky, can be computed as

$$
\xi_1^{\text{WA, full sky}} = \frac{3}{2}\int_{-1}^{1} \xi^{\rm WA}(r, d, \mu) L_1(\mu) d\mu. \tag{B.9}
$$

In figure B.2 we compare the full-sky dipole (dashed lines) and the corresponding dipole in the flat-sky approximation (continuous lines). Figure B.2a represents the Doppler dipole, computed from (B.5) (full-sky) and (6.11) (flat-sky), while figure B.2b shows the wide-angle dipole, computed from (B.9) (full-sky) and (6.19) (flat-sky). Different colors refer to different values of the magnification bias. In the top panel we plot the absolute value of the dipole. At small scale, the flat-sky approximation fairly reproduces the full-sky signal, while the full-sky dipoles significantly departs from the one in flat-sky at large scales. Note that the full-sky signal always results to be smaller, in absolute value, than the approximated one. Indeed two sources appear closer in the flat-sky limit and therefore they seems to be more correlated. In the bottom panel we plot the relative difference between the full-sky and the flat-sky dipole, in percentage. The black dashed line denotes a 3% difference between the two quantities, which we set as a threshold for the flat-sky approximation to be valid. We see that for scales larger than roughly 200Mpc/h the relative difference is beyond the threshold value and that it rapidly



(a) Doppler dipole. **(b)** Wide-angle dipole.

Figure B.2: Comparison between the dipole of the cross-correlation HI-galaxies (right panels) and the wide-angle dipole (left panels) computed within the flat-sky approximation (continuous line) and the full-sky dipole signal (dashed line). Top panels: absolute value of the dipole. Bottom panels: relative difference between the full signal and the flat-sky approximation. The dashed line denotes a difference of relative difference of 3%. Different colors denote different values of the galaxy magnification bias. The dotted vertical lines denotes the distance $d = r$.

increases at larger scales. Nevertheless the amount of physical information at these scales is irrelevant as we see from figure 6.10.

In figure B.3 we show the difference between the estimator defined in (6.20), which is unbiased from wide-angle effect, in full-sky and the one in flat-sky, at lower order in the d/r expansion. We see that at scales $d \leq 200\,\mathrm{Mpc/h}$ the approximated dipole fairly reproduces the full-sky quantity. Therefore, we set the maximum scale of our analysis to be $d_{\mathrm{max}} = 200\,\mathrm{Mpc/h}$.

B.2 Contribution to the wide-angle correction to the covariance

The estimator we used in the signal-to-noise analysis, which is unbiased from wide-angle effects (at least at the leading order in the d/r expansion), involved the measurement of both the dipole of the cross-correlation of galaxies and HI and the quadrupole of the autocorrelations of both tracers

$$\hat{\xi}_1(d,r) = \hat{\xi}_1(d,z_1) - \frac{3}{10}\left(\hat{\xi}_2^{gal}(d,z_1) - \hat{\xi}_2^{\mathrm{HI}}(d,z_1)\right)\frac{d}{r}. \tag{B.10}$$

We split the covariance for the estimator in equation (B.10) in two contributions [1],

$$\mathrm{COV}^{\mathrm{est}}(d_1,d_2) = \mathrm{COV}^{\mathrm{dip}}(d_1,d_2) + \mathrm{COV}^{\mathrm{WA}}(d_1,d_2), \tag{B.11}$$

[1] We assume the redshift to be constant and we drop the redshift dependence from our notation

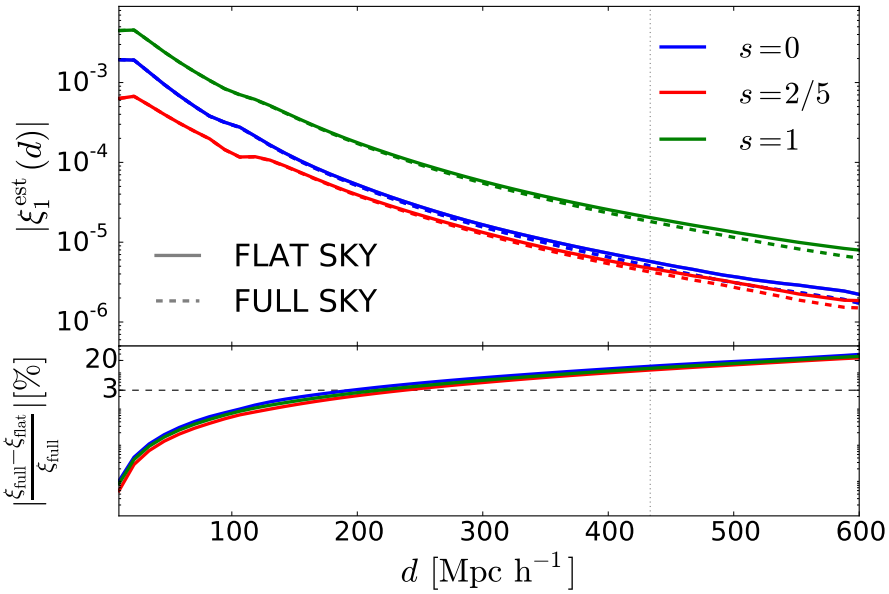


Figure B.3: Comparison between the estimator of the cross-correlation HI-galaxies computed within the flat-sky approximation (continuous line) and the full-sky dipole signal (dashed line). Top panel: absolute value of the estimator. Bottom panel: relative difference between the full signal and the approximated signal. Different colors denote different values of the galaxy magnification bias.

where $\mathrm{COV}^{\rm dip}(d_1,d_2) = \langle\hat{\xi}_1(d_1)\hat{\xi}_1(d_2)\rangle$ is the dipole contribution to the estimator it is given by equation (6.21), while $\mathrm{COV}^{\rm WA}(d_1,d_2)$ is the contributions of the wide-angle correction,

$$\begin{aligned}\mathrm{COV}^{\rm WA}(d_1,d_2) = \; & \left(\tfrac{3}{10r}\right)^2 d_1\, d_2 \left(\mathrm{COV}_{2,2}^{\rm gal,\ gal} + \mathrm{COV}_{2,2}^{\rm HI,\ HI} + 2\,\mathrm{COV}_{2,2}^{\rm gal,\ HI}\right) \\ & - \tfrac{3}{10}\tfrac{d_1+d_2}{r}\left(\mathrm{COV}_{1,2}^{\rm Dip,\ gal} - \mathrm{COV}_{1,2}^{\rm Dip,\ HI}\right), \end{aligned} \tag{B.12}$$

where

$$\begin{aligned}\mathrm{COV}_{2,2}^{\rm gal,\ gal}(d_1,d_2) &= \left\langle\hat{\xi}_2^{gal}(d_1)\hat{\xi}_2^{gal}(d_2)\right\rangle - \left\langle\hat{\xi}_2^{gal}(d_1)\right\rangle\left\langle\hat{\xi}_2^{gal}(d_2)\right\rangle, \\ \mathrm{COV}_{2,2}^{\rm HI,\ HI}(d_1,d_2) &= \left\langle\hat{\xi}_2^{HI}(d_1)\hat{\xi}_2^{HI}(d_2)\right\rangle - \left\langle\hat{\xi}_2^{HI}(d_1)\right\rangle\left\langle\hat{\xi}_2^{HI}(d_2)\right\rangle, \\ \mathrm{COV}_{2,2}^{\rm gal,\ HI}(d_1,d_2) &= \left\langle\hat{\xi}_2^{gal}(d_1)\hat{\xi}_2^{HI}(d_2)\right\rangle - \left\langle\hat{\xi}_2^{gal}(d_1)\right\rangle\left\langle\hat{\xi}_2^{HI}(d_2)\right\rangle, \\ \mathrm{COV}_{1,2}^{\rm Dip,\ gal}(d_1,d_2) &= \left\langle\hat{\xi}_1(d_1)\hat{\xi}_2^{gal}(d_2)\right\rangle - \left\langle\hat{\xi}_1(d_1)\right\rangle\left\langle\hat{\xi}_2^{gal}(d_2)\right\rangle, \\ \mathrm{COV}_{1,2}^{\rm Dip,\ HI}(d_1,d_2) &= \left\langle\hat{\xi}_1(d_1)\hat{\xi}_2^{HI}(d_2)\right\rangle - \left\langle\hat{\xi}_1(d_1)\right\rangle\left\langle\hat{\xi}_2^{HI}(d_2)\right\rangle. \end{aligned} \tag{B.13}$$

The first line in (B.12) is the autocorrelation of the wide-angle contribution to the dipole, and the second one is the cross-correlation of the dipole with the wide-angle correction. Each term of the wide-angle contribution to the covariance

can be computed similarly to the dipole contribution by applying the general formula in (Hall & Bonvin 2017).

The cosmic variance contribution to the autocorrelation can be written as

$$\mathrm{COV}^{\mathrm{WA,\,auto}}_{\mathrm{CVCV}} = \left(\frac{3}{10r}\right)^2 \frac{50 d_1 d_2}{V} \int \frac{k^2 dk}{2\pi^2} j_2(kd_1) j_2(kd_2) \sum_{L_1,L_2} G_{22}^{L1L2} [P_{L1}^{\mathrm{gal}} P_{L2}^{\mathrm{gal}} + P_{L1}^{\mathrm{HI}} P_{L2}^{\mathrm{HI}} - 2 P_{L1}^{\mathrm{gal}} P_{L2}^{\mathrm{HI}}]; \tag{B.14}$$

the mixed cosmic variance – noise contribution to the autocorrelation, including both shot-noise and interferometer noise, is given by

$$\mathrm{COV}^{\mathrm{WA,\,auto}}_{\mathrm{CVNoise}} = \left(\frac{3}{10r}\right)^2 \frac{50 d_1 d_2}{V} \int \frac{k^2 dk}{2\pi^2} j_2(kd_1) j_2(kd_2) \sum_{L} G_{22}^{0L} [1/n_{\mathrm{gal}} P_L^{\mathrm{gal}} + (N_{\mathrm{HI}} + P_{\mathrm{N}}) P_L^{\mathrm{HI}}]; \tag{B.15}$$

while the noise contribution to the covariance is

$$\mathrm{COV}^{\mathrm{WA,\,auto}}_{\mathrm{NoiseNoise}} = \left(\frac{3}{10r}\right)^2 \frac{10}{4\pi L_p V} \delta_{d_1,d_2} \left[(N_{\mathrm{HI}} + P_{\mathrm{N}})^2 + \frac{1}{n_{\mathrm{gal}}^2}\right]. \tag{B.16}$$

The coefficients G_{22}^{L1L2} in (B.14) and (B.15) are defined in terms of the Wigner $3j$ symbols

$$G_{22}^{L1L2} = \sum_L (2L+1) \begin{pmatrix} 2 & 2 & L \\ 0 & 0 & 0 \end{pmatrix}^2 \begin{pmatrix} L_1 & L_2 & L \\ 0 & 0 & 0 \end{pmatrix}^2. \tag{B.17}$$

The cosmic variance contribution to the cross-correlation term in (B.12) can be written as

$$\mathrm{COV}^{\mathrm{WA,\,cross}}_{\mathrm{CVCV}} = \left(\frac{3}{10} \frac{d_1 + d_2}{r}\right) \frac{30}{V} \int \frac{k^2 dk}{2\pi^2} j_1(kd_1) j_2(kd_2) \sum_{L_1,L_2} G_{12}^{L1L2} P_{L1} [P_{L2}^{\mathrm{gal}} - P_{L2}^{\mathrm{HI}}]; \tag{B.18}$$

the cosmic variance x noise contribution to the cross-correlation is given by

$$\mathrm{COV}^{\mathrm{WA,\,cross}}_{\mathrm{CVNoise}} = \left(\frac{3}{10} \frac{d_1 + d_2}{r}\right) \frac{30}{V} \int \frac{k^2 dk}{2\pi^2} j_1(kd_1) j_2(kd_2) \sum_{L} \left(\frac{1}{n_{\mathrm{gal}}} - N_{\mathrm{HI}} - P_{\mathrm{N}}\right) G_{12}^{0L} P_L. \tag{B.19}$$

The coefficients G_{12}^{L1L2} are defined similarly to the coefficients involved in the autocorrelation term,

$$G_{12}^{L1L2} = \sum_L (2L+1) \begin{pmatrix} 1 & 2 & L \\ 0 & 0 & 0 \end{pmatrix}^2 \begin{pmatrix} L_1 & L_2 & L \\ 0 & 0 & 0 \end{pmatrix}^2. \tag{B.20}$$

The purely noise contribution to the cross-correlation term vanishes.

Appendix C

THE GENERATION OF VORTICITY IN COSMOLOGICAL LARGE SCALE STRUCTURE

C.1 Finite-box effect

In this section we test the effect of the finite box employed for our simulations on our power spectra.

We run two simulations with the same spatial resolution and different box size. Their parameters are summarized in table C.1. Both SIM1 and BIGSIM1 adopt the *rescaled* method to extract the velocity field. The physical box size L of BIGSIM1 is twice larger than the one of SIM1.

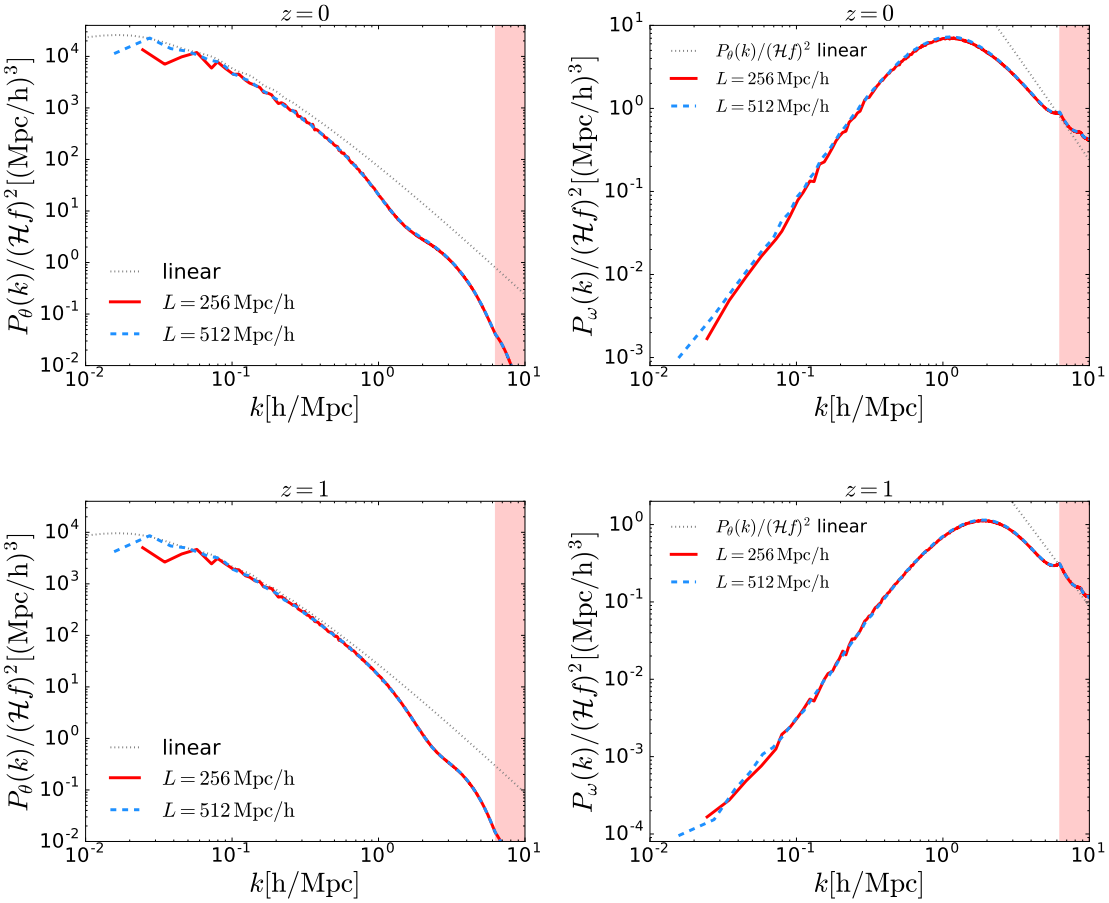
	$N_{\rm grid}$	$N_{\rm part}$	L [Mpc/h]	Grid-Cell size [Mpc/h]
SIM1	512	1024	256	0.5
BIGSIM1	1024	2048	512	0.5

Table C.1: Summary of the simulations used for testing the effect of a finite box.

In figure C.1 we compare the power spectra for the two simulations. Even if the BIGSIM1 is able to test a largest range of scales with respect to SIM1, we see that the power spectra of the two simulations agree in the common range of scales. Therefore, we conclude that a comoving box size of $L = 256$ Mpc/h is sufficiently large for the estimation of the divergence and vorticity power spectra on subhorizon scales.

C.2 Comparison with the Delaunay Tessellation Field Estimator method

In this section we compare our method for computing the vorticity power spectrum with a standard method used in the literature to extract the velocity field from



(a) Power spectrum - θ **(b)** Power spectrum - ω

Figure C.1: Power spectrum of the divergence θ (left panel) and the vorticity ω (right panel) of the velocity field at $z = 0$ (top panels) and $z = 1$ (bottom panels) for the simulations SIM1 and BIGSIM1 in table C.1. The red shaded region represents the modes beyond the Nyquist frequency.

N-body simulations, the Delaunay Tessellation Field Estimator (DTFE) method (Schaap & van de Weygaert 2000; van de Weygaert & Schaap 2007).

	N_{grid}	N_{part}	L [Mpc/h]	Grid-Cell size [Mpc/h]
SIM1	256	512	256	1
SIM2	512	1024	256	0.5

Table C.2: Summary of the simulations used in section C.2.

The parameters of the simulations used for the comparison are given in table C.2. The vorticity power spectrum is computed in *gevolution* with the *rescaled* method described in the previous sections. This is our reference method that we compare to the DTFE method. A snapshot of the simulation is then post-processed with the DTFE public code (Cautun & van de Weygaert 2011) in order to obtain

the vorticity field from the Delaunay tessellation. The DTFE method employs a linear interpolation scheme in order to estimate the requested fields in each point of a regular grid, once the values of the field is known in the vertices of the Delaunay tetrahedron in which the point is located. In the case of the vorticity field, the code computes the gradient of the velocity field, from which both vorticity and velocity divergence can be easily extracted. The DTFE code (Cautun & van de Weygaert 2011) employs a refined method, i.e. the fields are interpolated in $N_{\rm sample}$ points inside the tetrahedron and their values in the grid point is computed as the volume-weighted average of the field inside the tetrahedron. For our simulations we used $N_{\rm sample} = 100$, which is large enough for the vorticity field to reach convergence.

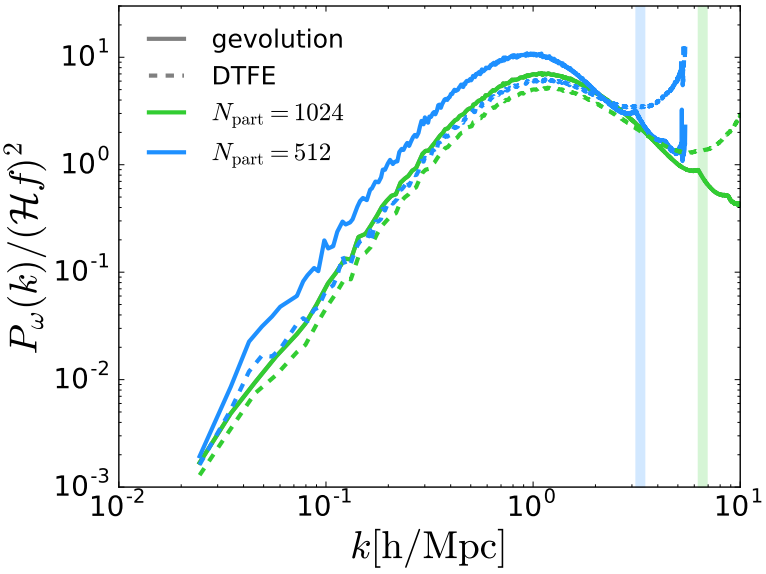


Figure C.2: Comparison of the vorticity power spectrum estimated with our *rescaled* method implemented in *gevolution* (solid lines) and the Delaunay Tessellation Field Estimator (DTFE) method (dashed lines). We show the comparison for the simulations SIM1 and SIM2 in table C.2 in blue and green, respectively. The power spectra are computed at $z = 0$. The shaded regions mark the Nyquist frequency of the two simulations.

In figure C.2 we compare the two methods at $z = 0$, for the two simulations SIM1 and SIM2 in table C.2. We see that for the low resolution simulations the two methods reproduce the same slope of the power spectrum, but a slightly different amplitude. We also notice that the difference in amplitude is smaller for the simulations with better resolution. This suggests that our method, based on the CIC projection of the momentum and density fields, introduces some extra projection effect with respect to the tessellation method and therefore it has a slower convergence. However, this spurious effect decays when improving the mass

and spatial resolution of the simulations. This is in agreement with the analysis presented in the appendix of (Pueblas & Scoccimarro 2009). For the SIM2 in table C.2 the difference is smaller than a factor two, therefore we expect the difference to be negligible for the high-resolution simulations discussed in section 7.3.3.

Bibliography

A. Alpher, R. & C. Herman, R. 1950, Theory of the Origin and Relative Abundance Distribution of the Elements, Vol. 22, 153–212

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, Phys. Rev. Lett., 119, 161101

Abbott, B. P. et al. 2016a, Phys. Rev. Lett., 116, 061102

Abbott, B. P. et al. 2016b, Phys. Rev. Lett., 116, 221101

Abbott, B. P. et al. 2017b, Nature, 551, 85

Abbott, B. P. et al. 2017c, Phys. Rev. Lett., 118, 221101

Abbott, B. P. et al. 2017d, Phys. Rev. Lett., 119, 161101

Abdalla, F. B., Bull, P., Camera, S., et al. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 17

Abel, T., Hahn, O., & Kaehler, R. 2012, Mon. Not. Roy. Astron. Soc., 427, 61

Abramo, L. R. & Bertacca, D. 2017, Phys. Rev. D[arXiv:1706.01834]

Adamek, J., Daverio, D., Durrer, R., & Kunz, M. 2016a, Nature Phys., 12, 346

Adamek, J., Daverio, D., Durrer, R., & Kunz, M. 2016b, JCAP, 1607, 053

Adamek, J., Durrer, R., & Kunz, M. 2014, Class. Quant. Grav., 31, 234006

Adamek, J., Durrer, R., & Kunz, M. 2017, JCAP, 1711, 004

Addison, G. E., Huang, Y., Watts, D. J., et al. 2016, Astrophys. J., 818, 132

Ade, P. A. R. et al. 2016a, Astron. Astrophys., 594, A13

Ade, P. A. R. et al. 2016b, Astron. Astrophys., 594, A14

Aghanim, N. et al. 2018a, ArXiv e-prints [arXiv:1807.06209]

Aghanim, N. et al. 2018b, ArXiv e-prints [arXiv:1807.06210]

Akrami, Y. et al. 2018a, ArXiv e-prints [arXiv:1807.06205]

Akrami, Y. et al. 2018b, ArXiv e-prints [arXiv:1807.06211]

Alam, S., Zhu, H., Croft, R. A. C., et al. 2017a, Mon. Not. Roy. Astron. Soc., 470, 2822

Alam, S. et al. 2017b, Mon. Not. Roy. Astron. Soc., 470, 2617

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

Alcubierre, M. 2008, Introduction to 3+1 Numerical Relativity (Oxford University Press, UK)

Alonso, D., Bellini, E., Ferreira, P. G., & Zumalacárregui, M. 2017, Phys. Rev., D95, 063502

Alonso, D., Bull, P., Ferreira, P. G., Maartens, R., & Santos, M. 2015, Astrophys. J., 814, 145

Alonso, D., Bull, P., Ferreira, P. G., & Santos, M. G. 2015, MNRAS, 447, 400

Alonso, D. & Ferreira, P. G. 2015, Phys. Rev., D92, 063525

Amati, L. & Della Valle, M. 2013, Int. J. Mod. Phys., D22, 1330028

Amendola, L. et al. 2013, Living Rev. Rel., 16, 6

Asorey, J., Crocce, M., Gaztanaga, E., & Lewis, A. 2012, Mon. Not. Roy. Astron. Soc., 427, 1891

Aubourg, É. et al. 2015, Phys. Rev., D92, 123516

Audren, B., Lesgourgues, J., Bird, S., Haehnelt, M. G., & Viel, M. 2013, JCAP, 1301, 026

Baker, T. & Bull, P. 2015, Astrophys. J., 811, 116

Bandura, K. et al. 2014, Proc. SPIE Int. Soc. Opt. Eng., 9145, 22

Barausse, E., Matarrese, S., & Riotto, A. 2005, Phys. Rev., D71, 063537

Bardeen, J. M. 1980, Phys. Rev., D22, 1882

Bartelmann, M. 1995, Astron. Astrophys., 298, 661

Bartolo, N., Komatsu, E., Matarrese, S., & Riotto, A. 2004, Phys. Rept., 402, 103

Baumann, D., Nicolis, A., Senatore, L., & Zaldarriaga, M. 2012, JCAP, 1207, 051

Baumgarte, T. W. & Shapiro, S. L. 2010, Numerical Relativity: Solving Einstein's Equations on the Computer (Cambridge University Press)

Bautista, J. E. et al. 2017, Astron. Astrophys., 603, A12

Ben-Dayan, I., Gasperini, M., Marozzi, G., Nugier, F., & Veneziano, G. 2013a, JCAP, 1306, 002

Ben-Dayan, I., Gasperini, M., Marozzi, G., Nugier, F., & Veneziano, G. 2013b, Phys. Rev. Lett., 110, 021301

Ben-Dayan, I., Marozzi, G., Nugier, F., & Veneziano, G. 2012, JCAP, 1211, 045

Bentivegna, E. & Bruni, M. 2016, Phys. Rev. Lett., 116, 251302

Berger, M. J. & Colella, P. 1989, Journal of Computational Physics, 82, 64

Bernal, J. L., Verde, L., & Riess, A. G. 2016, JCAP, 1610, 019

Bernardeau, F., Colombi, S., Gaztanaga, E., & Scoccimarro, R. 2002, Phys. Rept., 367, 1

Bernardeau, F. & van de Weygaert, R. 1996, Mon. Not. Roy. Astron. Soc., 279, 693

Bertacca, D. 2015, Class. Quant. Grav., 32, 195011

Bertacca, D., Maartens, R., & Clarkson, C. 2014, JCAP, 1409, 037

Bertacca, D., Maartens, R., Raccanelli, A., & Clarkson, C. 2012, JCAP, 1210, 025

Bertacca, D., Raccanelli, A., Bartolo, N., et al. 2017, Phys. Rev. D[arXiv:1705.09306]

Bertschinger, E., Dekel, A., Faber, S. M., Dressler, A., & Burstein, D. 1990, Astrophysical Journal, 364, 370

Beutler, F. et al. 2017, Mon. Not. Roy. Astron. Soc., 464, 3409

Bharadwaj, S. 1999, Astrophys. J., 516, 507

Bilicki, M. et al. 2016, ApJS[arXiv:1607.01182], [Astrophys. J. Suppl.225,5(2016)]

Blas, D., Lesgourgues, J., & Tram, T. 2011, JCAP, 1107, 034

Bloomfield, J. K., Flanagan, E. E., Park, M., & Watson, S. 2013, JCAP, 1308, 010

Bona, C. & Palenzuela-Luque, C. 2005, Elements of Numerical Relativity: From Einstein's Equations to Black Hole Simulations, Lecture Notes in Physics (Berlin: Springer)

Bonvin, C. 2008, Phys. Rev., D78, 123530

Bonvin, C. 2014, Class. Quant. Grav., 31, 234002

Bonvin, C., Andrianomena, S., Bacon, D., et al. 2017, Mon. Not. Roy. Astron. Soc., 472, 3936

Bonvin, C., Clarkson, C., Durrer, R., Maartens, R., & Umeh, O. 2015, JCAP, 1506, 050

Bonvin, C. & Durrer, R. 2011, Phys. Rev., D84, 063505

Bonvin, C., Durrer, R., & Gasparini, M. A. 2006, Phys. Rev., D73, 023523, [Erratum: Phys. Rev.D85,029901(2012)]

Bonvin, C., Durrer, R., Khosravi, N., Kunz, M., & Sawicki, I. 2018, JCAP, 1802, 028

Bonvin, C. & Fleury, P. 2018, JCAP, 1805, 061

Bonvin, C., Hui, L., & Gaztanaga, E. 2014, Phys. Rev., D89, 083535

Bonvin, C., Hui, L., & Gaztanaga, E. 2016, JCAP, 1608, 021

Borzyszkowski, M., Bertacca, D., & Porciani, C. 2017, Mon. Not. Roy. Astron. Soc., 471, 3899

Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., & Lemson, G. 2009, Mon. Not. Roy. Astron. Soc., 398, 1150

Breton, M.-A., Rasera, Y., Taruya, A., Lacombe, O., & Saga, S. 2018, ArXiv e-prints [arXiv:1803.04294]

Bull, P., Ferreira, P. G., Patel, P., & Santos, M. G. 2015, Astrophys. J., 803, 21

Bull, P. et al. 2016, Phys. Dark Univ., 12, 56

Calabrese, E., Slosar, A., Melchiorri, A., Smoot, G. F., & Zahn, O. 2008, Phys. Rev., D77, 123531

Caldwell, R. R., Dave, R., & Steinhardt, P. J. 1998, Phys. Rev. Lett., 80, 1582

Camera, S., Harrison, I., Bonaldi, A., & Brown, M. L. 2017, Mon. Not. Roy. Astron. Soc., 464, 4747

Camera, S., Maartens, R., & Santos, M. G. 2015a, Mon. Not. Roy. Astron. Soc., 451, L80

Camera, S., Santos, M. G., & Maartens, R. 2015b, Mon. Not. Roy. Astron. Soc., 448, 1035

Cardona, W., Durrer, R., Kunz, M., & Montanari, F. 2016, Phys. Rev., D94, 043007

Carrasco, J. J. M., Foreman, S., Green, D., & Senatore, L. 2014, JCAP, 1407, 057

Carrasco, J. J. M., Hertzberg, M. P., & Senatore, L. 2012, JHEP, 09, 082

Carroll, S. M., Press, W. H., & Turner, E. L. 1992, Annual Review of Astronomy and Astrophysics, 30, 499

Carucci, I. P., Villaescusa-Navarro, F., & Viel, M. 2017, JCAP, 4, 001

Castorina, E. & Villaescusa-Navarro, F. 2017, Mon. Not. Roy. Astron. Soc., 471, 1788

Cautun, M. C. & van de Weygaert, R. 2011, ArXiv e-prints [arXiv:1105.0370]

Challinor, A. & Lewis, A. 2011, Phys. Rev., D84, 043516

Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, Nature, 466, 463

Chiba, T., Okabe, T., & Yamaguchi, M. 2000, Phys. Rev., D62, 023511

Chuang, C.-H. et al. 2013, MNRAS[arXiv:1312.4889]

Clarkson, C. 2012, Comptes Rendus Physique, 13, 682

Clifton, T., Ferreira, P. G., Padilla, A., & Skordis, C. 2012, Phys. Rept., 513, 1

Clowe, D., Bradac, M., Gonzalez, A. H., et al. 2006, Astrophys. J., 648, L109

Cooray, A. & Sheth, R. K. 2002, Phys. Rept., 372, 1

Crocce, M., Cabre, A., & Gaztanaga, E. 2011, Mon. Not. Roy. Astron. Soc., 414, 329

Cusin, G., Tansella, V., & Durrer, R. 2017, Phys. Rev., D95, 063527

Dai, L., Kamionkowski, M., Kovetz, E. D., Raccanelli, A., & Shiraishi, M. 2016, Phys. Rev., D93, 023507, [Phys. Rev.D93,023507(2016)]

DESI. 2015, DESI Technical Design Report Part I: Science,Targeting, and Survey Design

Desjacques, V., Jeong, D., & Schmidt, F. 2018, Phys. Rept., 733, 1

Di Dio, E. 2017, JCAP, 1703, 016

Di Dio, E., Durrer, R., Marozzi, G., & Montanari, F. 2014a, JCAP, 1412, 017, [Erratum: JCAP1506,no.06,E01(2015)]

Di Dio, E., Durrer, R., Marozzi, G., & Montanari, F. 2016a, JCAP, 1601, 016

Di Dio, E., Montanari, F., Durrer, R., & Lesgourgues, J. 2014b, JCAP, 1401, 042

Di Dio, E., Montanari, F., Lesgourgues, J., & Durrer, R. 2013, JCAP, 1311, 044

Di Dio, E., Montanari, F., Raccanelli, A., et al. 2016b, JCAP, 1606, 013

Dodelson, S. 2003, Modern cosmology (San Diego, CA: Academic Press)

Dolag, K. & Bartelmann, M. 1997, Mon. Not. Roy. Astron. Soc., 291, 446

du Mas des Bourboux, H. et al. 2017, Astron. Astrophys., 608, A130

Durrer, R. 1994, Fund. Cosmic Phys., 15, 209

Durrer, R. 2008, The Cosmic Microwave Background (Cambridge University Press)

Durrer, R. & Caprini, C. 2003, JCAP, 0311, 010

Dyer, C. C. & Roeder, R. C. 1972, Astrophys. J., 174, L115

Dyer, C. C. & Roeder, R. C. 1973, Astrophys. J., 180, L31

Dyer, C. C. & Roeder, R. C. 1974, Astrophys. J., 189, 167

Efstathiou, G., Davis, M., White, S. D. M., & Frenk, C. S. 1985, The Astrophysical Journal Supplement Series, 57, 241

Eisenstein, D. J., Seo, H.-j., Sirko, E., & Spergel, D. 2007, Astrophys. J., 664, 675

Eisenstein, D. J. et al. 2005, Astrophys. J., 633, 560

Everitt, C. W. F. et al. 2011, Phys. Rev. Lett., 106, 221101

Everitt, C. W. F. et al. 2015, Class. Quant. Grav., 32, 224001

Evoli, C., Salucci, P., Lapi, A., & Danese, L. 2011, ApJ, 743, 45

Fanizza, G., Gasperini, M., Marozzi, G., & Veneziano, G. 2013, JCAP, 1311, 019

Fidler, C., Kleinjohann, A., Tram, T., Rampf, C., & Koyama, K. 2018, ArXiv e-prints [arXiv:1807.03701]

Fidler, C., Tram, T., Rampf, C., et al. 2016, JCAP, 1609, 031

Fidler, C., Tram, T., Rampf, C., et al. 2017, JCAP, 1712, 022

Fisher, R. A. 1935, Annals Eugen., 6, 391

Fixsen, D. J., Cheng, E. S., Gales, J. M., et al. 1996, Astrophys. J., 473, 576

Fonseca, J., Camera, S., Santos, M., & Maartens, R. 2015, Astrophys. J., 812, L22

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, Publ. Astron. Soc. Pac., 125, 306

Frenk, C. S., Colberg, J. M., Couchman, H. M. P., et al. 2000, ArXiv e-prints, astro

Friedmann, A. 1922, Zeitschrift fur Physik, 10, 377

Frieman, J., Turner, M., & Huterer, D. 2008, Ann. Rev. Astron. Astrophys., 46, 385

Frisch, U., Grimberg, G., & Villone, B. 2017, European Physical Journal H, 42 [arXiv:1707.01882]

Fry, J. N. & Gaztanaga, E. 1993, Astrophys. J., 413, 447

Gasperini, M., Marozzi, G., Nugier, F., & Veneziano, G. 2011, JCAP, 1107, 008

Gaztanaga, E., Bonvin, C., & Hui, L. 2017, JCAP, 1701, 032

Geach, J. E. et al. 2010, Mon. Not. Roy. Astron. Soc., 402, 1330

Giblin, J. T., Mertens, J. B., & Starkman, G. D. 2017, Class. Quant. Grav., 34, 214001

Gourgoulhon, E. 2007, ArXiv e-prints [arXiv:gr-qc/0703035]

GRAVITY Collaboration, Abuter, R., Amorim, A., et al. 2018, A&A, 615, L15

Grieb, J. N., Sánchez, A. G., Salazar-Albornoz, S., & Dalla Vecchia, C. 2016, Mon. Not. Roy. Astron. Soc., 457, 1577

Gubitosi, G., Piazza, F., & Vernizzi, F. 2013, JCAP, 1302, 032, [JCAP1302,032(2013)]

Guth, A. H. 1981, Phys. Rev. D, 23, 347

Guzzo, L. et al. 2008, Nature, 451, 541

Hahn, O., Angulo, R. E., & Abel, T. 2015, Mon. Not. Roy. Astron. Soc., 454, 3920

Hall, A. & Bonvin, C. 2017, Phys. Rev., D95, 043530

Hall, A., Bonvin, C., & Challinor, A. 2013, Phys. Rev., D87, 064026

Hamilton, A. J. S. 1997, in Ringberg Workshop on Large Scale Structure Ringberg, Germany, September 23-28, 1996

Heavens, A. F., Kitching, T. D., & Verde, L. 2007, Mon. Not. Roy. Astron. Soc., 380, 1029

Helmholtz, H. 1858, Journal für die reine und angewandte Mathematik, 55, 25

Heymans, C. et al. 2013, Mon. Not. Roy. Astron. Soc., 432, 2433

Hinshaw, G. et al. 2013, Astrophys. J. Suppl., 208, 19

Hockney, R. W. & Eastwood, J. W. 1988, Computer Simulation Using Particles (Bristol, PA, USA: Taylor & Francis, Inc.)

Hogg, D. W. 1999, ArXiv Astrophysics e-prints [astro-ph/9905116]

Horndeski, G. W. 1974, Int. J. Theor. Phys., 10, 363

Hubble, E. 1929, Proc. Nat. Acad. Sci., 15, 168

Iršič, V., Di Dio, E., & Viel, M. 2016, JCAP, 1602, 051

Jeong, D., Schmidt, F., & Hirata, C. M. 2012, Phys. Rev., D85, 023504

Jolicoeur, S., Umeh, O., Maartens, R., & Clarkson, C. 2017, JCAP, 1709, 040

Kaiser, N. 1987, Monthly Notices of the Royal Astronomical Society, 227

Kaiser, N. 2013, Mon. Not. Roy. Astron. Soc., 435, 1278

Kaiser, N. & Peacock, J. A. 2016, Mon. Not. Roy. Astron. Soc., 455, 4518

Kantowski, R. 2003, Phys. Rev., D68, 123516

Khoury, J. & Weltman, A. 2004, Phys. Rev., D69, 044026

Kim, H.-S., Wyithe, J. S. B., Baugh, C. M., et al. 2017, Monthly Notices of the Royal Astronomical Society, 465, 111

Kitching, T. D., Amara, A., Abdalla, F. B., Joachimi, B., & Refregier, A. 2009, Mon. Not. Roy. Astron. Soc., 399, 2107

Knox, L., Scoccimarro, R., & Dodelson, S. 1998, Phys. Rev. Lett., 81, 2004

Kodama, H. & Sasaki, M. 1984, Prog. Theor. Phys. Suppl., 78, 1

Kopp, M., Vattis, K., & Skordis, C. 2017, Phys. Rev., D96, 123532

Kovac, J. M., Leitch, E. M., Pryke, C., et al. 2002, Nature, 420, 772

Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. 2004, Astrophys. J., 609, 35

Kunz, M. 2012, Comptes Rendus Physique, 13, 539

Lahav, O., Lilje, P. B., Primack, J. R., & Rees, M. J. 1991, Mon. Not. Roy. Astron. Soc., 251, 128

Lemaître, G. 1931, MNRAS, 91, 490

Lepori, F., Di Dio, E., Viel, M., Baccigalupi, C., & Durrer, R. 2017, JCAP, 1702, 020

Lesgourgues, J. 2011, ArXiv e-prints [arXiv:1104.2932]

Levi, M. et al. 2013, ArXiv e-prints [arXiv:1308.0847]

Lidz, A., Zahn, O., Furlanetto, S. R., et al. 2009, ApJ, 690, 252

Lombriser, L., Yoo, J., & Koyama, K. 2013, Phys. Rev., D87, 104019

Lorenz, C. S., Alonso, D., & Ferreira, P. G. 2017, Phys. Rev. D[arXiv:1710.02477]

Lu, T. H.-C., Ananda, K., Clarkson, C., & Maartens, R. 2009, JCAP, 0902, 023

Luciano, R. & Zanotti, O. 2013, Relativistic hydrodynamics (Oxford University Press)

Ma, C.-P. & Bertschinger, E. 1995, Astrophys. J., 455, 7

Macpherson, H., Lasky, P. D., & Price, D. J. 2018, ArXiv e-prints [arXiv:1807.01714]

Mancuso, C., Lapi, A., Prandoni, I., et al. 2017, ApJ, 842, 95

Mandel, K. S. & Zaldarriaga, M. 2006, Astrophys. J., 647, 719

Marulli, F., Carbone, C., Viel, M., Moscardini, L., & Cimatti, A. 2011, Mon. Not. Roy. Astron. Soc., 418, 346

Masui, K. W. et al. 2013, Astrophys. J., 763, L20

Matsubara, T. 2000, Astrophys. J., 535, 1

Matsubara, T. 2004, Astrophys. J., 615, 573

McDonald, P. 2009, JCAP, 0911, 026

McDonald, P. & Vlah, Z. 2018, Phys. Rev., D97, 023508

Mehta, K. T., Seo, H.-J., Eckel, J., et al. 2011, Astrophys. J., 734, 94

Mocz, P., Lancaster, L., Fialkov, A., Becerra, F., & Chavanis, P.-H. 2018, Phys. Rev., D97, 083519

Monaghan, J. J. 1992, Annual Review of Astronomy and Astrophysics, 30, 543

Montanari, F. & Durrer, R. 2012, Phys. Rev., D86, 063503

Montanari, F. & Durrer, R. 2015, JCAP, 1510, 070

Moré, J. J. 1978, Numerical Analysis: Proceedings of the Biennial Conference Held at Dundee, June 28–July 1, 1977 (Berlin, Heidelberg: Springer Berlin Heidelberg), 105–116

Motloch, P. & Hu, W. 2018, Phys. Rev., D97, 103536

Mukhanov, V. F., Feldman, H. A., & Brandenberger, R. H. 1992, Phys. Rept., 215, 203

Namikawa, T., Okamura, T., & Taruya, A. 2011, Phys. Rev., D83, 123514

Nielsen, J. T. & Durrer, R. 2017, JCAP, 1703, 010

Norris, R. P. et al. 2011, Publ. Astron. Soc. Austral., 28, 215

Padmanabhan, H. & Refregier, A. 2017, Mon. Not. Roy. Astron. Soc., 464, 4008

Padmanabhan, H., Refregier, A., & Amara, A. 2017, Mon. Not. Roy. Astron. Soc., 469, 2323

Papai, P. & Szapudi, I. 2008, Mon. Not. Roy. Astron. Soc., 389, 292

Peacock, J. A. & Smith, R. E. 2000, Mon. Not. Roy. Astron. Soc., 318, 1144

Penzias, A. A. & Wilson, R. W. 1965, Astrophys. J., 142, 419

Perlmutter, S. et al. 1998, Nature, 391, 51

Peterson, J. B. et al. 2009, ArXiv e-prints [arXiv:0902.3091]

Piazza, F. & Vernizzi, F. 2013, Class. Quant. Grav., 30, 214007

Pietroni, M. 2018, J. Cosmology Astropart. Phys.[arXiv:1804.09140]

Pueblas, S. & Scoccimarro, R. 2009, Phys. Rev., D80, 043504

Raccanelli, A., Bertacca, D., Doré, O., & Maartens, R. 2014, JCAP, 1408, 022

Raccanelli, A., Bertacca, D., Jeong, D., Neyrinck, M. C., & Szalay, A. S. 2016a, Physics of the Dark Universe [arXiv:1602.03186]

Raccanelli, A., Bertacca, D., Maartens, R., Clarkson, C., & Doré, O. 2016b, Gen. Rel. Grav., 48, 84

Raccanelli, A., Bertacca, D., Pietrobon, D., et al. 2013, Mon. Not. Roy. Astron. Soc., 436, 89

Raccanelli, A., Montanari, F., Bertacca, D., Doré, O., & Durrer, R. 2016c, JCAP, 1605, 009

Raccanelli, A., Samushia, L., & Percival, W. J. 2010, Mon. Not. Roy. Astron. Soc., 409, 1525

Rampf, C. & Frisch, U. 2017, Mon. Not. Roy. Astron. Soc., 471, 671

Renk, J., Zumalacarregui, M., & Montanari, F. 2016, JCAP, 1607, 040

Riess, A. G. et al. 1998, Astron. J., 116, 1009

Riess, A. G. et al. 2016, Astrophys. J., 826, 56

Risaliti, G. & Lusso, E. 2015, Astrophys. J., 815, 33

Robertson, H. P. 1935, Astrophys. J., 82, 284

Ross, A. J. et al. 2012, Mon. Not. Roy. Astron. Soc., 424, 564

Ross, A. J. et al. 2017, Mon. Not. Roy. Astron. Soc., 464, 1168

Rubin, V. C., Ford, Jr., W. K., & Thonnard, N. 1980, ApJ, 238, 471

Sadeh, I., Feng, L. L., & Lahav, O. 2015, Phys. Rev. Lett., 114, 071103

Saga, S., Taruya, A., & Colombi, S. 2018, ArXiv e-prints [arXiv:1805.08787]

Samushia, L., Percival, W. J., & Raccanelli, A. 2012, Mon. Not. Roy. Astron. Soc., 420, 2102

Samushia, L. et al. 2014, Mon. Not. Roy. Astron. Soc., 439, 3504

Sanchez, A. G. et al. 2014, Mon. Not. Roy. Astron. Soc., 440, 2692

Sanchez, E., Alonso, D., Sanchez, F. J., Garcia-Bellido, J., & Sevilla, I. 2013, Mon. Not. Roy. Astron. Soc., 434, 2008

Sanchez, E., Carnero, A., Garcia-Bellido, J., et al. 2011, Mon. Not. Roy. Astron. Soc., 411, 277

Santos, M. G., Alonso, D., Bull, P., Silva, M., & Yahya, S. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14) [arXiv:1501.03990]

Sanz, J. L., Martinez-Gonzalez, E., & Benitez, N. 1997, Mon. Not. Roy. Astron. Soc., 291, 418

Sasaki, M. 1987, Mon. Not. Roy. Astron. Soc., 228, 653

Schaap, W. E. & van de Weygaert, R. 2000, Astron. Astrophys., 363, L29

Schmidt, F. & Jeong, D. 2012, Phys. Rev., D86, 083527

Scoccimarro, R. 2001, Annals N. Y. Acad. Sci., 927, 13

Scoccimarro, R., Sheth, R. K., Hui, L., & Jain, B. 2001, Astrophys. J., 546, 20

Scranton, R. et al. 2005, Astrophys. J., 633, 589

Seljak, U. 2000, Mon. Not. Roy. Astron. Soc., 318, 203

Shandarin, S., Habib, S., & Heitmann, K. 2012, Phys. Rev., D85, 083005

Sousbie, T. & Colombi, S. 2016, J. Comput. Phys., 321, 644

Springel, V. 2005, Mon. Not. Roy. Astron. Soc., 364, 1105

Springel, V. 2010, Mon. Not. Roy. Astron. Soc., 401, 791

Springel, V., Frenk, C. S., & White, S. D. M. 2006, Nature, 440, 1137

Springel, V. et al. 2005, Nature, 435, 629

Springel, V. et al. 2018, Mon. Not. Roy. Astron. Soc., 475, 676

Szalay, A. S., Matsubara, T., & Landy, S. D. 1998, Astrophys. J., 498, L1

Szapudi, I. 2004, Astrophys. J., 614, 51

Tansella, V., Bonvin, C., Durrer, R., Ghosh, B., & Sellentin, E. 2017, JCAP [arXiv:1708.00492]

Taylor, A. N., Kitching, T. D., Bacon, D. J., & Heavens, A. F. 2007, Mon. Not. Roy. Astron. Soc., 374, 1377

Taylor, A. R. & Jagannathan, P. 2016, Mon. Not. Roy. Astron. Soc., 459, L36

Taylor, J. H. & Weisberg, J. M. 1989, Astrophys. J., 345, 434

Tegmark, M., Hamilton, A. J. S., Strauss, M. A., Vogeley, M. S., & Szalay, A. S. 1998, Astrophys. J., 499, 555

Thomas, D. B., Bruni, M., & Wands, D. 2015, Mon. Not. Roy. Astron. Soc., 452, 1727

Tinker, J. L., Kravtsov, A. V., Klypin, A., et al. 2008, Astrophys. J., 688, 709

Tinker, J. L., Kravtsov, A. V., Klypin, A., et al. 2010, The Astrophysical Journal, 724, 878

Troxel, M. A. & Ishak, M. 2014, Phys. Rept., 558, 1

Umeh, O., Clarkson, C., & Maartens, R. 2014a, Class. Quant. Grav., 31, 202001

Umeh, O., Clarkson, C., & Maartens, R. 2014b, Class. Quant. Grav., 31, 205001

Umeh, O., Jolicoeur, S., Maartens, R., & Clarkson, C. 2017, JCAP, 1703, 034

van de Weygaert, R. & Schaap, W. 2007, ArXiv e-prints [arXiv:0708.1441], [Lect. Notes Phys.665,291(2009)]

Vargas-Magana, M. et al. 2016, ArXiv e-prints [arXiv:1610.03506]

Villa, E., Di Dio, E., & Lepori, F. 2018, JCAP, 1804, 033

Villaescusa-Navarro, F., Alonso, D., & Viel, M. 2017, Mon. Not. Roy. Astron. Soc., 466, 2736

Villaescusa-Navarro, F., Bull, P., & Viel, M. 2015, Astrophys. J., 814, 146

Villaescusa-Navarro, F., Planelles, S., Borgani, S., et al. 2016, MNRAS, 456, 3553

Villaescusa-Navarro, F., Viel, M., Alonso, D., et al. 2015, JCAP, 3, 034

Villaescusa-Navarro, F., Viel, M., Datta, K. K., & Choudhury, T. R. 2014, JCAP, 1409, 050

Villaescusa-Navarro, F. et al. 2016, Mon. Not. Roy. Astron. Soc., 456, 3553

Villone, B. & Rampf, C. 2017, European Physical Journal H, 42 [arXiv:1707.01883]

Vogelsberger, M., Genel, S., Springel, V., et al. 2014, Mon. Not. Roy. Astron. Soc., 444, 1518

Walker, A. G. 1937, Proceedings of the London Mathematical Society, s2-42, 90

Weinberg, S. 1972, Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity (New York, NY: Wiley)

Weinberg, S. 1989, Rev. Mod. Phys., 61, 1

White, M. J., Hernquist, L., & Springel, V. 2001, Astrophys. J., 550, L129

Widrow, L. M. & Kaiser, N. 1993, Astrophys. J., 416, L71

Will, C. M. 2014, Living Rev. Rel., 17, 4

Wojtak, R., Hansen, S. H., & Hjorth, J. 2011, Nature, 477, 567

Yang, X., Mo, H. J., Bosch, F. C. v. d., Zhang, Y., & Han, J. 2012, Astrophys. J., 752, 41

Yoo, J. 2010, Phys.Rev.D, 82, 083508

Yoo, J., Fitzpatrick, A. L., & Zaldarriaga, M. 2009, Phys.Rev., D80, 083514

Yoo, J., Hamaus, N., Seljak, U., & Zaldarriaga, M. 2012, Phys. Rev., D86, 063514

Yoo, J. & Seljak, U. 2015, Mon. Not. Roy. Astron. Soc., 447, 1789

Yoo, J. & Zaldarriaga, M. 2014, Phys. Rev., D90, 023513

Zehavi, I. et al. 2011, Astrophys. J., 736, 59

Zhang, X., Huang, Q.-G., & Li, X.-D. 2018, ArXiv e-prints [arXiv:1801.07403]

Zhao, G.-B., Pogosian, L., Silvestri, A., & Zylberberg, J. 2009, Phys. Rev., D79, 083513

Zhao, H., Peacock, J. A., & Li, B. 2013, Phys. Rev. D, 88, 043013

Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, Astrophys. J., 633, 791

Zheng, Z., Coil, A. L., & Zehavi, I. 2007, Astrophys. J., 667, 760

Zhu, H., Alam, S., Croft, R. A. C., Ho, S., & Giusarma, E. 2017a, Mon. Not. Roy. Astron. Soc., 471, 2345

Zhu, H.-M., Yu, Y., & Pen, U.-L. 2017b, Phys. Rev. D[arXiv:1711.03218]

Zhu, W. & Feng, L.-L. 2017, Astrophys. J., 838, 21

Zhu, W., long Feng, L., & Fang, L.-Z. 2010, The Astrophysical Journal, 712, 1

Zwicky, F. 1937, Astrophysical Journal, 86, 217