Università
degli Studi
di Padova

Dipartimento
di Fisica
e Astronomia
Galileo Galilei

Corso Di Dottorato In Fisica

Relativistic redshift-space distortions in stage IV galaxy surveys

Mohmaed Yousry Elkhashab

tesi di dottorato

Supervisore:
Dr. Daniele Bertacca
Co-Supervisore:
Prof. Sabino Matarrese

Coordinatore:
Prof. Giulio Monaco

xxxvi ciclo

To Manal, the original Yusry, and Marioma.

Abstract

We map the Universe through galaxy redshift surveys. Statistically analysing these galaxy maps allows us access to the nature of dark matter and dark energy. To interpret these galaxy maps, however, the measured redshifts are converted into distances assuming a homogeneous and isotropic universe. That assumption ignores the rich structure of the Universe and thus contaminates the estimated distances through redshift space distortions (RSD). The leading order RSD is caused by the peculiar velocities of the sources which in turn are sourced by gravitational instability. These RSD change the observed clustering patterns of galaxies which allows us to probe the nature of gravity. However, the photons emitted by distant galaxies are also affected by the inhomogeneities they encounter on their way to the observer. These additional relativistic RSD include multiple effects that alter the radial and transversal positions as well as the observed fluxes of galaxies. While these effects are subdominant at small scales, they gain prominence at the larger scales that will be probed by Stage IV surveys such as *Euclid* and `SKA02`, which are poised to explore unprecedentedly vast regions of the late Universe.

This thesis investigates these relativistic RSD in-depth, addressing two central questions: Can they be detected in Stage IV galaxy surveys? and can they be used as novel probes to measure cosmological parameters? To these ends, we adopt the `LIGER` method which produces galactic mock catalogues tailored to match user-defined surveys after accounting for relativistic RSD at linear order in cosmological perturbations.

To answer the first inquiry, we analyze the impact of relativistic RSD on the angular power spectrum, two-point correlation function, and power spectrum multipoles measured from mock catalogues that match the radial and angular distribution of the *Euclid* Wide Spectroscopic Survey. We quantitatively assess the significance of relativistic RSD, concentrating specifically on wide-angle effects, weak lensing and velocity of the observer effects. We use the latter's impact on the power spectrum multipoles, dubbed the 'Finger of the Observer' (FOTO) effect, as our key in answering the second inquiry. Using mock *Euclid* and `SKA02` catalogues, we find that the FOTO effect can be used to measure both the magnitude and direction of the peculiar velocity of the observer as well as constrain cosmological parameters in optimistic conditions.

Acknowledgments

This document marks the conclusion of a four-year exciting – albeit sometimes exhausting – journey. I would like to dedicate a small part of it to thank my journey companions whose support and input have been necessary for its completion. First and foremost, I am thankful to my supervisor and mentor Daniele Bertacca. More than just a mentor, Daniele has been a trusted friend whose influence extended beyond the realm of physics education. His pedagogical approach and profound insights in physics coupled with personal care and genuine kindness have significantly shaped the trajectory of my research endeavours. They are something that I will strive for in my future work. Additionally, I extend my heartfelt appreciation to Sabino Matarrese for his unwavering support, invaluable insights into cosmology, and consistent guidance over the past three years. His expertise in the field has greatly enriched my understanding and contributed significantly to the development of my research. I am also grateful to Cristiano Porciani, whose attention to detail and illuminating physics discussions have had a profound influence on my approach as a researcher. I would like to extend my gratitude to the Padova cosmology group staff from whom I have learnt a great deal; Nicola Bartolo, Michele Liguori, Alvise Raccanelli and Angelo Ricciardone. I am also thankful to the fellow members of the Padova cosmology group Pritha Bari, Alessandro Greco, Lorenzo Valbusa dall'Armi, Federico Semenzato, Jessie De Kruijf, Andrea Begnoni, Gabriele Perna and Francesco Spezzati who have made my stay in Padova truly enjoyable and who I wish the best in future endeavours. I would also like to thank my friends in Padova, Marah Jamil, Prajwal Padmanabha and Ivana Batkovic.

Of course, I can not forget the friendships I have made during my period abroad. To my friends, Davit Alkhanishvili, Anna Pugno, Mandar Karandikar, Yvonne Fichtner, Prachi Khatri, Alex Eggemeier, Jacopo Salvalaggio and Emilio Romano-Díaz, I thank you all for the enjoyable moments and all the rant sessions.

I am also thankful to my friends who were patient enough to read drafts of my thesis, Heba Elkilany and Mohamed Samir. And to my friends in distant lands, Ahmed Elmenshawy, Abdelrahaman Gaber, Mahmoud Aboutaleb, Mohamed Galal, Mohamed Mamdouh and Jaber Zain, I appreciate that geographical limitations have only made our friendships stronger.

I would like to state that I am eternally grateful to my parents, Yusry and Manal.

Their love, patience and sacrifices are the only reason I have made it this far. I would also like to thank my brothers Ahmed and Abdellatif for always looking out for me and my little sister Nada for being the coolest sister anybody could ask for. Finally, to the person who not only tolerated but supported me during all the long nights and the stressful months and who always brings a part of Egypt everywhere we live, my wife Mariam, I am truly grateful.

Mohamed Yousry Elkhashab
April 1,2024

List of Publications

The research carried out through this work has lead to one peer-reviewed publication, an under review key project in the *Euclid* collaboration and a to be submitted publication:

- Paper I: **Elkhashab, M. Y.**, Porciani, C., & Bertacca, D. 2021, *The large-scale monopole of the power spectrum in a Euclid-like survey: wide-angle effects, lensing, and the 'finger of the observer'*, Monthly Notices of the Royal Astronomical Society, 509, 1626.

- Paper II: **Elkhashab, M. Y.**, Porciani, C., & Bertacca, D. , *Measuring the observer velocity from spectroscopic redshift surveys* **To be submitted to JCAP**.

- Paper III: **Elkhashab, M. Y.**, Bertacca, D., Porciani, C., Salvalaggio, J. et al., *Euclid preparation. The impact of relativistic redshift-space distortions on two-point clustering statistics from the Euclid Wide Spectroscopic Survey.* **Submitted to the ECEB.**

Contents

Chapter 1

INTRODUCTION

Galaxy redshift surveys measure the redshifts, fluxes, and angular positions of distant light-emitting sources. These measurements are converted into comoving distances, assuming a Friedmann–Lemaître–Robertson–Walker (FLRW) universe. Despite its convenience, this assumption neglects the rich structure of the Universe (Sargent & Turner, 1977), leading to distortions in the estimated distances due to various projection effects collectively known as Redshift Space Distortions (RSD). As shown by Kaiser (1987), the leading order RSD is caused by the peculiar motion of the light-emitting objects which in turn is sourced by gravitational instability.

The impact of the peculiar velocity effect can be directly incorporated into the galaxy clustering models of two-point statistics (the main object of study of this work) under the so-called Global Plane-Parallel (GPP) approximation. The GPP assumes that all galaxies in a particular survey are sufficiently distant that their lines of sight are effectively parallel. Kaiser (1987) showed that the impact of velocity shift RSD under the GPP approximation on the galaxy power spectrum can be cast into a simple amplitude boost that depends on the angle between the line of sight and the wavevector (see also Hamilton & Culhane, 1996, Hamilton, 1998, 2000, for extensions to configuration space). That effect has been incorporated in the modelling and analysis of subsequent relatively small redshift surveys to great success (e.g. Peacock et al., 2001, Wilson et al., 2017, Beutler et al., 2014).

However, the upcoming generation of surveys will measure galaxy clustering on unprecedentedly large scales. For example, The recently launched *Euclid* satellite (Laureijs et al., 2011) will observe tens of millions of galaxies within its spectroscopic sample (Euclid Collaboration: Scaramella et al., 2022) and more than a billion in its photometric counterpart (Euclid Collaboration: Tanidis et al., 2023) covering 15 000 $\deg^2$. The study of galaxy clustering in these large scales requires the consideration of two primary modelling challenges. Firstly, (*i*) the model needs to go beyond the GPP approximation to account for galaxy pairs with significant angular separation (Szalay et al., 1998). Secondly (*ii*), the model needs to account for other sources of RSD (Hereafter relativistic RSD Challinor & Lewis, 2011).

Focusing on the first challenge, corrections beyond the GPP approximation, collectively referred to as wide-angle effects, become significant when exploring the

clustering dynamics of galaxies at larger angular separations. For surveys with large angular separations, wide-angle effects have been shown to impact the configuration space two-point correlation function (2PCF) (see e.g. Matsubara, 2000a, Szapudi, 2004, Pápai & Szapudi, 2008, Raccanelli et al., 2010, Samushia et al., 2012, Raccanelli et al., 2018) and the power spectrum multipoles (see e.g. Zaroubi & Hoffman, 1996, Reimberg et al., 2016, Castorina & White, 2018, Castorina et al., 2019). A proposed treatment to wide-angle effects in the power spectrum is the so-called Local Plane Parallel (LPP) approximation which partially accounts for the variable lines of sight of galaxies in the estimator and the survey window function treatment (Beutler et al., 2014, Wilson et al., 2017).

The second complication in the examination of clustering statistics at large scales arises from the relativistic RSD that are not sourced by the peculiar velocities of the sources. As we only detect galaxies on our past lightcone, the null geodesics of the photons emitted by these galaxies are distorted by the inhomogeneities between the source galaxy and the Observer. To accurately account for these effects, a fully relativistic treatment of RSD effects has been developed up to first order in cosmological perturbations (Yoo et al., 2009, Bonvin & Durrer, 2011, Challinor & Lewis, 2011, Jeong et al., 2012) as well as up to second order (Bertacca et al., 2014a,b, Bertacca, 2015). Subsequently, the impact of these effects on the angular power spectrum was integrated into two prominent Boltzmann solvers, namely `CAMB` (Lewis & Bridle, 2002) and `CLASS` (Di Dio et al., 2013). The impact of these relativistic RSD on 3D clustering statistics has also been investigated in the context of the 2PCF (e.g. Bertacca, 2015, Raccanelli et al., 2016b, Tansella et al., 2018, Bertacca, 2020, Jelic-Cizmek et al., 2021). For the power spectrum, perturbative analytical treatments compared the relative importance of different relativistic RSD (Castorina & Di Dio, 2022, Foglieni et al., 2023, Noorikuhani & Scoccimarro, 2023). Another approach is based on numerical studies of mock galaxy lightcones (Borzyszkowski et al., 2017, Breton et al., 2022). In particular, for the power spectrum and Fourier-based statistics, the loss of translational invariance complicates the very definition of the statistics. A possible solution to this issue is to use mixed-space statistics (or the 'local' power spectrum Bianchi et al., 2015, Scoccimarro, 2015, Reimberg et al., 2016). Another possible prescription which solves that issue and naturally incorporates relativistic RSD is to forgo Cartesian Fourier statistics entirely and opt for the "Spherical-Fourier-Bessel " (SFB) formalism. It was first introduced by Peebles (1973), then extended to redshift space by Heavens & Taylor (1995), and applied to real data in Percival et al. (2004). Finally, it was extended to include GR effects in Yoo & Desjacques (2013), Bertacca et al. (2018). However, we do not adopt this prescription in this project as we aim to measure the impact of relativistic RSD in the standard statistics used in stage-IV surveys.

These relativistic RSD are the main object of study of this work. In particular, this thesis is designed to answer two main questions: *i)* Can relativistic RSD be detected in Stage IV galaxy surveys? *ii)* Can these effects be directly used to extract cosmological information with constraints competitive to standard probes? To answer both questions, we numerically study relativistic RSD through their impact

on the two-point statistics of realistic galaxy surveys using the LIGER method. The LIGER method (Borzyszkowski et al., 2017, Paper I, Paper II, Paper III) is a numerical technique that takes in N-body simulations and produces survey-tailored galaxy catalogues that include the impact of relativistic RSD at first order in cosmological perturbations.

Concentrating on the first question, we construct 140 galaxy catalogues specifically tailored to match the anticipated wide spectroscopic sample of the recently launched Euclid mission. However, we build different iterations of these galaxy catalogues, each incorporating progressively more RSD. Subsequently, we employ a likelihood ratio hypothesis test to quantitatively assess the significance of one or multiple relativistic RSD. Firstly, We study the importance of wide-angle effects. To measure the significance of wide-angle effects, we compare the power spectra of mock datasets containing solely the Kaiser velocity-gradient term against two power spectrum models: the GPP model and the LPP model. Notably for this analysis, we analyse simultaneously disconnected patches of the sky to reach the largest scales accessible to the *Euclid* survey as opposed to the standard method which treats disconnected patches of the sky independently.

Next, we delve into the significance of different relativistic RSD effects on the two-point statistics measured from the *Euclid* Wide Spectroscopic Survey (Euclid Collaboration: Scaramella et al., 2022). In particular, we concentrate on the weak gravitational lensing and the observer velocity sourced terms. We evaluate the importance of these RSD across three two-point statistics: the angular power spectrum, the 2PCF multipoles, and the power spectrum multipoles.

We then dedicate a significant portion of this thesis to the relativistic RSD sourced by the velocity of the observer, which result a kinematic dipole in the observed distribution of galaxies on the sky via aberration and Doppler effects (Ellis & Baldwin, 1984). If the large-scale distribution of galaxies is isotropic and the rest frames of matter and galaxies coincide with the Cosmic Microwave Background (CMB) rest frame, then the amplitude of that kinematic dipole should match the observed kinematic dipole of the CMB. Half a century of CMB measurements have shown that, under the kinematic interpretation, the heliocentric peculiar velocity $v_\odot$ is approximately 370 km s^{-1}, pointing towards $\ell \approx 263°$ and $b \approx 48°$ in Galactic coordinates (Fixsen et al., 1996, Hinshaw et al., 2009, Planck Collaboration: Aghanim et al., 2020b, Delouis et al., 2021). However, measurements of the kinematic dipole from radio surveys have given discrepant results (see Peebles, 2022, for a recent review). Recent studies consistently measure a dipole anisotropy that matches the CMB direction but produces a significantly larger amplitude (Tiwari et al., 2015, Siewert et al., 2021). Similar results have been obtained for quasars. In this case, the discrepancy with the CMB dipole reaches a high 5σ confidence level (Singal, 2021, Secrest et al., 2021, 2022). That amplitude discrepancy can be diminished by accounting for the redshift-dependent magnification and evolution bias in the projected signal (Gibelyou & Huterer, 2012, Dalang & Bonvin, 2022, Guandalin et al., 2023). In this thesis, we introduce a novel method to measure that kinematic dipole through the velocity of the observer contribution to the ob-

served galaxy contrast that does not depend on a spherical surface projection of the signal but uses the full 3D information of the galaxy density field.

To that end, we devote special attention to the oscillatory signal induced by the velocity of the observer in the observed power spectrum multipoles, making it the focal point of our investigation for the second aspect of this study. Referred to as the Finger Of The Observer (FOTO) signal, we analytically derive its impact on any multipole of the power spectrum. In order to validate our analytical findings numerically, we create two mock survey catalogues using the `LIGER` method: one resembling the `SKA02` (Bacon et al., 2020) survey and another resembling the *Euclid* survey. Furthermore, we demonstrate both analytically and numerically that the FOTO effect is actually amplified not cancelled by a radial redshift correction. Leveraging that finding, we introduce the aforementioned methodology aimed at extracting both the magnitude and direction of the observer's velocity from the FOTO effect. Finally, under favourable conditions, we illustrate that the FOTO effect holds promise for measuring cosmological parameters, in particular the matter and dark energy density parameters.

This thesis is structured as follows. In Chapter 2, we briefly outline the standard cosmological model and present the theoretical treatment of RSD both classically and accounting for relativistic projection effects. Following that, we describe how we create the mock galaxy catalogues using the new version of the `LIGER` method implemented in this work. In Chapter 4, we outline the estimators employed in measuring the two-point statistics used in this work as well as the statistical methods we use to measure the significance of relativistic RSD. Subsequently, we devote Chapter 5 to wide-angle effects and comparing the mock power spectra to the Kaiser model using both the GPP and LPP approximations. We then address the first main objective of this work in Chapter 6 where we measure the detectability of different relativistic RSD for the *Euclid* survey. In Chapter 7, we introduce analytically and numerically the FOTO effect and we then use it to measure the magnitude and direction of the velocity of the Observer as well as the matter density parameter. Finally, we summarise the thesis and conclude in Chapter 8.

Chapter 2

Theoretical Background

Cosmology is the scientific endeavour dedicated to the study of the origin, evolution, and ultimate fate of the Universe. At the core of contemporary cosmological understanding lies the standard ΛCDM model. Assuming General Relativity is the gravitational theory that describes our Universe and that the Universe is homogeneous and isotropic at large enough scales, the ΛCDM model consists of two fundamental components: cold dark matter (DM) and the cosmological constant (Λ), representing dark energy (DE) responsible for the Universe's accelerated expansion. The latest observations indicate that 68% of the cosmic makeup can be attributed to DE, 27% to DM, while all baryonic matter accounts for only 5% (Planck Collaboration: Aghanim et al., 2020a). Consequently, the nature of the vast majority of the makeup of the Universe still eludes us. A primary objective of contemporary and forthcoming cosmological investigations is to deepen our understanding of dark energy and dark matter. Given our access to just one test sample – the Universe we inhabit – cosmology differs from other scientific disciplines. In other words, we can not carry out repeated experiments/observations to form a statistical sample. It relies on statistical descriptions of hypothetical universes in comparison with our observations of this singular test sample.

In this chapter, we (briefly) introduce the standard model of cosmology and the statistical description of the Universe. We then devote two sections to the main topic of this thesis: Redshift Space Distortions (RSD). We dedicate a section to each of the following: the classical description of redshift distortions and the relativistic treatment of the phenomenon.

Conventions

Throughout this thesis, we adopt Einstein's summation convention and define the space-time metric tensor to have the signature $(-,+,+,+)$. Greek indices signify space-time components (i.e. run from 0 to 3) while Latin indices label spatial components (i.e. run from 1 to 3). Furthermore, the Dirac delta and the Kronecker delta functions are denoted by the symbols δ_{D} and δ^{K}, respectively. Finally, the symbols c and G denote the speed of light in vacuum and the gravitational constant

respectively.

2.1 The ΛCDM model

2.1.1 The Homogeneous Universe

Cosmological Principle

The Cosmological Principle (CP) states that on sufficiently large separations the Universe looks about the same from any point of view. In other words, that the Universe is Homogeneous and Isotropic at any point to an observer who is at rest with the cosmic fluid, i.e. a 'Fundamental' observer (e.g., Peebles, 1980). A consequence of the CP is that the physical distance r between any two fundamental observers could only have a temporal dependence

$$r(t) = a(t)\, x\,, \tag{2.1}$$

where $a(t)$ is a generic temporal function known as the scale factor and x is the distance at some arbitrary point in time t_0 (usually called comoving distance). Conventionally, the scale factor is set to one for the current point in time t_0, i.e. $a(t_0) = 1$.

Taking the time derivative of Eq. (2.1), we can derive the Hubble law (Hubble, 1929) which relates the recession-velocity of any object to its position

$$\boldsymbol{v}_{\rm rec} \equiv \dot{r} = \dot{a}\, x = H(t)\, r(t)\,. \tag{2.2}$$

We define the Hubble parameter as $H(t) \equiv \dot{a}/a$, where the dot denotes the time derivative. We note that the Hubble constant is defined as $H_0 = H(t_0)$.

In order to describe the evolution of the Universe as a whole, we turn to the Einstein's field equations

$$G_{\mu\nu} + \Lambda\, g_{\mu\nu} = \frac{8\pi\, G}{c^4} T_{\mu\nu}\,. \tag{2.3}$$

The left side of the previous equation describes the geometry of the spacetime under consideration. It consists of the metric tensor $g_{\mu\nu}$, the Einstein tensor $G_{\mu\nu}$ (which can be constructed from $g_{\mu\nu}$ and its derivatives) and a cosmological constant Λ. The right hand side of Eq. (2.3) on the other hand describes the matter content of the Universe through the stress-energy tensor $T_{\mu\nu}$. In summary, Eq. (2.3) mathematically describes the intricate relationship between matter and the spacetime it inhabits. The first necessary ingredient for Eq. 2.3 is the metric tensor $g_{\mu\nu}$ which is a second order tensor that defines the inner product in (pseudo)Riemannian spaces. Equipped with a metric tensor, the line element $\mathrm{d}s^2$ can be defined as

$$\mathrm{d}s^2 = g_{\mu\nu}\, \mathrm{d}x^\mu\, \mathrm{d}x^\nu\,, \tag{2.4}$$

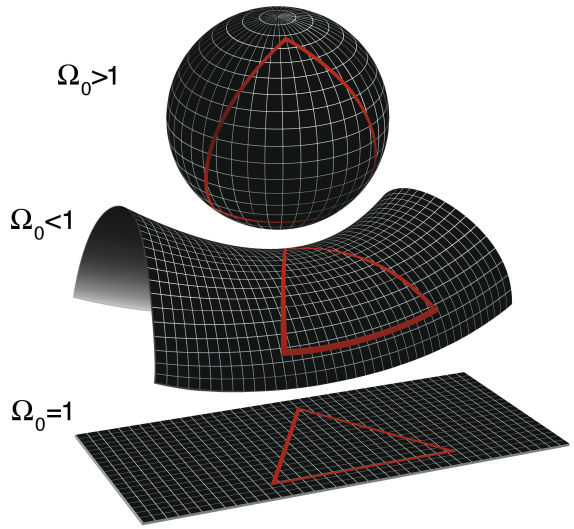


Figure 2.1: A diagram depicting the three possible curvatures of the Universe allowed by the line element, defined in Eq. (2.4). The top diagram represents a positive curvature ($k > 0$) and a spherical universe, the bottom depicts a flat universe with no curvature ($k = 0$) and finally the middle panel displays a hyperbolic universe, with negative curvature ($k < 0$). The labels indicate how the matter-energy of the universe determines each scenario (see Eq. 2.16). This Figure is taken from the NASA website.

where $\mathrm{d}x^\nu$ represents the infinitesimal change in the coordinates. A universe that follows the CP can be described by the FLRW metric (for a symmetry based derivation, see Hobson et al., 2006)

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a^2\left[\mathrm{d}x^2 + f_k^2(x)\mathrm{d}\Omega_x\right] , \tag{2.5}$$

where Ω_x is the solid angle, $k \in \{-1, 0, 1\}$ and f_k is the angular diameter distance in curved spacetimes. The functional form of $f_k(x)$ changes depending on the curvature of the metric

$$f_{\mathrm{k}}(x) = \begin{cases} R\,\sin(x/R) & k = 1 \text{ (Spherical)} \\ x & k = 0 \text{ (Flat)} \\ R\,\sinh(x/R) & k = -1 \text{ (Hyperbolic)} \end{cases} , \tag{2.6}$$

where R is the radius of curvature, with length dimensions, in the non-zero k cases. The adopted standard model of cosmology (ΛCDM) posits that the Universe is spatially flat ($k = 0$), based on multiple observations from independent probes (e.g., based on the Cosmic Microwave background Bennett et al., 2013, Planck Collaboration: Aghanim et al., 2020a).

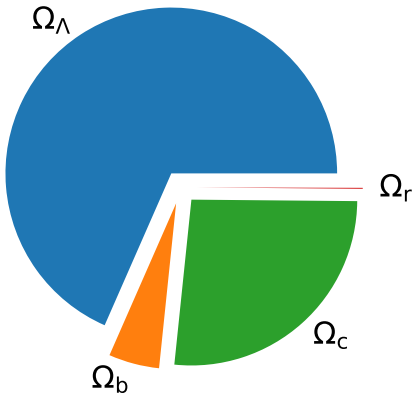


Figure 2.2: A pie chart describing the energy-density makeup of the Universe according to the ΛCDM model.

Energy-density makeup of the Universe.

Equipped with the metric, it is possible to solve Eq. (2.3) if the stress-energy tensor is modelled as a perfect fluid, i.e. a fluid can be solely described by the fluid's energy density ρ and isotropic pressure P. Mathematically, it is defined as

$$T^{\mu\nu} = \begin{pmatrix} \rho\, c^2 & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix} . \tag{2.7}$$

In that case, Eq. (2.3) reduces to the Friedmann–Lemaître (FL) equations

$$H^2(t) = \frac{8\pi\, G}{3}\rho(t) - \frac{k\, c^2}{a(t)^2} + \frac{\Lambda}{3} , \tag{2.8}$$

$$\frac{\ddot{a}(t)}{a(t)} = \frac{4\pi\, G}{3}\left[\rho(t) + \frac{3P(t)}{c^2}\right] + \frac{\Lambda}{3} , \tag{2.9}$$

$$\dot{\rho}(t) = -3\, H(t)\left[\rho(t) + \frac{P(t)}{c^2}\right] . \tag{2.10}$$

Solving the previous three equations, requires an equation of state for the fluid under consideration. The ΛCDM model admits three types of fluids, each has the following equation of state

$$P = w\, \rho\, c^2 , \tag{2.11}$$

where w is a constant value. The evolution behaviour of such a fluid can be derived from Eq. (2.10) to be

$$\rho(t) = \rho(t_0) a(t)^{-3(1+w)} . \tag{2.12}$$

The three energy densities that make up the ΛCDM model are

- *Matter* energy density ρ_m: Matter travels with non-relativistic speed and

consequently has vanishing pressure ($w = 0$) and consequently $\rho_{\rm m} \propto a^{-3}$. It encompasses both cold dark matter $\rho_{\rm c}$ which interacts only gravitationally and baryonic matter $\rho_{\rm b}$ which interacts both gravitationally and electromagnetically.

- *Radiation* energy density $\rho_{\rm r}$: Radiation describes massless particles (photons) and massive (e.g., neutrinos) particles that travel at relativistic speeds. $w = 1/3$ for these particles and $\rho_{\rm r} \propto a^{-4}$.

- *Dark* energy density ρ_Λ : The dark energy density is sourced by the cosmological constant in the FL equations ρ_Λ =const as $w = -1$. That energy density was added to the model to explain the ground-breaking observations of the accelerating expansion of the Universe (Perlmutter et al., 1999).

The total energy density of the Universe is simply the sum of all the different contributions

$$\rho = \rho_{\rm m} + \rho_{\rm r} + \rho_\Lambda \,, \tag{2.13}$$

where we show their relative contribution to the total energy density of the Universe in Fig. 2.2. It is standard practice to define the energy density parameters [1]

$$\Omega_{\rm i}(t) = \frac{\rho_{\rm i}(t)}{\rho_{\rm crit}} \,, \tag{2.14}$$

where the critical density is defined as $\rho_{\rm crit} = 3H(t)^2/(8\pi\, G)$. Eq. 2.8 can then be recast into

$$H^2(a) = H_0^2[\Omega_{\rm r,0}/a^4 + \Omega_{\rm m,0}/a^3 + (1-\Omega_0)/a^2 + \Omega_\Lambda] \,. \tag{2.15}$$

We note that we use the subscript 0 to denote the present day values of the density parameters and that we define

$$\Omega_0 = 1 - k\, c^2/H_0^2 = 1 - (\Omega_{\rm r,0} + \Omega_{\rm m,0} + \Omega_\Lambda) \,. \tag{2.16}$$

Cosmological distances

Finally, we define the various cosmological distance proxies used in this work. Starting with the redshift, it is defined as the fractional change in the wavelength of the received light due to the expansion of the Universe. It can be written as

$$z = \frac{\lambda_{\rm o} - \lambda_{\rm e}}{\lambda_{\rm e}} \,, \tag{2.17}$$

where $\lambda_{\rm e}$ and $\lambda_{\rm o}$ are the wavelength of the light emitted by a source at the time of emission and observation respectively. Using the metric defined in Eq. (2.5) and the geodesic equations of a particle with a null worldline, it can be shown that

[1] We note that we use $\Omega_{\rm i}$ with a Roman subscript to denote the energy density parameter while Ω_x with an italic subscript symbolizes the solid angle of the vector $\boldsymbol{x}$.

$\lambda \propto a^{-1}$(Peebles, 1980). Consequently, for the FLRW metric, the following relation can be derived

$$a = \frac{1}{1+z}\,. \tag{2.18}$$

Cosmological distances vary depending on the method of measurement. Starting with the comoving distance defined in Eq. (2.1), it can be computed for an FLRW metric using the measured redshift through

$$x(z) = \int_0^z \frac{c\,\mathrm{d}z'}{H(z')}\,. \tag{2.19}$$

The other two types of distances of interest to this work are the

- **Luminosity distance** D_L: the luminosity distance relates the observed flux bolometric[2] F of a light emitting source to its source bolometric Luminosity L through $L = 4\pi\, D_\mathrm{L}^2\, F$. The luminosity distance is given by (Weinberg, 1972)

$$D_\mathrm{L} = (1+z)\,x\,. \tag{2.20}$$

- **Angular distance** D_A: The angular (diameter) distance relates an objects angular extension θ to its physical radius R through $D_\mathrm{A} = 2R/\theta$. It is computed via (Weinberg, 1972)

$$D_\mathrm{A} = f_k(x)/(1+z)\,. \tag{2.21}$$

2.1.2 The Inhomogeneous Universe

The CP (fortunately) only holds for scales larger than $70\,h^{-1}\,\mathrm{Mpc}$ (e.g., see Yadav et al., 2005). On smaller scales, the Universe is full of diverse structures, ranging from planetary systems and stellar clusters at the intragalactic level to galaxy clusters as seen in the Fig. 2.3 created using data from the *SDSS* survey. The theory of structure formation posits that gravitational instability is the primary force driving the creation of these intricate objects (Zel'dovich, 1970, Dodelson, 2003). According to the standard cosmological model, minute quantum fluctuations undergo amplification during the inflationary epoch. Consequently, overdense regions stemming from these stochastic fluctuations initiate matter aggregation, progressively expanding in size and gravitational influence. As more matter falls into these expanding potential wells, it collapses into distinct, bound structures, detaching from cosmological expansion. This process results in the emergence of halos, voids, and, on the largest scales, the cosmic web. It's noteworthy that at both galactic and intergalactic scales, the evolutionary trajectory of structures is further influenced by baryonic matter, which interacts gravitationally and electromagnetically, in contrast to dark matter's exclusively gravitational interactions.

[2] after integrating over all frequencies.

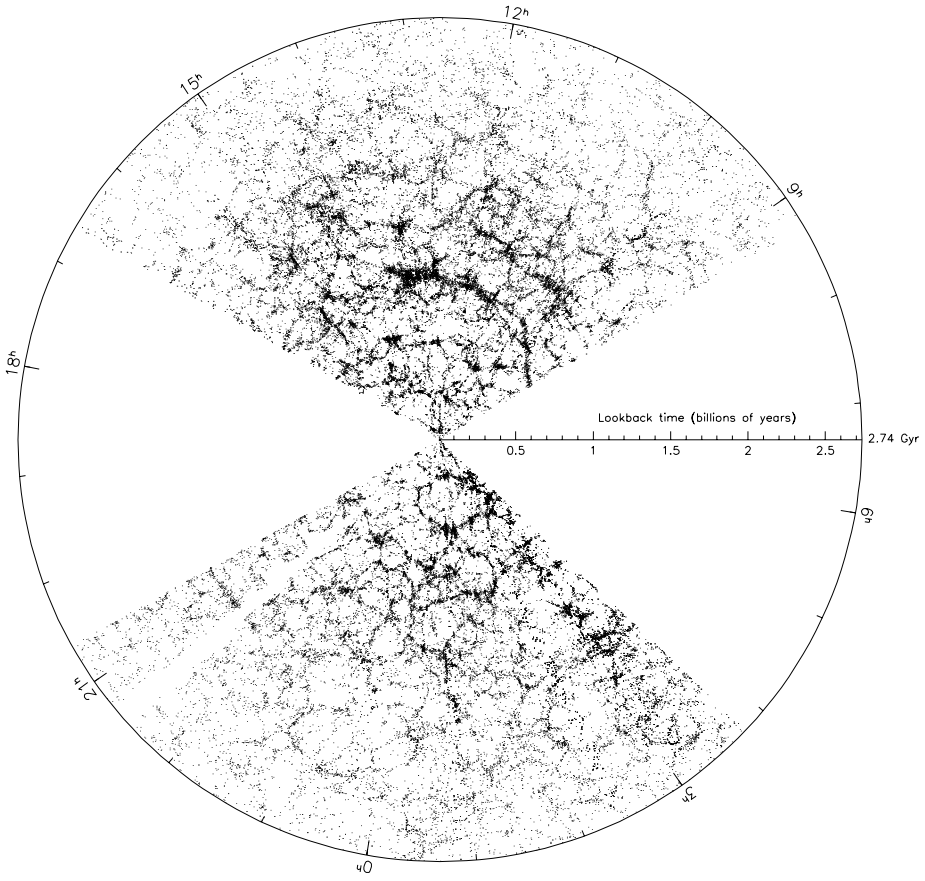


Figure 2.3: An illustrative figure of the structure of the local universe compiled from the data of the *SDSS* survey. The figure is taken from Gott III et al. (2005).

Linear perturbation theory

Due to the non-linear nature of gravity, solving the Einstein system of equations (see Eq. 2.3) for realistic matter models that have a spatial dependence is extremely difficult due to the inherent non-linearity of the system. One of the standard approaches to tackle that issue is the so-called Newtonian fluid approximation. The Newtonian treatment is only valid at scales smaller than the horizon, for larger scales a fully general relativistic treatment is needed (see Sec. 2.3.1 for the relation between the relativistic and Newtonian treatment). In the Newtonian fluid approximation, the DM particle distribution is approximated by a fluid with a density $\rho(\boldsymbol{x}, t)$ that follows the following system of equations (hereafter, the fluid equations Peebles, 2020)

$$\text{(Continuity)} \qquad \frac{\partial \rho(\boldsymbol{r}, t)}{\partial t} + \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot (\rho\, \boldsymbol{u}(\boldsymbol{r}, t)) = 0 \quad , \tag{2.22}$$

$$\text{(Momentum)} \qquad \frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \boldsymbol{\nabla}_r)\boldsymbol{u} = -\boldsymbol{\nabla}_r \Phi \quad , \tag{2.23}$$

$$\text{(Poisson)} \qquad \nabla_r^2 \Phi = 4\pi G \rho - \Lambda \quad , \tag{2.24}$$

where $\boldsymbol{u}$ denotes the DM velocity field and Φ is the Newtonian gravitational potential. Equation (2.22) is derived from the conservation of matter. Equation (2.23) (also known as the Euler equation) describes the conservation of momentum. Finally, Eq. (2.24) relates the gravitational potential to the matter density. The fluid approximation breaks down at small scales after shell-crossing occurs. As it is no longer possible to provide a well-defined velocity field (Zel'dovich, 1970). The previous equations are also invalid at Horizon scales, $d_H = c/H(t)$ where the Newtonian approximation fails.

The fluid equations (FE) remain too complex for analytical solutions. To tackle the FE system, perturbation theory is usually employed as follows. First, the system is rewritten in terms of the matter overdensity

$$\delta(\boldsymbol{r}, t) \equiv \frac{\rho(\boldsymbol{r}, t) - \bar{\rho}(t)}{\bar{\rho}(t)} \, . \tag{2.25}$$

A change of variables is also employed to rewrite the FE system in comoving coordinates instead of proper coordinates, i.e.

$$\boldsymbol{r} = a(t)\boldsymbol{x} \, , \tag{2.26}$$

$$\rho(\boldsymbol{r}, t) = \rho(\boldsymbol{x}, t) \, , \tag{2.27}$$

$$\boldsymbol{u}(\boldsymbol{r}, t) = \dot{a}\boldsymbol{x} + \boldsymbol{v}(\boldsymbol{x}, t) \, , \tag{2.28}$$

$$\varphi(\boldsymbol{x}, t) = \Phi(a(t)\boldsymbol{x}, t) + \frac{\ddot{a}(t)a(t)}{2}|\boldsymbol{x}|^2 \, . \tag{2.29}$$

Where we define the comoving potential φ in the synchronous-comoving gauge and the peculiar velocity field $\boldsymbol{v} = \mathrm{d}x/\mathrm{d}t$.

Plugging these relations into the fluid equations, they transform to

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\boldsymbol{\nabla}_{\boldsymbol{x}} \cdot [(1+\delta)]\, \boldsymbol{v} = 0\,, \tag{2.30}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + \frac{\dot{a}}{a}\boldsymbol{v} + \frac{1}{a}(\boldsymbol{v}\cdot\boldsymbol{\nabla}_x)\,\boldsymbol{v} = -\frac{1}{a}\boldsymbol{\nabla}_x\varphi\,, \tag{2.31}$$

$$\nabla_x^2\varphi = \frac{3H_0^2\Omega_{\rm m}}{2a}\delta\ . \tag{2.32}$$

If we restrict our interest to large scales, we can consider only regimes where δ and v are perturbatively small. In that case, we can linearize the fluid equations (Sahni, 1995)

$$\frac{\partial \delta}{\partial t} = -\frac{1}{a}\boldsymbol{\nabla}_{\boldsymbol{x}} \cdot \boldsymbol{v}\,, \tag{2.33}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + \frac{\dot{a}}{a}\boldsymbol{v} = -\frac{1}{a}\boldsymbol{\nabla}_x\varphi\,, \tag{2.34}$$

$$\nabla_x^2\varphi = \frac{3H_0^2\Omega_{\rm m}}{2a}\delta\,. \tag{2.35}$$

Combining the linear fluid equations, we find an equation describing the temporal evolution of the matter overdensity

$$\frac{\partial^2\delta}{\partial t^2} + \frac{2\dot{a}}{a}\frac{\partial\delta}{\partial t} - \frac{3H_0^2\Omega_m}{2a^3}\delta = 0\,. \tag{2.36}$$

The general solution to Eq. (2.36) is a linear combination of products of purely spatial and purely temporal functions (Dodelson, 2003), i.e.

$$\delta(\boldsymbol{x},t) = D_+(t)\,\Delta_+(\boldsymbol{x}) + D_-(t)\,\Delta_-(\boldsymbol{x})\,. \tag{2.37}$$

Only the temporal functions can be determined from Eq. (2.36) as the equation has no spatial dependence. Ignoring the decaying solution, the growing solution for generic values of the cosmological parameters, in terms of the scale factor a, is given by (Peebles, 1980)

$$D_+(a) = \frac{5\Omega_m}{2}H_0^2\,H(a)\int_0^a \frac{{\rm d}a'}{[a'H(a')]^3}\,. \tag{2.38}$$

Assuming that the velocity field $\boldsymbol{v}(\boldsymbol{x},t)$ can be written as the gradient of a scalar potential[3] $\boldsymbol{v} = \boldsymbol{\nabla}_x\phi_v$, we can use the linear continuity equation (Eq. 2.33) to solve for the velocity field. Plugging in only the growing mode for the overdensity $\delta(\boldsymbol{x},t) = D_+(t)\Delta_+(\boldsymbol{x})$ into the continuity equation

$$(\boldsymbol{\nabla}\cdot v) = \nabla^2\phi_v = -a\frac{\dot{D}_+}{D_+}\delta\,. \tag{2.39}$$

[3]This assumption is valid at linear scales as rotational modes can be shown to vanish from Eq. (2.33) (Sahni, 1995, Dodelson, 2003).

Defining the growth parameter

$$f = \frac{\mathrm{d}\ln D_+}{\mathrm{d}\ln a}\,, \tag{2.40}$$

we can solve Eq. (2.39) using Green functions to find

$$\boldsymbol{v}(\boldsymbol{x}, a) = \frac{f\, a\, H(a)}{4\pi} \int \delta(\boldsymbol{x'}, a) \frac{\boldsymbol{x'} - \boldsymbol{x}}{|\boldsymbol{x'} - \boldsymbol{x}|^3} \mathrm{d}^3 x'\,. \tag{2.41}$$

2.1.3 Two-Point Statistics

Two-point Correlation function

Due to the complexity of the Universe, we can only study the spatial part of the field $\delta(x, z)$ statistically[4]. To that end, we employ the theory of random fields. In that paradigm, the Universe is treated as a single realisation of a Gaussian (or nearly Gaussian) random process. A Gaussian random field is completely described by its 2PCF

$$\langle \delta(\boldsymbol{x}_1, z)\, \delta(\boldsymbol{x}_2, z) \rangle = C(\boldsymbol{x}_1, \boldsymbol{x}_2, z)\,. \tag{2.42}$$

If we enforce the CP on that random field, it implies that the field δ is both translationally and rotationally invariant. Consequently, the 2PCF depends only on the separation of the two points at which it is computed

$$\langle \delta(\boldsymbol{x}_1, z)\, \delta(\boldsymbol{x}_2, z) \rangle = \xi(|\boldsymbol{x}_2 - \boldsymbol{x}_1|, z) \equiv \xi(r, z)\,, \tag{2.43}$$

where $\boldsymbol{r} = \boldsymbol{x}_2 - \boldsymbol{x}_1$.

Another possible definition for the 2PCF –originally proposed in (Peebles, 1980) – is the excess probability of finding two objects located at $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ respectively at a separation $|\boldsymbol{x}_2 - \boldsymbol{x}_1|$ w.r.t. a Poisson distribution

$$\mathrm{d}P = \bar{n}^2(z)[1 + \xi(r, t)]\mathrm{d}^3 x_1\, \mathrm{d}^3 x_2\,, \tag{2.44}$$

where $\bar{n}$ is the average number density. The definition in Eq. (2.44) has an understandable physical interpretation. It is easy to see that for strongly clustered particles, $\xi > 0$ while it drops below zero for empty regions. Having no correlation ($\xi = 0$) implies the objects are randomly distributed.

The previous analysis focused on correlating the dark matter density contrast. However, in many scenarios, direct observation of δ is unattainable. Instead, we typically observe proxies or tracers of dark matter. In this study, we concentrate on galaxy redshift surveys. Hence, our aim is to establish a connection between the observed galaxy overdensity δ_g and the underlying dark matter overdensity. We

[4] We use the redshift to parameterize the time dependence of the overdensity

define the the observed galaxy overdensity as

$$\delta_{\rm g}(\boldsymbol{x}, z) = \frac{n_{\rm g}(\boldsymbol{x}, z) - \bar{n}_{\rm g}(z)}{\bar{n}_{\rm g}(z)} \,. \tag{2.45}$$

where $n_{\rm g}(\boldsymbol{x}, z)$ are the galaxy number density and $\bar{n}_{\rm g}(z)$ is its corresponding spatial average. To achieve this, we adopt a straightforward linear biasing relation, derived from the threshold biasing approach (Peacock & Heavens, 1985):

$$\delta_{\rm g}(\boldsymbol{x}, z) = b(z)\, \delta(\boldsymbol{x}, z) \,, \tag{2.46}$$

where b is the linear bias parameter. We note the relation is only valid for linear scales (where $|\delta| \ll 1$) (For a review of non-linear perturbation theory for biased tracers, see Bernardeau et al., 2002). Consequently, the galaxy 2PCF is related to its matter counterpart through

$$\xi_{\rm g}(\boldsymbol{r}, z) = b^2(z)\, \xi(\boldsymbol{r}, z) \,. \tag{2.47}$$

An example of the galaxy 2PCF is shown in Fig. 2.4b constructed using the `CAMB` (Lewis & Bridle, 2002) and `HANKL` (Karamanis & Beutler, 2021) libraries, it is clear that the galaxies are correlated at small separations as expected from the gravitational instability paradigm. At $r = 100\, h^{-1}\,$Mpc, we see the imprint of the baryonic acoustic oscillations (Cole et al., 2005, Eisenstein et al., 2005). Subsequently, the 2PCF crosses zero at $r = 120\, h^{-1}\,$Mpc to show anticorrelation at large scales.

Power Spectrum

Another statistic of great importance to cosmology is the power spectrum. Widely utilized in extracting cosmological information from galaxy redshift surveys, it serves as a cornerstone for parameter inference and analysis (e.g., Eisenstein et al., 2005, Yamamoto et al., 2006, Wilson et al., 2017, Beutler et al., 2014). To establish its foundation, let us start by defining the Fourier transform of the overdensity field:

$$\tilde{\delta}(\boldsymbol{k}, z) = \int \delta(\boldsymbol{x}, z) e^{{\rm i}\boldsymbol{k}\cdot\boldsymbol{x}}\, {\rm d}^3x \,, \tag{2.48}$$

we can then define the power spectrum as the Fourier-space autocorrelation function, i.e.

$$\langle \tilde{\delta}(\boldsymbol{k}, z) \tilde{\delta}(\boldsymbol{k'}, z) \rangle = (2\pi)^3 \delta_{\rm D}(\boldsymbol{k} + \boldsymbol{k'}) P(k, z) \,, \tag{2.49}$$

where we again see the implication of the isotropy and homogeneity of the overdensity field. As it leads to the diagonality of the power spectrum as shown in Eq. (2.49), in other words, the Fourier space autocorrelation only depends on the magnitude of the wavevector k.

Using Eqs. (2.49) and (2.43), it is simple to show that the power spectrum and

the 2PCF are Hankel pairs

$$P(k,z) = 4\pi \int r^2 j_0(k\,r)\xi(r,z)\,. \tag{2.50}$$

Finally, regarding the redshift evolution of the power spectrum, it can be shown from Eq. (2.36) that $\delta(k,z) = D_+\delta(k,z=0)$, consequently the evolution of the linear power spectrum

$$P(k,z) = D_+^2(z)\,P(k,z=0)\,. \tag{2.51}$$

We can then build the galaxy power spectrum using Eq. (2.46) and the previous relation to find

$$P_{\rm g}(k,z) = b^2(z)\,D_+^2(z)\,P(k,z=0)\,. \tag{2.52}$$

An example of the galaxy power spectrum is shown in Fig. 2.4a. As seen in the Figure, The power spectrum peaks at $k \approx 0.1\,h\,\mathrm{Mpc}^{-1}$. We can then see that the BAO causes oscillations in Fourier space for $k > 0.05\,h\,\mathrm{Mpc}^{-1}$ instead of a peak as seen in the 2PCF. Finally, we note that the power spectrum is positive-definite by construction.

Angular Power Spectrum

The last two-point statistic of interest to this work is the angular power spectrum of the large-scale-structure (LSS, Peebles, 1973) . It measures the clustering power of the galaxy distribution as a function of the angular scale, i.e. after projecting on a spherical surface. It is usually employed in photometric surveys where the redshift measurements have large statistical errors (Loureiro et al., 2019).

We start with the projected galaxy overdensity

$$\Sigma_{\rm g}(\hat{\boldsymbol{x}}) = \int \delta_{\rm g}(\hat{\boldsymbol{x}}, x(z))\,\mathcal{P}(z)\,\mathrm{d}z\,, \tag{2.53}$$

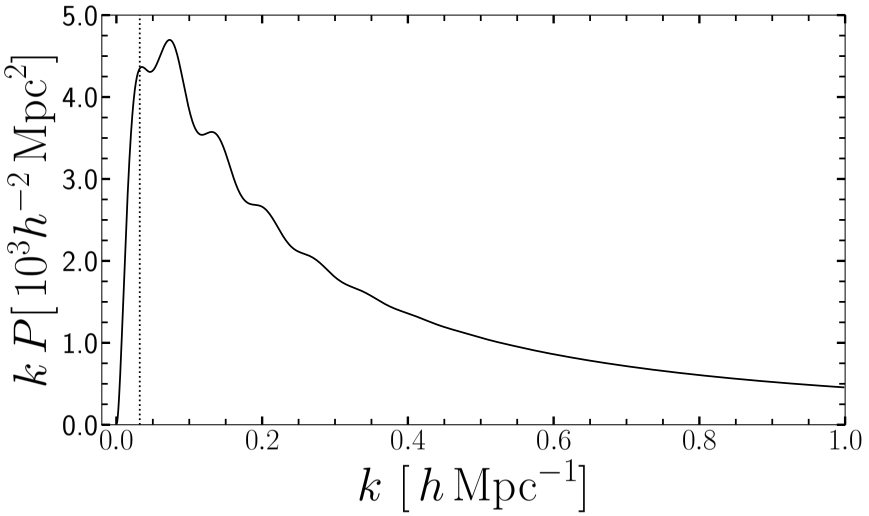
where $\mathcal{P}(z)$ is the normalized radial galaxy distribution, i.e. $\mathcal{P}(z) = \bar{n}_{\rm g}(z)/\int \bar{n}_{\rm g}(z)\,\mathrm{d}z$. We can then compute its harmonic expansion on the sphere

$$a_{\ell m} = \int \Sigma(\hat{\boldsymbol{x}})\,Y_{\ell m}(\hat{\boldsymbol{x}})\,\mathrm{d}\Omega_x\ , \tag{2.54}$$

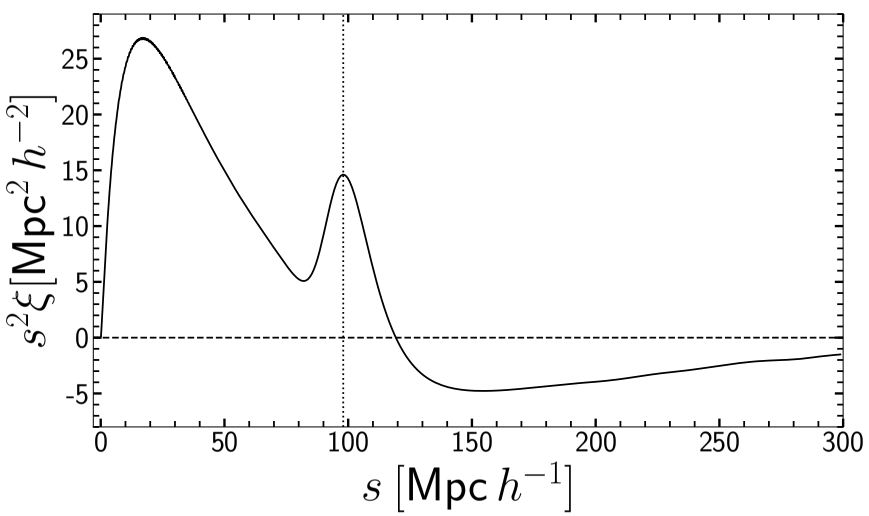
where $Y_{\ell m}(\hat{\boldsymbol{x}})$ are the spherical harmonic functions. We are then able to define the angular power spectrum as

$$C_\ell = \left\langle |a_{\ell m}|^2 \right\rangle\ . \tag{2.55}$$

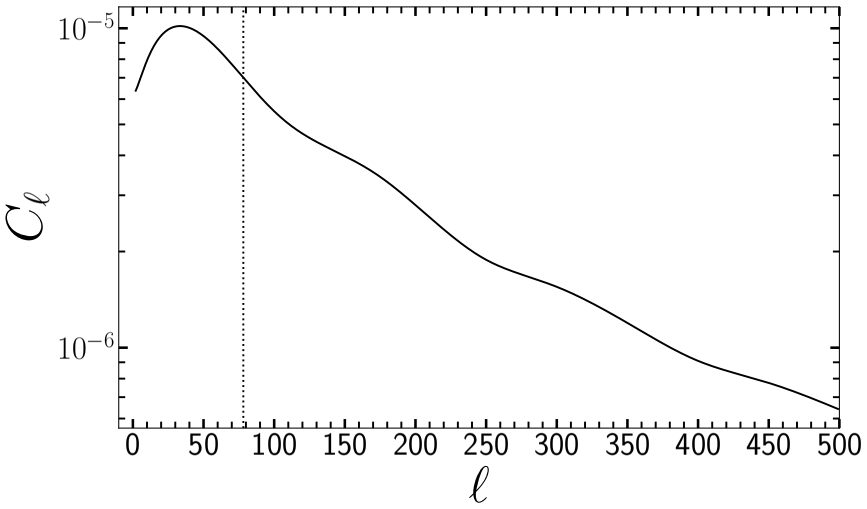
The angular power spectrum can be theoretically related to the power spectrum through (e.g., Fisher et al., 1994, Scharf et al., 1992, Padmanabhan et al., 2007,



(a) Example of the galaxy power spectrum.



(b) Example of the galaxy 2PCF.



(c) Example of the galaxy angular power spectrum.

Figure 2.4: We show examples of the three two-point statistics considered in this thesis. All three panels are computed at $z = 1.0$, with bias $b = 1.46$. These figures were created using the `CAMB` and `HANKL` libraries. The dotted lines mark the first (and only in the 2PCF case) BAO signature.

Loureiro et al., 2019)

$$C_\ell = \frac{2}{\pi} \int \mathcal{W}_\ell^2(k)\, k^2\, P(k, z=0)\, \mathrm{d}k\,. \tag{2.56}$$

The weight function is defined as

$$\mathcal{W}_\ell(k) = \int b(z)\, \mathcal{P}(z) D(z)\, j_\ell(k\, x(z)) \mathrm{d}z\,, \tag{2.57}$$

where j_ℓ are the spherical Bessel functions.

An example of the angular power spectrum is shown in Fig. 2.4c. As seen in the figure, the angular power spectrum peaks at $\ell \approx 40$, after which we observe the imprint of BAO as oscillations.

2.2 Redshift Space Distortions

In galaxy redshift surveys, our observations are limited to the angular positions, $\boldsymbol{n}$, and the redshifts denoted as z, of light-emitting galaxies. Constructing maps of the Universe necessitates a conversion from these measured redshifts and angular coordinates to physical distances. For a universe described by the FLRW metric (see Eq. 2.4), this conversion is is given by

$$x(z) = \int_0^z \frac{c\, \mathrm{d}z'}{H(z')}\,. \tag{2.58}$$

However, the above expression, denoted as Eq. (2.58), oversimplifies the intricacies of the Universe's structure (refer to Section 2.1.2), which significantly affect the observed redshifts of galaxies (Sargent & Turner, 1977). Consequently, the observed redshifts are not solely sourced by the cosmological expansion. The discrepancies between the observed galaxy positions and those predicted by Eq. (2.58) are collectively referred to as Redshift Space Distortions (RSD). In this section, we delineate the classical approach to RSD initially developed by Kaiser (1987) in his pioneering work (also see Hamilton, 1998, Hamilton & Culhane, 1996). Subsequently, we delve into the relativistic treatment that was subsequently developed to account for all RSD up to first order in cosmological perturbations (Yoo, 2010, Bonvin & Durrer, 2011, Challinor & Lewis, 2011, Jeong et al., 2012). This relativistic framework serves as the foundation for the `LIGER` technique (see Chapter 3); the main numerical method employed throughout this work.

2.2.1 'Classical' Approach.

The classical RSD treatment developed by Kaiser in 1987 focuses primarily on the alteration of the observed redshift resulting from the peculiar velocities of galaxies and the observer. This change in the observed redshift can be modelled as a relativistic Doppler shift.

Co-ordinate Map

Consider a source that is comoving with the cosmic expansion, emitting light at a specific wavelength, λ_e. Upon reception by a comoving observer, as discussed in Section 2.1.1, this light is perceived at a different wavelength, λ_o due to the expansion of the Universe. Specifically, the ratio of these wavelengths, $\lambda_o/\lambda_e = (1 + z_{cos})$, defines the cosmological redshift, denoted as z_{cos} (as indicated in the FLRW model shown in Eq. 2.17).

If the source has a peculiar velocity $\boldsymbol{v}_e$ at emission, then the observer measures a different wavelength for the light signals. In this case, the observed redshift z_{obs} is given by the expression

$$1 + z_{obs} = \lambda_o/\lambda_e = (\lambda_o/\lambda'_e)\,(\lambda'_e/\lambda_e) = (1 + z_{cos})\,(1 + z_{pec,e})\,, \tag{2.59}$$

where λ'_e denotes the wavelength measured by a comoving observer at the event of emission, which can be computed from λ_e using a simple Lorentz transformation, i.e.

$$z_{pec,e} = \frac{1 + (\boldsymbol{v}_e \cdot \hat{\boldsymbol{x}})/c}{\sqrt{1 - (v_e/c)^2}} - 1 \simeq \frac{\boldsymbol{v}_e \cdot \hat{\boldsymbol{x}}}{c} \equiv \frac{v_{\parallel e}}{c}\,. \tag{2.60}$$

The peculiar velocity of the observer similarly changes the observed redshift, $1 + z_{obs} = \lambda_o/\lambda_e = (\lambda'_o/\lambda_e)\,(\lambda_o/\lambda'_o) = (1 + z_{cos})\,(1 + z_{pec,o})$ where λ'_o denotes the wavelength measured by a comoving observer at the event of detection. In this case,

$$z_{pec,o} = \frac{\sqrt{1 - (v_o/c)^2}}{1 + (\boldsymbol{v}_o \cdot \hat{\boldsymbol{x}})/c} - 1 \simeq -\frac{\boldsymbol{v}_o \cdot \hat{\boldsymbol{x}}}{c} = -\frac{v_{\parallel o}}{c}\,. \tag{2.61}$$

Combing both distortions to the redshift, the observed redshift is given by

$$1 + z_{obs} = (1 + z_{cos})\,(1 + z_{pec,e})\,(1 + z_{pec,o})\,. \tag{2.62}$$

Consequently, the change in the observed redshift is given by

$$\boxed{\frac{\Delta z}{1 + z} = \frac{(v_{\parallel,e} - v_{\parallel,o})}{c}\,.} \tag{2.63}$$

It is important to note that by using Eq. (2.58), we implicitly set the scale factor

$$a = \frac{1}{(1 + z_{obs})}\,. \tag{2.64}$$

As evident by Eq. (2.63), this implies that $a \neq (1 + z_{cos})^{-1}$. We articulate that the equations derived below are written in terms of the **observed** redshift position, i.e. in terms of z_{obs} and its corresponding FLRW distance $\boldsymbol{x}$. At linear order, that fact leads to a difference in the average comoving density (Hamilton, 1998, Bahr-Kalus

et al., 2021) and at higher orders it changes the form of the shift equations (Bertacca et al., 2014a).

We can use the redshift change shown in Eq. (2.63) to calculate how the actual (real-space) position of the galaxy, $\boldsymbol{x}_{\rm r}$, relates to the observed (redshift-space) position $\boldsymbol{x}$

$$\boldsymbol{x}_{\rm r} = \boldsymbol{x} + \delta\boldsymbol{x}\,, \tag{2.65}$$

where $\delta\boldsymbol{x}$ is the coordinate map. We use Eq. (2.58) to derive $\delta\boldsymbol{x}$ as follows

$$\begin{aligned}
\boldsymbol{x} &= x\,\hat{\boldsymbol{x}} \\
&= \left[\int_0^{z_{\rm obs}} \frac{c}{H(z')}\,{\rm d}z'\right]\hat{\boldsymbol{x}} \\
&= \left[\int_0^{z_{\rm cos}} \frac{c}{H(z')}\,{\rm d}z' + \int_{z_{\rm cos}}^{z_{\rm obs}} \frac{c}{H(z')}\,{\rm d}z'\right]\hat{\boldsymbol{x}} \\
&\approx \left[x_{\rm r} + \frac{c\,(z_{\rm obs} - z_{\rm cos})}{H(z_{\rm obs})}\right]\hat{\boldsymbol{x}}\,.
\end{aligned} \tag{2.66}$$

We thus find that the peculiar velocities only change the observed positions along the line of sight. Analytically, we can write the radial change as

$$\delta\boldsymbol{x} = \delta x\,\hat{\boldsymbol{x}}\,, \tag{2.67}$$

where

$$\boxed{\delta x = (1+z_{\rm obs})\,\frac{v_{\parallel,{\rm e}} - v_{\parallel,{\rm o}}}{H(z_{\rm obs})} \equiv \frac{v_{\parallel,{\rm e}} - v_{\parallel,{\rm o}}}{a\,H(z_{\rm obs})}\,.} \tag{2.68}$$

We display a schematic of Eq. (2.67) in Fig. 2.5.

Impact on the observed galaxy contrast

In this section, we derive the impact of this shift on the observed galaxy density and the observed galaxy density contrast at linear order. It is pivotal to clarify the range of scales where the results presented in this section are valid. We present the impact of RSD at different scales in Fig. 2.6. In the figure, a spherical galaxy distribution is shown in both realspace (left) and redshift space (right). At all scales, only the red and green dots have a non-zero velocity component along the line of sight. Starting with the large scales (the top row), the spherical overdensity looks squashed as compared to its real space counterpart in what is known as the Kaiser effect (Kaiser, 1987). As the scales considered decrease, the peculiar velocity increases leading to the shell collapsing to a single sheet as seen in the middle row. These scales are aptly named the turnaround scales (Hamilton, 1998). Finally, as we go to even smaller scales, the positions of galaxies are flipped leading to the so-called "Finger of God" effect (Jackson, 1972, Sargent & Turner, 1977). For the purposes of this thesis and its related publications, we are only interested in linear scales.

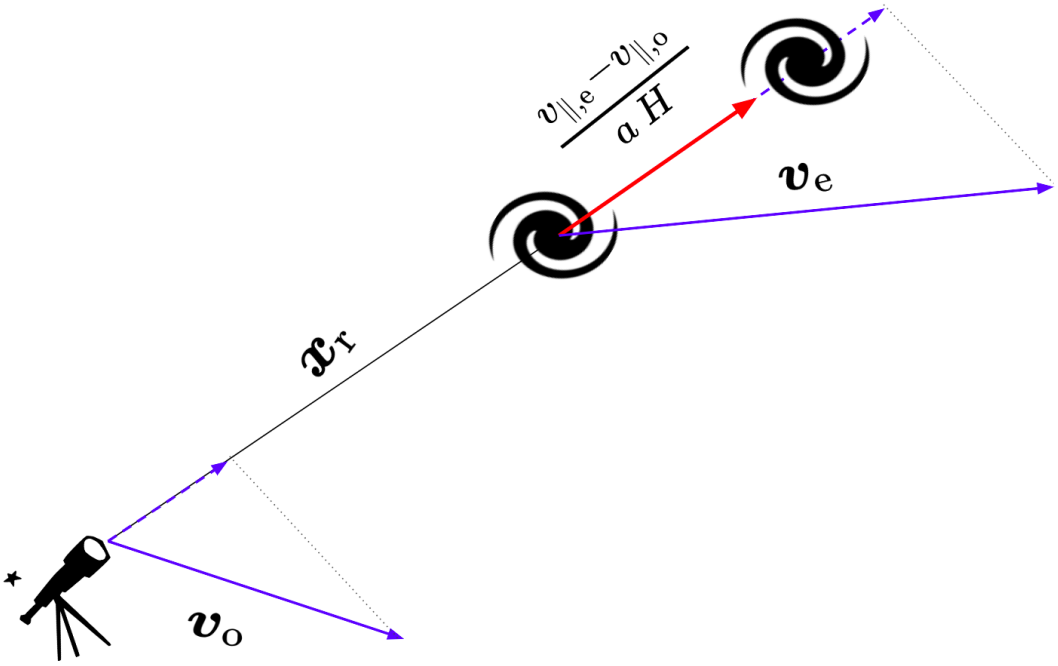


Figure 2.5: A schematic representation of Eq. (2.67). A galaxy at a distance $\boldsymbol{x}_\mathrm{r}$ is shifted by the peculiar velocities of both the source and the observer along the line of sight of the source.

We now turn to the derivation of the impact of RSD on the observed galaxy overdensity $\delta_{\mathrm{g,s}}$. The natural starting point for the derivation of the impact of the peculiar velocity is the conservation of galaxy number counts as RSD only affect the observed positions of galaxies. In other words, the total number of galaxies $\mathcal{N}$ can be expressed as

$$\mathcal{N} = \int n_{\mathrm{g,s}}(\boldsymbol{x})\mathrm{d}^3x = \int n_{\mathrm{g}}(\boldsymbol{x}_\mathrm{r})\mathrm{d}^3x_\mathrm{r}\,, \tag{2.69}$$

where $n_{\mathrm{g,s}}$ and n_{g} are the galaxy number density in redshift and real-space respectively. Equating the number of galaxies in an infinitesimal volume, we can derive the observed galaxy density from Eq. (2.69) as follows

$$\begin{aligned} n_{\mathrm{g,s}}(\boldsymbol{x})\mathrm{d}^3x &= n_{\mathrm{g}}(\boldsymbol{x}_\mathrm{r})\mathrm{d}^3x_\mathrm{r}\,, \\ \bar{n}_{\mathrm{g,s}}(x)[1+\delta_{\mathrm{g,s}}(\boldsymbol{x})]\mathrm{d}^3x &= \bar{n}_{\mathrm{g}}(x_\mathrm{r})[1+\delta_{\mathrm{g}}(\boldsymbol{x}_\mathrm{r})]\mathrm{d}^3x_\mathrm{r}\,, \end{aligned} \tag{2.70}$$

where we used $n_\mathrm{i} = \bar{n}_\mathrm{i}\,(1+\delta_\mathrm{i})$ for both n_g and $n_{\mathrm{g,s}}$.

Isolating the redshift space galaxy overdensity, $\delta_{\mathrm{g,s}}$, we find

$$\delta_{\mathrm{g,s}}(\boldsymbol{x}) = \frac{\bar{n}_{\mathrm{g}}(x_\mathrm{r})}{\bar{n}_{\mathrm{g,s}}(x)}[1+\delta_{\mathrm{g}}(\boldsymbol{x}_\mathrm{r})]\left|\frac{\mathrm{d}^3x_\mathrm{r}}{\mathrm{d}^3x}\right| - 1\,. \tag{2.71}$$

From the previous relation, we can see that deriving $\delta_{\mathrm{g,s}}$ requires three ingredients; the real-space galaxy contrast δ_g which is predicted through perturbation theory

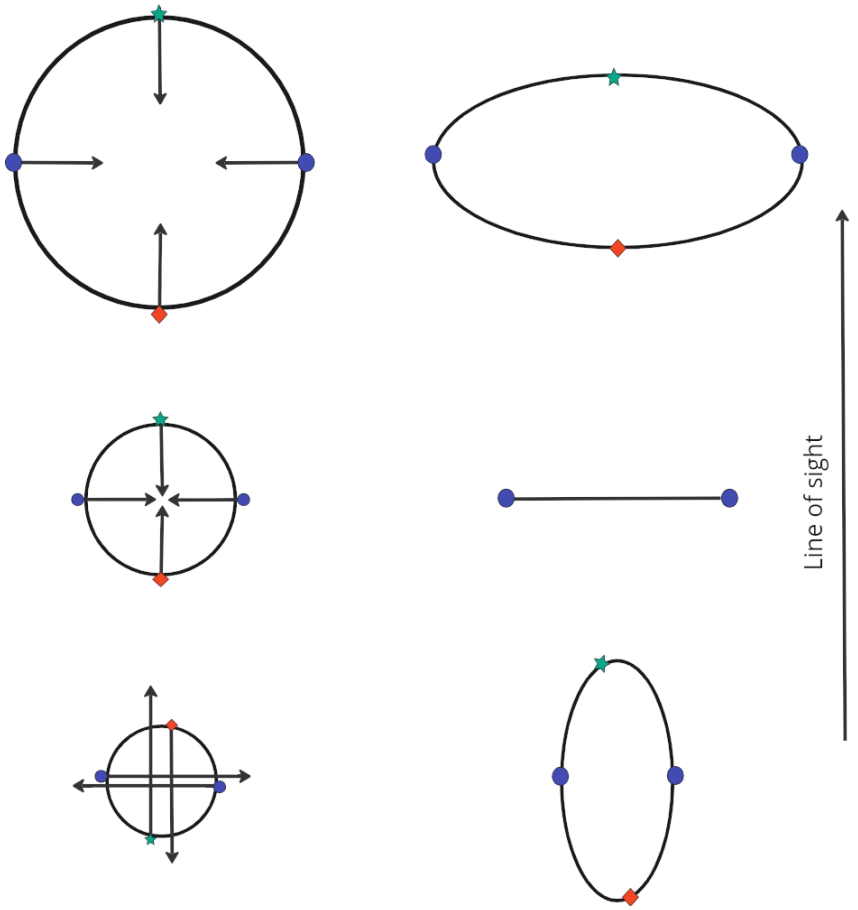


Figure 2.6: A schematic representation of RSD at different scales. The left column shows the real space positions of a hypothetical spherical galaxy distribution. Each row represents a spherical shell sampled by the dots, while the arrows denote the peculiar velocity field of that shell. The right column displays the observed positions for the spherical overdensity (their redshift space positions). This plot is based on a similar figure in (Hamilton, 1998)

(see Sec. 2.1.2), the Jacobian $|\mathrm{d}^3x_\mathrm{r}/\mathrm{d}^3x|$, and the ratio of average number densities $\bar{n}_\mathrm{g}(x_\mathrm{r})/\bar{n}_\mathrm{g,s}(x)$.

Starting with the latter, let us assume a flux-limited survey with a flux limit F_lim that targets a population of galaxies described by the luminosity function $\Theta(x, L)$. We define the average number density (or the selection function) as

$$\bar{n}_\mathrm{g,s}(x) = \int_{L_\mathrm{min}(x)}^{\infty} \Theta(x, L)\,\mathrm{d}L\,, \tag{2.72}$$

where $L_\mathrm{min}(x) = 4\pi D_\mathrm{L}^2(x)\,F_\mathrm{lim}$. The change in the average comoving density is solely caused by the shift in the coordinates. Consequently, we can derive by applying the coordinate map in Eq. (2.67) to the previous equation, we find (Paper II)

$$\begin{aligned}
\bar{n}_\mathrm{g}(x_\mathrm{r}) = \bar{n}_\mathrm{g,s}(x-\delta x) &= \int_{L_\mathrm{min}(x-\delta x)}^{\infty} \Theta(x-\delta x, L)\,\mathrm{d}L \\
&= \int_{L_\mathrm{min}(x)}^{\infty} \left[\Theta(x,L) - \delta x \frac{\partial\Theta(x,L)}{\partial x}\right]\mathrm{d}L \\
&\qquad + \frac{\mathrm{d}L_\mathrm{min}}{\mathrm{d}x}\delta x\left[\Theta(x,L_\mathrm{min}) + \delta x\frac{\partial\Theta(x,L_\mathrm{min})}{\partial x}\right] \\
&\approx \bar{n}_\mathrm{g}(x) - \frac{\delta x}{x}\frac{\mathrm{d}\,\bar{n}_\mathrm{g}(x)}{\mathrm{d}\ln x}\,.
\end{aligned} \tag{2.73}$$

The only missing component in Eq. (2.71) is the Jacobian $|\mathrm{d}^3x_\mathrm{r}/\mathrm{d}^3x|$. To simplify the notation used in the computation, we define the gradient operator ∇ w.r.t. x as well as its respective parallel and orthogonal components

$$\nabla_\parallel = \hat{\boldsymbol{x}}\cdot\boldsymbol{\nabla}\ , \qquad\qquad \boldsymbol{\nabla}_\perp = \boldsymbol{\nabla} - \hat{\boldsymbol{x}}(\hat{\boldsymbol{x}}\cdot\boldsymbol{\nabla}). \tag{2.74}$$

The Jacobian at linear order can then be computed as (Bertacca et al., 2014b)

$$\left|\frac{\mathrm{d}^3x_\mathrm{r}}{\mathrm{d}^3x}\right| = \left[1 - \nabla_\parallel\delta x_\parallel - \delta x_\parallel(\boldsymbol{\nabla}\cdot\hat{\boldsymbol{x}}) - \boldsymbol{\nabla}_\perp\delta\boldsymbol{x}_\perp\right]\,. \tag{2.75}$$

The first term on the RHS of the previous relation splits into two parts as follows

$$\begin{aligned}
\nabla_\parallel\delta x_\parallel =& \frac{\partial(\delta x_\parallel)}{\partial x} \\
=& (v_{\parallel,\mathrm{e}} - v_{\parallel,\mathrm{o}})\frac{\mathrm{d}z}{\mathrm{d}x}\frac{\partial(a\,H)^{-1}}{\partial z} + \frac{1}{a\,H}\frac{\partial v_{\parallel,\mathrm{e}}}{\partial x} \\
=& \left[1 - \frac{\mathrm{d}\ln H}{\mathrm{d}\ln(1+z)}\right]\frac{(v_{\parallel,\mathrm{e}} - v_{\parallel,\mathrm{o}})}{c} + \frac{1}{a\,H}\frac{\partial v_{\parallel,\mathrm{e}}}{\partial x}\,.
\end{aligned} \tag{2.76}$$

The second term can be directly computed to find

$$\delta x_{\parallel}(\boldsymbol{\nabla}\cdot\hat{\boldsymbol{x}}) = 2\frac{\delta x}{x} = 2\,\frac{(v_{\parallel,\mathrm{e}}-v_{\parallel,\mathrm{o}})}{x\,a\,H}\,. \tag{2.77}$$

The third term in the Jacobian identically vanishes as the shift is only along the line of sight (see Eq. 2.67).

Putting all three terms together, the Jacobian is then given by

$$\left|\frac{\mathrm{d}^3x_{\mathrm{r}}}{\mathrm{d}^3x}\right| = 1 - \frac{1}{a\,H}\frac{\partial v_{\parallel,\mathrm{e}}}{\partial x} - \frac{(v_{\parallel,\mathrm{e}}-v_{\parallel,\mathrm{o}})}{a\,H\,x}\left\{2+\frac{aHx}{c}\left[1-\frac{\mathrm{d}\ln H}{\mathrm{d}\ln(1+z)}\right]\right\}\,. \tag{2.78}$$

At this point, we are ready to derive the galaxy density contrast in redshift space. Putting the results of Eqs (2.78) and (2.73) in Eq. (2.71) , we find at first order

$$\begin{aligned}\delta_{\mathrm{g,s}}(x) = \delta_{\mathrm{g}}(x) &- \frac{1}{a\,H}\frac{\partial v_{\parallel,\mathrm{e}}}{\partial x} \\ &- \frac{(v_{\parallel,\mathrm{e}}-v_{\parallel,\mathrm{o}})}{a\,H\,x}\left\{\frac{\mathrm{d}\,\ln x^2\,\bar{n}_{\mathrm{g,s}}(x)}{\mathrm{d}\ln x}+\frac{Hx}{c(1+z)}\left[1-\frac{\mathrm{d}\ln H}{\mathrm{d}\ln(1+z)}\right]\right\}\,,\end{aligned} \tag{2.79}$$

which is usually expressed as

$$\boxed{\delta_{\mathrm{g,s}}(\boldsymbol{x}) = \delta_{\mathrm{g}}(\boldsymbol{x}) - \frac{1}{a\,H}\frac{\partial v_{\parallel,\mathrm{e}}}{\partial x} + \frac{(v_{\parallel,\mathrm{e}}-v_{\parallel,\mathrm{o}})}{a\,H\,x}\alpha_{\mathrm{c}}\,,} \tag{2.80}$$

where

$$\alpha_{\mathrm{c}}(x) = \left\{\frac{\mathrm{d}\,\ln x^2\,\bar{n}_{\mathrm{g,s}}(x)}{\mathrm{d}\ln x}+\frac{Hx}{c(1+z)}\left[1-\frac{\mathrm{d}\ln H}{\mathrm{d}\ln(1+z)}\right]\right\}\,. \tag{2.81}$$

We note that only the $\mathrm{d}\,\ln x^2\,\bar{n}_{\mathrm{g,s}}(x)/\mathrm{d}\ln x$ term is derived in the non-relativistic studies (Kaiser, 1987, Hamilton, 1998). This is mainly due to the shallowness of the era's surveys, consequently, $a\,H$ has been usually replaced with H_0. However, as shown in subsequent sections (see Chapter 7), this term significantly alters effects which are sensitive to the value of α_{c}.

2.2.2 Two-point Statistics in Redshift Space

2PCF

We next look at the impact of RSD on the two-point statistics introduced in Sec. 2.1.3. From Eq. (2.80), It is clear that the redshift space overdensity depends on the observer's position, which in turn breaks the translational invariance of the overdensity. Secondly, RSD mix the temporal and spatial coordinates, consequently the separation of Eq. (2.43) is no longer possible. Therefore, in general, the ensemble average of the product of the galaxy overdensity at different positions

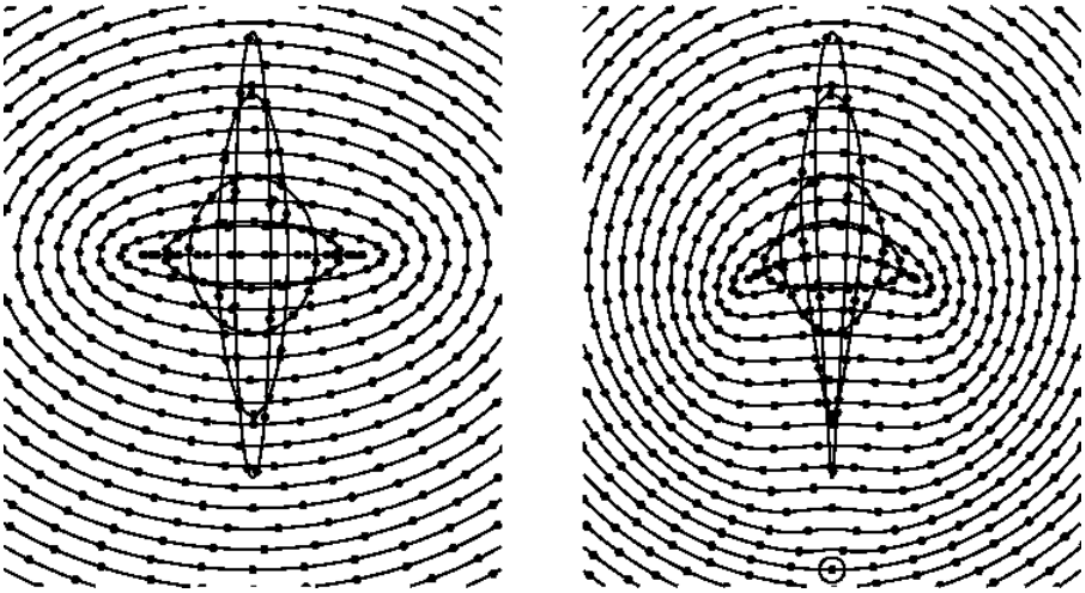
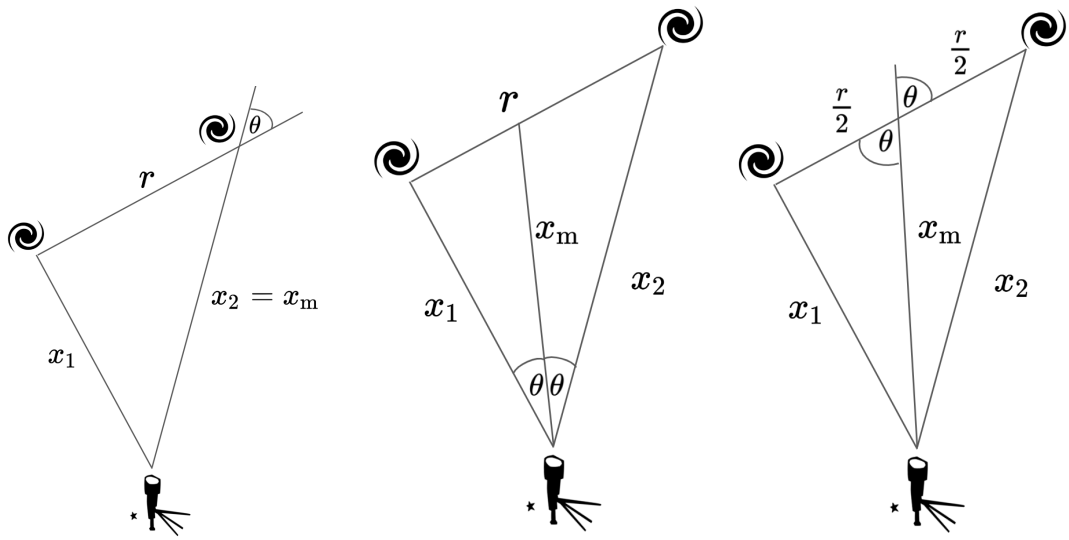


Figure 2.7: We show how a spherical overdensity in real-space is observed by a distant observer – where the GPP approximation can be applied– on the left and by a proximate observed (shown as a circle) on the right. The connected dots represent spherically symmetrical distributions in realspace. This figure is borrowed from (Hamilton, 1998).



(a) End-point. (b) Bisector. (c) Midpoint.

Figure 2.8: The possible parameterizations of the redshift space 2PCF function.

is no longer solely determined by r, i.e. it depends on both positions $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ (Szalay et al., 1998), and it should also be affected by unequal temporal correlation (e.g., Bertacca, 2020, Raccanelli & Vlah, 2023). In other words, the general form of the redshift-space 2PCF is given by ,

$$\langle \delta_{\rm g}(\boldsymbol{x}_1, z_1)\delta_{\rm g}(\boldsymbol{x}_2, z_2)\rangle = \xi_{\rm g}(\boldsymbol{x}_1, \boldsymbol{x}_2, z_1, z_2)\,. \tag{2.82}$$

To circumvent the complexity added by RSD, Kaiser (1987) proposed the Global Plane-Parallel (GPP) approximation. It posits that for a small volume, distant from the observer, we can assume all lines of sight are effectively parallel and replace the line of sight to all the objects by a constant line of sight, $\hat{\boldsymbol{x}}_{\rm c}$. As for the time dependence, it is set by the average redshift (denoted by z) of the volume considered. We can see how the observed galaxy distribution changes based on the location of the observer in Fig. 2.7. In the left panel, where the GPP approximation is applied, the spherically symmetric real-space distributions are seen as flattened circles (or pancakes) in redshift space. While for a proximate observer in redshift space, the redshift-space distributions are more boomerang (or croissant) shaped.

Assuming the GPP approximation, the 2PCF significantly simplifies to a dependence on the separation magnitude, the redshift and the angle between the separation $\boldsymbol{r}$ and the constant line of sight. Mathematically, it is given by (Hamilton, 1992, Hamilton & Culhane, 1996)

$$\xi_{\rm g}(r, \hat{\boldsymbol{r}}\cdot\hat{\boldsymbol{x}}_{\rm c}, z) = \sum_{l=0,2,4} \mathcal{F}_\ell(z)\, D_+^2(z)\, b^2(z)\, \zeta_\ell(r)\, \mathcal{L}_\ell(\hat{\boldsymbol{r}}\cdot\hat{\boldsymbol{x}}_{\rm c})\,, \tag{2.83}$$

where

$$\zeta_\ell(r) = \frac{i^\ell}{2\pi^2}\int k^2\, P(k, z=0) j_\ell(k\, r)\,, \tag{2.84}$$

and $\mathcal{L}_\ell$ are the Legendre polynomial.

Only the first three even multipoles of Eq. (2.83) survive in the GPP approximation

$$\mathcal{F}_0(z) = 1 + \frac{2}{3}\beta(z) + \frac{1}{5}\beta^2(z)\,, \tag{2.85}$$

$$\mathcal{F}_2(z) = \frac{4}{3}\beta(z) + \frac{4}{7}\beta^2(z)\,, \tag{2.86}$$

$$\mathcal{F}_4(z) = \frac{8}{35}\beta^2(z)\,, \tag{2.87}$$

where we have defined the linear RSD parameter as

$$\beta(z) = \frac{1}{b(z)}\left.\frac{{\rm d}\ln D_+}{{\rm d}\ln a}\right|_{a=(1+z)^{-1}}\,. \tag{2.88}$$

We show an example of the redshift-space GPP 2PCF multipoles , computed using

$$\xi_\ell(r,z) = (2\ell+1)\int \xi_{\rm g}(r,\hat{\boldsymbol{r}}\cdot\hat{\boldsymbol{x}}_{\rm c},z)\mathcal{L}_\ell(\hat{\boldsymbol{r}}\cdot\hat{\boldsymbol{x}}_{\rm c})\mathrm{d}\Omega_r\,, \tag{2.89}$$

in Fig. 2.9b, where Ω_r is the solid angle of the vector $\boldsymbol{r}$. It is clear that the monopole (red dashed line), ξ_0, is amplified w.r.t. both in correlation and anticorrelation with respect to the real-space 2PCF (black solid line). The quadropole, ξ_2, is negative definite due to the i^ℓ term in Eq. (2.84). As for the hexadecapole, ξ_4 exhibits a lower signal. Finally, we note that the BAO feature can be seen at the same scale $\approx 100\,h^{-1}$ Mpc for all three multipoles (although very weakly for ξ_4).

Going beyond the GPP approximation requires the so-called wide-angle formalism that accounts for the large angular separations between pairs (the right panel in Fig. 2.7). Szalay et al. (1998) finds that the 2PCF can be then defined by three parameters that determine the shape of the triangle formed by the two points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ and the observer. There are three widely-used conventions in the literature for the parametrisation of the aforementioned triangle. Firstly, the end-point parametrization (EP), shown in the left panel in Fig. 2.8, defines the line of sight as the direction to one of the pair. However, that parametrisation produces spurious odd multipoles due to its inherent asymmetry (Reimberg et al., 2016). Secondly, the Bisector parametrization (BP), shown in the middle panel, takes the angle bisector as the line of sight. The BP is the most employed convention in studies of wide-angle effects (e.g., Matsubara, 2000a, Szapudi, 2004, Pápai & Szapudi, 2008, Raccanelli et al., 2010, Bertacca et al., 2012, Samushia et al., 2012, Bertacca, 2015, Raccanelli et al., 2018) as well as relativistic RSD corrections (for eg. Bertacca et al., 2012, Raccanelli et al., 2016a, Tansella et al., 2018, Bertacca, 2020, Jelic-Cizmek et al., 2021). Finally, the mid-point parametrization (MP), shown in the right panel, employs the midpoint of the separation between the pair. The MP has been historically used in estimating the 2PCF (Fisher et al., 1994, Moore et al., 2001, Keihänen et al., 2019, Paviot et al., 2022). For a review of the impact of changing the adopted convention on the power spectrum multipoles, we recommend referring to Reimberg et al. (2016).

We adopt the MP in the analysis presented in this work as it is the standard convention of the *Euclid* collaboration which is the target survey we use for our results in Paper III. In the MP, we use the midpoint coordinate system to estimate $\xi(\boldsymbol{x}_1,\boldsymbol{x}_2)$ and, in this case, we define the separation $\boldsymbol{r}$ and midpoint $\boldsymbol{x}_{\rm m}$ vectors in the following way

$$\boldsymbol{r} = \boldsymbol{x_2} - \boldsymbol{x_1}\ \ ,\ \ \boldsymbol{x}_{\rm m} = \frac{\boldsymbol{x_2}+\boldsymbol{x_1}}{2}\,. \tag{2.90}$$

Instead the pair orientation angle θ (i.e. the angle between the $\boldsymbol{r}$ linear separation and the midpoint $\boldsymbol{x}_{\rm m}$) is given by

$$\mu = \cos\theta = \hat{\boldsymbol{r}}\cdot\hat{\boldsymbol{x}}_{\boldsymbol{m}}\,. \tag{2.91}$$

With that choice of coordinate system, it is natural to redefine the 2PCF function

as

$$\xi(r, x_{\rm m}, \mu) = \langle \delta_{\rm g}(\boldsymbol{x}_{\rm m} + \boldsymbol{r}/2)\delta_{\rm g}(\boldsymbol{x}_{\rm m} - \boldsymbol{r}/2)\rangle \ . \tag{2.92}$$

Power Spectrum

We next look at the impact of RSD on the power spectrum. Similar to the 2PCF, the introduction of RSD complicates the treatment of the power spectrum. In general, the ensemble average of the product of two Fourier modes of the galaxy overdensity is no longer diagonal, i.e. it depends on both wavenumbers $\boldsymbol{k}$ and $\boldsymbol{k'}$ (Zaroubi & Hoffman, 1996, Szalay et al., 1998), and it should also be affected by unequal temporal correlation. In other words, the general form of the Fourier space autocorrelation function is given by

$$\left\langle \tilde{\delta}_{\rm g}(\boldsymbol{k}, z)\, \tilde{\delta}_{\rm g}(\boldsymbol{k}', z') \right\rangle \equiv (2\,\pi)^3\, C(\boldsymbol{k}, \boldsymbol{k}', z,\, z')\,. \tag{2.93}$$

We can simplify this relation by assuming the GPP approximation. In that case, the Fourier space power spectrum of galaxies in redshift space becomes a simple analytical relation, and the map between the multipoles and the power spectrum in realspace is given by (similar to Eq. (2.83))

$$P_{\rm g,s}(\boldsymbol{k}, z) = \sum_{l=0,2,4} \mathcal{F}_\ell(z) D_+^2(z)\, b^2(z)\, P(k, z=0) \mathcal{L}_\ell(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{x}}_{\rm c})\,, \tag{2.94}$$

where the coefficients $\mathcal{F}_\ell$ are given in Eqs. (2.85 - 2.87). The previous result is usually called the Kaiser prediction (Kaiser, 1987).

Instead, for a thicker redshift interval, the evolution in redshift can be obtained by averaging over the entire redshift interval. In that case, the power spectrum can be rewritten in the following way (see Yamamoto et al., 1999, Pryer et al., 2022)

$$\boxed{P_{\rm g,s}(\boldsymbol{k}; z_i, z_f) = \sum_{l=0,2,4} F_{{\rm eff},\ell}(z_i, z_f) P(k, z=0) \mathcal{L}_\ell(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{x}}_{\rm c})\,,} \tag{2.95}$$

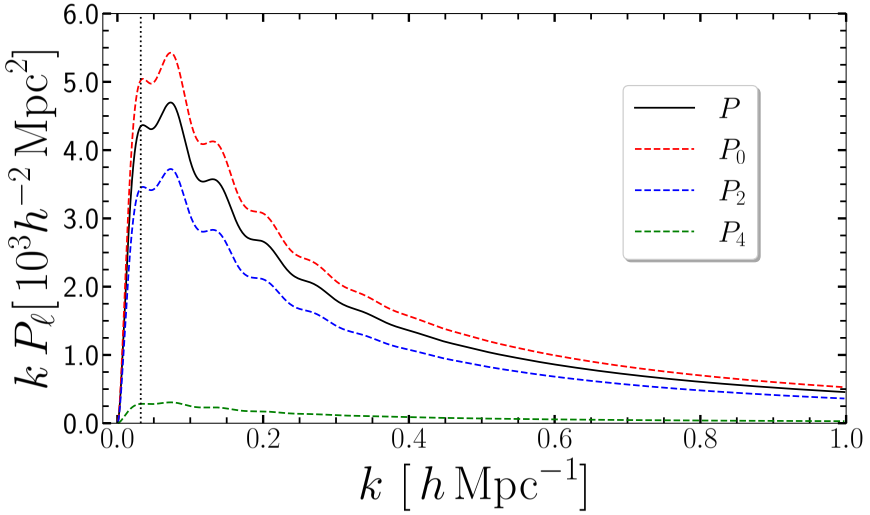
where

$$F_{{\rm eff},\ell}(z_i, z_f) = \frac{\int_{z_i}^{z_f} \mathcal{F}_\ell(z)\, b^2(z)\, D_+^2(z)\, \bar{n}_{\rm g}^2(z) (\mathrm{d}V_{\mathcal{S}}/\mathrm{d}z)\, \mathrm{d}z}{\int_{z_i}^{z_f} \bar{n}_{\rm g}^2(z) (\mathrm{d}V_{\mathcal{S}}/\mathrm{d}z)\, \mathrm{d}z}\,. \tag{2.96}$$

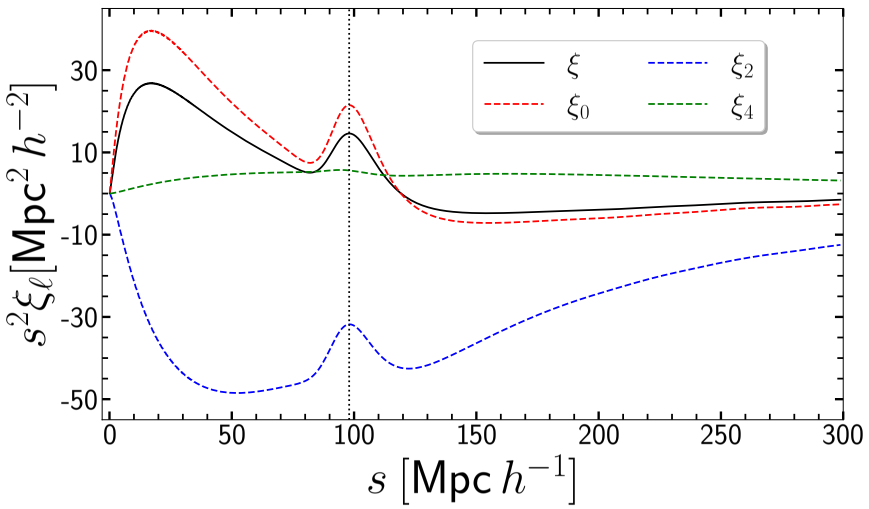
Where $\mathrm{d}V_{\rm S}$ is the comoving volume element and the limits of the integral correspond to the boundary redshifts of the small volume considered.

We show an example of the power spectrum multipoles, $P_\ell(k; z_i, z_f)$,[5] which we

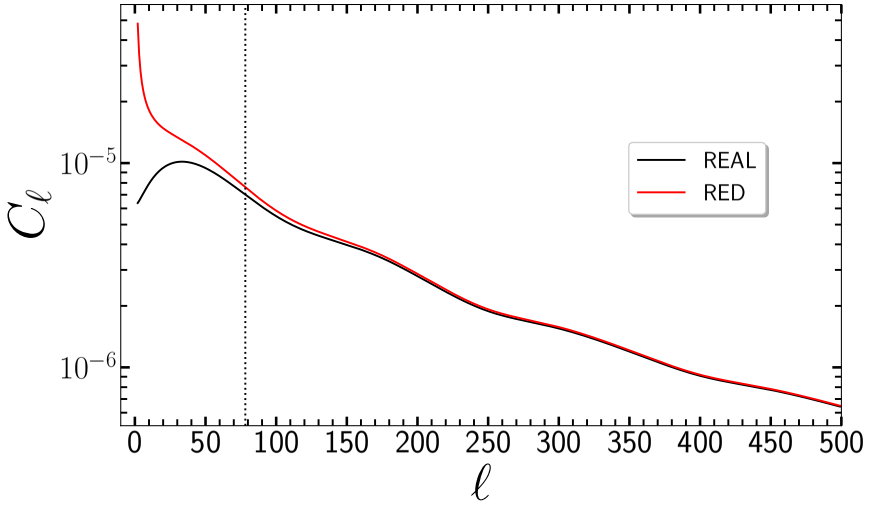
[5] We drop the g and s subscripts in the multipoles as the ℓ subscript is always used to describe the redshift-space galaxy power spectrum multipoles in this work.



(a) Example of the galaxy power spectrum in redshift space using the GPP approximation.



(b) Example of the galaxy 2PCF in redshift space using the GPP approximation.



(c) Example of the galaxy angular power spectrum in redshift-space.

Figure 2.9: Similar to Fig. 2.4 but we add the redshift space two-point statistics.

compute via

$$P_\ell(k; z_i, z_f) = (2\ell + 1) \int P_{\rm g,s}(\boldsymbol{k}, z_i, z_f) \mathcal{L}_\ell(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{x}}_{\rm c}) {\rm d}\Omega_k \,, \tag{2.97}$$

in Fig. 2.9a. It is clear that most of the power in redshift space in concentrated in the monopole shown as the dashed red line. The hexadecapole (dashed green line) has very little signal while the quadropole (dashed blue line) is in between. The k-dependence of the multipoles matches the real-space power spectrum (shown as a solid black line) as the GPP Kaiser prediction only boosts the amplitude.

Going beyond the model presented in Eq. (2.95) causes two complications: *i*) we lose the translational invariance symmetry (Zaroubi & Hoffman, 1996, Reimberg et al., 2016) and *ii*) the power spectrum of galaxies must include the impact of wide-angle spatial distortions, local and integrated relativistic corrections along the line of sight (see Sec. 2.3). To circumvent the first issue, Scoccimarro (2015) defines a 'local' (or mixed space) power spectrum, $P_{\rm loc}$, which is obtained by Fourier transforming the 2-point correlation function, defined in Eq. (2.43), at separation $\boldsymbol{r}$

$$P_{\rm loc}(\boldsymbol{x}_{\rm m}, \boldsymbol{k}) \equiv \int \langle \delta_{\rm g}(\boldsymbol{x}_{\rm m} + \boldsymbol{r}/2)\, \delta_{\rm g}(\boldsymbol{x}_{\rm m} - \boldsymbol{r}/2) \rangle \, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \, {\rm d}^3 r \,, \tag{2.98}$$

$$= \int \xi(r, x_{\rm m}, \mu) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \, {\rm d}^3 r \tag{2.99}$$

where $P_{\rm loc}(\boldsymbol{x}_{\rm m}, \boldsymbol{k})$ is the Fourier conjugate of the local contribution to the 2PCF function at the (midpoint) position $\boldsymbol{x}_{\rm m}$. The multipoles of the power spectrum can then be recovered through

$$P_\ell(k) = (2\ell + 1) \iint P_{\rm loc}(\boldsymbol{x}_{\rm m}, \boldsymbol{k})\, \mathcal{L}_\ell(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{x}}_{\rm m}) \frac{{\rm d}\Omega_k}{4\pi} \frac{{\rm d}^3 x_{\rm m}}{V_\mathcal{S}} \,, \tag{2.100}$$

where $V_\mathcal{S}$ is the volume of the considered region.

In estimating the power spectrum multipoles in actual surveys, the volume integral in Eq. (2.100) is limited to the shell of the survey in which we are computing the spectrum $P_\ell(k)$. We devote Chapter 5 to the study of the implications of limited survey volume and the GPP approximation.

As for the inclusion of wide-angle and relativistic RSD (discussed in Sec. 2.3), Castorina & White (2018) proposed to expand the multipoles of the power spectrum in the perturbative parameter $(k\, x)^{-1}$ and calculate corrections to the GPP prediction. In this approximate prescription, relativistic corrections can then be added to this expansion perturbatively (Beutler et al., 2019, Castorina & Di Dio, 2022, Noorikuhani & Scoccimarro, 2023). Note that this approach might fail when we consider wide-angle pairs of galaxies. Another possible approach which solves the above issues without approximations is to use the "spherical-Fourier Bessel" (SFB) formalism which was first computed by (Peebles, 1973), then in redshift space by (Heavens & Taylor, 1995), and applied to real data in (Percival et al., 2004). Fi-

nally, it was extended to include GR effects in (Yoo & Desjacques, 2013, Bertacca et al., 2018). In this work, we use the 'local' power spectrum approach defined in Eq. (2.98) as adopted by the *Euclid* collaboration as our aim to measure the importance of relativistic RSD in the standard statistics of galaxy clustering. It is important to note that this approach has been extremely successful in extracting parameters from smaller scales than the ones discussed in this work, i.e., linear and quasi-linear scales (For eg. Beutler et al., 2014) which is the main reason of the *Euclid* collaboration adoption (Amendola et al., 2018) of the 'local' power spectrum approach.

The Angular power spectrum

Finally, for the angular power spectrum, the impact of RSD distortions can be incorporated using a version of Eq. (2.56), i.e.,

$$C_\ell = \frac{2}{\pi}\int \mathcal{W}^2_{\mathrm{Tot},\ell}(k)\, k^2\, P(k, z=0)\, \mathrm{d}k\,. \tag{2.101}$$

Redefining however the weight functions $\mathcal{W}^{\mathrm{Tot}}$. The new weight function is given by (e.g., Fisher et al., 1994, Scharf et al., 1992, Padmanabhan et al., 2007, Loureiro et al., 2019)

$$\mathcal{W}_{\mathrm{Tot},\ell}(k) = \mathcal{W}_\ell(k) + \mathcal{W}_{\mathrm{RSD},\ell}(k)\,. \tag{2.102}$$

The first term on the RHS of the previous equation is the same one used for the real-space clustering (see Eq. 2.57). The second term incorporates the velocity shift and is given by

$$\begin{aligned}\mathcal{W}_{\mathrm{RSD},\ell}(k) &= \frac{\beta}{k}\int \frac{\mathrm{d}\mathcal{P}(x)}{\mathrm{d}x}\frac{\mathrm{d}j_\ell(k\,x)}{\mathrm{d}x}\,\mathrm{d}x \\ &= \beta\int \mathcal{P}(x)\left[\frac{(2\ell^2+2\ell+1)}{(2\ell+3)\,(2\ell-1)}j_\ell(k\,x) + \frac{\ell\,(\ell-1)}{(2\ell-1)\,(2\ell+1)}j_{\ell-2}(k\,x)\right. \\ &\qquad \left. + \frac{(\ell+1)\,(\ell+1)}{(2\ell+1)\,(2\ell+3)}j_{\ell+2}(k\,x)\right]\mathrm{d}x\,,\end{aligned} \tag{2.103}$$

where the spherical Bessel recurrence relation is used in the second equality.

We plot the Redshift-space angular power spectrum in Fig. 2.9c. As seen in the Figure, the peculiar velocity RSD strongly boosts the angular power spectrum for $\ell < 100$.

As for the inclusion of the relativistic RSD discussed in the following section, they can be incorporated by adjusting the weight function in Eq. (2.102) (Bonvin & Durrer, 2011, Challinor & Lewis, 2011). Subsequently, they have been integrated into two widely-used Boltzmann solvers, namely the `CAMB` (Lewis & Bridle, 2002) and `CLASS` codes (Blas et al., 2011, Di Dio et al., 2013).

2.3 Relativistic Redshift Space Distortions

The prescription described in Sec. 2.2 only accounts for the change in redshift due to the peculiar velocity. However, the light produced by distant galaxies is distorted by the intervening structure as it travels from the source galaxy to the observer Yoo (2010). In order to account correctly for all these effects, a relativistic prescription is required. These additional distortions have been studied up to 1st order in (Bonvin & Durrer, 2011, Challinor & Lewis, 2011, Jeong et al., 2012) and up to 2nd order in (Bertacca et al., 2014a,b).

In this section, we write an overview of these effects. We adopt the approach outlined in Bertacca (2015). Similar to the classical analysis presented in Sec. 2.2, we aim to relate the observed overdensity in redshift space to the real space one, i.e. produce a relativistic version of Eq. (2.80). The starting point of the derivation remains the same: the conservation of the number of galaxies. Let us assume that there is a population of galaxies with physical comoving coordinates $x^\mu_{\rm r}$, fluxes $F_{\rm r}$ and velocity field $u^\mu_{\rm r}(x^\mu)$ that exist in a space-time defined by the comoving metric $g^{\mu\nu}$ with determinant g. The physical number of galaxies of that population in a volume $\mathcal{V}$ is given by

$$\mathcal{N} = \int_{\mathcal{V}} a^3_{\rm r}(x^0_{\rm r}) n_{\rm g}(x^\mu_{\rm r}, F_{\rm r}) {\rm d}\mathcal{V}_{\rm r}\,, \tag{2.104}$$

where $n_{\rm g}$ is the physical number density and $a_{\rm r}$ is the real-space scale factor. We equate the previous equation to the number of galaxies observed in redshift space assuming a flux-limited survey where only objects with an observed flux $F_{\rm r}$ larger than the survey limit $F_{\rm lim}$ are counted. The number of galaxies in redshift space is by construction

$$\mathcal{N} = \int_{\mathcal{V}} a^3(x^0)\, n_{\rm g}(x^\mu, F > F_{\rm lim})\, {\rm d}\mathcal{V}\,, \tag{2.105}$$

where x^μ and $a = (1+z_{\rm obs})^{-1}$ are the redshift space (or observed) coordinates, and scale factor respectively. To correctly relate the redshift and real space overdensity, we follow the following steps.

i) Derive the coordinate map that relates the physical (real-space) position $x^\mu_{\rm r}$ to the redshift position x^μ, i.e. the relativistic analogue of Eq. (2.67).

ii) Calculate the change in the volume element, i.e. relate ${\rm d}\mathcal{V}_{\rm r}$ to ${\rm d}\mathcal{V}$ (the classical volume change is given in Eq. 2.78).

iii) Evaluate the change in the number density due to the coordinate map as well as the change in the flux (the relativistic version of Eq. 2.73).

iv) Finally, put the results together to produce the redshift space overdensity.

We dedicate a subsection to each step. It's worth noting that we assume no velocity bias; in other words, we presume that galaxy velocities follow the matter

velocity field. However, before delving into the transformation from real to redshift space, we must establish a connection between the cosmological functions discussed in Sec. 2.1.2 and their relativistic counterparts. This is crucial for accurately defining the galaxy overdensity $\delta_{\rm g}$ in redshift space.

2.3.1 Interpreting Newtonian Simulations In General Relativity

In Sec. 2.1.2, we provided a linear order Newtonian description of structure formation. To accurately compute the impact of Relativistic (RSD) using a relativistic treatment, it is essential to elucidate how the fluid equations (see to Eqs. 2.30–2.32) can be interpreted in a relativistic framework for a spatially flat universe. In other words, we demonstrate how, at linear order, the Newtonian potential φ, matter overdensity $\delta_{\rm N}$, and [6] peculiar velocity field $\boldsymbol{v}$ can be mapped to general relativistic degrees of freedom defined by the metric in a specific gauge.

Synchronous-Comoving Gauge

Translating the Newtonian quantities to relativistic functions depends on the gauge chosen for defining the metric. In this section, we adopt the synchronous-comoving gauge, following the prescription outlined in (Bruni et al., 2014, also see Matarrese et al., 1998, Bartolo et al., 2010). Similar treatments using the conformal Newtonian gauge, also known as the Poisson restricted gauge, are discussed in Chisari & Zaldarriaga (2011), while the longitudinal gauge is explored in Green & Wald (2012). In the treatment presented in this subsection, we adopt natural units. The synchronous-comoving gauge can be written as follows

$$\mathrm{d}s^2 = a^2(\tau)\left[-\mathrm{d}\tau^2 \gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j\right], \tag{2.106}$$

where τ is the conformal time and γ_{ij} is the spatial metric.

We assume irrotational dust flow with fundamental observers whose four-velocity can be defined as $u_\mu = [-a, 0, 0, 0]$. This choice enables us to align the geodesics of the matter with the worldlines of our "fundamental" observers while ensuring that cosmic time coincides with the proper time of the fluid elements. In this sense, the reference frame is Lagrangian (Villa et al., 2011). However, we note that the synchronous and comoving conditions can only be simultaneously applied for irrotational dust.

We define the deformation tensor as

$$\mathcal{D}^\mu_\nu := a u^\mu_{;\nu} - \mathcal{H}\delta^{\mathrm{K}\ \mu}_\nu\,, \tag{2.107}$$

where the semicolon denotes the covariant derivative, $\mathcal{H} = a'/a$ and the prime denotes the conformal time derivative. The trace $\mathcal{D}$ and traceless components $\mathcal{D}^{\mathrm{TS}\mu}_\nu$

[6] we denote the Newtonian overdensity with the subscript N.

of the deformation tensor denote the inhomogeneous and anisotropic parts of the volume expansion, respectively. The deformation matrix in our case is purely spatial and can be related to the spatial metric using

$$\mathcal{D}^i_j = \frac{1}{2}\gamma^{ik}\gamma'_{kj}\,. \tag{2.108}$$

Using the relativistic energy conservation equation, a relativistic continuity equation can be written as

$$\delta'_{\rm SC} + (1+\delta_{\rm SC})\mathcal{D} = 0\,, \tag{2.109}$$

where we use the subscript SC to differentiate the relativistic matter density contrast in the synchronous-comoving gauge from its Newtonian counterpart. Subsequently, by combining the pure temporal element of the Einstein field equations (see Eq. 2.3) with the trace of the spatial components, an evolution equation for the trace of the deformation tensor can be derived

$$\mathcal{D}^2 + \mathcal{H}\mathcal{D} + \mathcal{D}^i_j\mathcal{D}^j_i + \frac{3}{2}\mathcal{H}^2\Omega_m\delta_{\rm N} = 0\,. \tag{2.110}$$

The only missing ingredient in order to match the relativistic and Newtonian prescription is an analogue to the Poisson equation. To that end, we linearize the spatial metric as follows

$$\gamma_{ij} = (1-2\zeta)\delta^{\rm K}_{ij} + \left(\partial_i\partial_j - \delta^{\rm K}_{ij}\nabla^2\right)\chi\,, \tag{2.111}$$

where ζ and χ are scalar potentials[7]. The scalar potentials can be combined to form the gauge invariant Bardeen potentials

$$\Phi_{\rm I.B.} = -\frac{1}{2}\left(\chi'' + \mathcal{H}\chi'\right)\,, \tag{2.112}$$

$$\Psi_{\rm I.B.} = \zeta + \frac{1}{6}\nabla^2\chi + \frac{1}{2}\mathcal{H}\chi\,. \tag{2.113}$$

The Bardeen potential $\Phi_{\rm I.B.}$ can be computed using a Poisson-like equation sourced by the matter overdensity, i.e. (Bardeen, 1980)

$$\nabla^2\Phi_{\rm I.B.} = \frac{3}{2}\mathcal{H}^2\Omega_m\delta_{\rm SC}\,. \tag{2.114}$$

The relativistic system of equations is given by Eqs. (2.109), (2.110) and (2.114). To relate that system of equations to the Newtonian fluid equations presented in Sec. 2.1.2, we rewrite the Newtonian equations in terms of the conformal time

[7] We ignore vector and tensor degrees of freedom as they are linearly independent at first order in cosmological perturbations.

($\tau = t/a$) and combine the momentum and Poisson equations to find

$$\frac{\partial \delta_{\rm N}}{\partial \tau} + \nabla \cdot [(1+\delta_{\rm N})\,\boldsymbol{v}] = 0\,, \tag{2.115}$$

$$\frac{\partial \nabla \cdot \boldsymbol{v}}{\partial \tau} + \mathcal{H}\,\nabla \cdot \boldsymbol{v} + \nabla \cdot (v \cdot \nabla)\,\boldsymbol{v} + \frac{3}{2}\mathcal{H}^2 \Omega_m \delta_{\rm N} = 0\,, \tag{2.116}$$

$$\nabla^2 \varphi = \frac{3}{2}\mathcal{H}^2 \Omega_m \delta_{\rm N}\,. \tag{2.117}$$

In order to match the relativistic equations, we then rewrite the Newtonian fluid equations in Lagrangian coordinates using

$$\frac{\mathrm{d}}{\mathrm{d}\tau} = \frac{\partial}{\partial \tau} + \boldsymbol{v} \cdot \nabla\,. \tag{2.118}$$

By defining $\mathcal{D}^i_{Nj} = \partial^i v_j = \partial^i \partial_j \phi_v$, where ϕ_v is a scalar potential (see Sec. 2.1.2), the Newtonian equations can then be recast as

$$\frac{\mathrm{d}\delta_{\rm N}}{\mathrm{d}\tau} + (z+\delta_{\rm N})\mathcal{D}_{\rm N} = 0\,, \tag{2.119}$$

$$\frac{\mathrm{d}\mathcal{D}_{\rm N}}{\mathrm{d}\tau} + \mathcal{H}\mathcal{D}_{\rm N} + \mathcal{D}^i_{{\rm N}\,j}\mathcal{D}^j_{{\rm N}\,i} + \frac{3}{2}\mathcal{H}^2\Omega_m\delta_{\rm N} = 0\,, \tag{2.120}$$

$$\nabla^2\varphi = \frac{3}{2}\mathcal{H}^2\Omega_m\delta_{\rm N}\,, \tag{2.121}$$

where the Poisson equation is unchanged.

Comparing Eqs. (2.109), (2.110) and (2.114) to the Newtonian case shown in Eqs. (2.119), (2.120) and (2.121), it is clear that the dictionary that translate the Newtonian case to the relativistic linear order solution in the synchronous-comoving gauge is given by

$$\begin{aligned} \frac{1}{\mathrm{d}\tau} &\to \frac{\partial}{\partial\tau} \quad , \quad \partial^i v_j \to \mathcal{D}^i_j \\ \delta_{\rm N} &\to \delta_{\rm SC} \quad , \quad \varphi \to \Phi_{\rm I.B.}\,. \end{aligned} \tag{2.122}$$

The main feature of the synchronous-comoving gauge is that it allows for a direct (lagrangian) definition of the the local bias through a peak background split approach (Jeong et al., 2012, Yoo, 2023), i.e., $\delta_{\rm g} = b\delta_{\rm SC}$. However, there are two difficulties when relating to numerical simulations, firstly the potential computed using the relativistic Poisson equation (see Eq. 2.114) does not correspond directly to the perturbative metric scalar degrees of freedom. Secondly, the simulation peculiar velocity field is not directly matched to the relativistic peculiar velocity as the latter is vanishing by construction in the synchronous-comoving gauge. The scalar-restricted Poisson gauge, on the other hand, given by

$$\mathrm{d}s^2 = a^2(\tau)\left[(1+\psi)\,\mathrm{d}\tau^2 + (1+\phi)\delta^K_{ij}\mathrm{d}x^i\mathrm{d}x^j\right]\,, \tag{2.123}$$

simultaneously solves these two issues. As the metric perturbation ψ coincides

exactly with the Bardeen potentials, i.e.

$$\psi = \Phi_{\text{I.B.}} \tag{2.124}$$
$$\phi = \Psi_{\text{I.B.}}\,, \tag{2.125}$$

and the peculiar velocity at first order matches directly with the Newtonian peculiar velocity (assuming no velocity bias, Chisari & Zaldarriaga, 2011). As for accounting for the linear bias, it can be properly accounted at linear order above by a gauge transformation (Challinor & Lewis, 2011, Jeong et al., 2012)

$$\delta_{\text{g,P}} = b\delta_{\text{SN}} + (3 - \mathcal{E})\,\mathcal{H}\Phi_{\text{v}}\,. \tag{2.126}$$

2.3.2 Coordinate Transformation

General linear order Map

Let us define the following notation that will simplify the expressions in this section, for a vector $\boldsymbol{A}$, we define its parallel and orthogonal components as

$$\boldsymbol{n}\cdot\boldsymbol{A} = n_i\,A^i = A_{\parallel} \quad , \quad A^i_{\perp} = A^i - n^i A_{\parallel}\,. \tag{2.127}$$

Similarly for the gradient operation

$$\frac{\partial}{\partial x^i} = \partial_i \qquad , \qquad \boldsymbol{n}\cdot\boldsymbol{\nabla} = n^i\,\partial_i = \partial_{\parallel}\,, \tag{2.128}$$
$$\partial_{\perp,i} = \partial_i - n^i\partial_{\parallel} \quad , \quad \boldsymbol{\nabla}^2_{\perp} = \partial_{\perp,i}\partial^i_{\perp} = \delta^{\text{K}\,ij}\partial_i\,\partial_j - \partial^2_{\parallel} - \frac{2}{x}\partial_{\parallel}\,. \tag{2.129}$$

In order to outline the coordinate map between the redshift space and real space, we adopt the "cosmic rulers" prescription developed in Schmidt & Jeong (2012). In that prescription, the null geodesic of a photon x^μ from a distant light source in observed (redshift) space is, by construction, described by the following coordinates (assuming a spatially flat universe) [8]

$$x^\mu = (c\,\eta_0 - x,\; x\,\boldsymbol{n})\,. \tag{2.130}$$

Where $\eta_0 = \eta + x/c$ is the present-day value of the conformal time η. x is computed using the background universe, i.e. through Eq. (2.58) using the observed redshift z. Finally $\boldsymbol{n} \equiv \boldsymbol{x}/x$ is the observed direction. The collection of these coordinates define the observed (redshift) frame. It is clear from Eq. (2.130) that the inner products of these null vectors is measured through the conformal version of the FLRW metric (see Eq. 2.5)[9]

$$\mathrm{d}s^2 = a^2(\eta)\left[-c^2\mathrm{d}\eta^2 + \delta^{\text{K}}_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j\right]\,. \tag{2.131}$$

[8] As the equations derived in this section are implemented in the code, we use physical units, where the speed of light is given by c.

[9] Which can be derived by defining the conformal time $\eta = a\,t$.

The affine parameter adopted for all calculations is the redshift space comoving distance x, consequently, the total derivative of any quantity is given by

$$\frac{\mathrm{d}}{\mathrm{d}x} = \frac{-1}{c}\frac{\partial}{\partial\eta} + \boldsymbol{n}\cdot\boldsymbol{\nabla} = \frac{-1}{c}\frac{\partial}{\partial\eta} + \partial_{\parallel}\,. \tag{2.132}$$

On the other hand, the real-space coordinates are defined by the physical coordinates x^{μ}_{r}. The cosmic ruler approach relates the physical coordinates to their redshift space counterparts perturbatively at the physical comoving distance x_{r} through

$$x^{\mu}_{\mathrm{r}}(x_{\mathrm{r}}) = x^{\mu}(x) + \Delta x^{\mu}(x)\,. \tag{2.133}$$

In principle, the coordinate shift can be generally derived using the cosmic ruler prescription without defining the perturbed metric (see Schmidt & Jeong, 2012, Bertacca et al., 2014a,b, Bertacca, 2015). However, as the `LIGER` method and implemented code adopt the conformal Newtonian gauge (see Sec. 2.3.1), we opt to provide all results directly in that gauge. We are then able to directly write the perturbed space-time metric for the physical (real) space coordinates as, neglecting vector and tensor perturbations,

$$\mathrm{d}s^2 = a^2_{\mathrm{r}}(\eta)\left[-c^2(1\,+\,2\psi)\mathrm{d}\eta^2 + (1\,-\,2\phi)\delta^{\mathrm{K}}_{ij}\,\mathrm{d}x^i_{\mathrm{r}}\,\mathrm{d}x^j_{\mathrm{r}}\right]\,, \tag{2.134}$$

where ψ and ϕ denote the dimensionless Bardeen potentials.

In order to derive the transformation $\Delta x^{\mu}(x)$, we write $x_{\mathrm{r}} = x + \delta x$ and evaluate the change in the coordinates in two distinct steps. First relating the redshift and real space position at the realspace comoving distance x_{r}

$$x^{\mu}_{\mathrm{r}}(x_{\mathrm{r}}) = x^{\mu}(x_{\mathrm{r}}) + \delta x^{\mu}(x_{\mathrm{r}})\,. \tag{2.135}$$

then we perturb the comoving position

$$x^{\mu}_{\mathrm{r}}(x_{\mathrm{r}}) = x^{\mu}(x) + \frac{\mathrm{d}x^{\mu}}{\mathrm{d}x}\delta x + \delta x^{\mu}(x)\,, \tag{2.136}$$

where the $(\mathrm{d}\delta x^{\mu}/\mathrm{d}x)\delta x$ is ignored at first order. Combining Eqs. (2.136) & (2.133), we find the form of the coordinate map

$$\Delta x^{\mu}(x) = \frac{\mathrm{d}x^{\mu}}{\mathrm{d}x}\delta x + \delta x^{\mu}(x)\,. \tag{2.137}$$

We note that the classical analysis presented in Sec. 2.2.1 only accounts for the δx term in the previous equations, assuming that the redshift and realspace positions overlap. Consequently, it ignores the change in the two coordinate systems due to the structure of the universe.

From Eq. (2.132), it is immediately clear that the geodesic vector $k^{\mu} = \mathrm{d}x^{\mu}/\mathrm{d}x =$

$(-1, \boldsymbol{n})$, consequently the shifts of Eq. (2.137) are given by

$$\Delta x^0(x) = -\delta x + \delta x^0(x)\,, \tag{2.138}$$
$$\Delta x^i(x) = n^i \delta x + \delta x^i(x)\,. \tag{2.139}$$

The real-space geodesic vector $k^\mu_{\rm r}$ on the other hand can be derived using Eq. (2.135)

$$\begin{aligned} k^\mu_{\rm r} \equiv \frac{{\rm d}x^\mu_{\rm r}}{{\rm d}x} &= \frac{\rm d}{{\rm d}x}(x^\mu + \delta x^\mu) \\ &= \left(-1 + \delta f, n^i + \delta n^i\right)\,, \end{aligned} \tag{2.140}$$

where we defined

$$\frac{{\rm d}\delta x^0}{{\rm d}x} \equiv \delta f \;\text{ and }\; \frac{{\rm d}\delta x^i}{{\rm d}x} \equiv \delta n^i\,. \tag{2.141}$$

Redshift and scale factor

We can derive the temporal coordinate map of Eq. (2.138) by understanding how the scale factors can be related between real and redshift space. To that end, we outline how the scale factor $a_{\rm r}$ is perturbed. The scale factor $a_{\rm r}$ in physical space is given by

$$\begin{aligned} a_{\rm r}(x^0_{\rm r}) = a(x^0 + \Delta x^0) &= a(x^0) + \frac{\partial a}{\partial x^0}\Delta x^0 \\ &= a(x^0)\left[1 + \frac{1}{a(x^0)}\frac{\partial a}{\partial x^0}\Delta x^0\right] \\ &= a(x^0)\left[1 + \mathcal{H}\Delta x^0\right]\,, \end{aligned} \tag{2.142}$$

where we defined

$$\mathcal{H} = \frac{1}{a(x^0)}\frac{\partial a}{\partial x^0} = a\,H\,. \tag{2.143}$$

The fractional change in the scale factor is thus given by

$$\delta \ln a = \frac{a_{\rm r}}{a} - 1 = \mathcal{H}\Delta x^0 \tag{2.144}$$

In order to explicitly calculate $\delta \ln a$ (and in turn Δx^0), we use the standard definition of the observed redshift (Schmidt & Jeong, 2012, Jeong et al., 2012)

$$(1 + z_{\rm obs}) := (1 + z) = \frac{(g_{\mu\alpha} u^\mu_{\rm r} p^\alpha_{\rm r})|_{\rm e}}{(g_{\mu\alpha} u^\mu_{\rm r} p^\alpha_{\rm r})|_{\rm o}} \tag{2.145}$$

where the subscript e and o denote evaluating the inner product at emission (at $x_{\rm r}$) and observation (a_0=0), respectively. $p^\mu_{\rm r}$ is the four-momentum of the photon given by

$$p^\mu_{\rm r} = \frac{f(a_{\rm r})}{a_{\rm r}} k^\mu_{\rm r}\,, \tag{2.146}$$

where $f(a)$ is the frequency of the photon.

In order to evaluate the observed redshift, we need to derive the four-velocity vector $u^\mu_{\rm r}$. The velocity field in the perturbed metric is defined as the derivative with respect to the proper time, i.e. $u^\mu_{\rm r} = {\rm d}x^\mu_{\rm r}/{\rm d}s$. Using the metric defined in Eq. (2.134), we directly write

$$u^0_{\rm r} = \frac{1}{a_{\rm r}}(1-\psi) \quad , \quad u^i_{\rm r} = \frac{v^i}{a_{\rm r}} \,, \tag{2.147}$$

where $\boldsymbol{v}$ is the 3D peculiar velocity field.

We can then rewrite Eq (2.145) as

$$(1+z) = \frac{(g_{\mu\alpha}u^\mu_{\rm r}p^\alpha_{\rm r})|_{\rm e}}{(g_{\mu\alpha}u^\mu_{\rm r}p^\alpha_{\rm r})|_{\rm o}} = \frac{f(a_{\rm r})a_0}{a_{\rm r}f(a_0)}\left[\frac{(g_{\mu\alpha}u^\mu_{\rm r}k^\alpha_{\rm r})|_{\rm e}}{(g_{\mu\alpha}u^\mu_{\rm r}k^\alpha_{\rm r})|_{\rm o}}\right] = \frac{a_0^2}{a_{\rm r}^2}\left[\frac{(g_{\mu\alpha}u^\mu_{\rm r}k^\alpha_{\rm r})|_{\rm e}}{(g_{\mu\alpha}u^\mu_{\rm r}k^\alpha_{\rm r})|_{\rm o}}\right] . \tag{2.148}$$

Equipped with Eqs. (2.154), (2.140), (2.146) and (2.131), setting $a_0 = 1$ and using $a = (1+z)^{-1}$, we can then derive $\delta \ln a$ as follows

$$\begin{aligned} 1 + \delta\ln a &= a_{\rm r}^{-1}\frac{-a_{\rm r}(1+2\psi_{\rm e})(1-\psi_{\rm e})(-1+\delta f_{\rm e}) + a_{\rm r}\,(1-2\phi_{\rm e})\,\delta^{\rm K}_{ij}\,v^i_{\rm e}\,(n^j+\delta n^j_{\rm e})/c}{-(1+2\psi_{\rm o})(1-\psi_{\rm o})(-1+\delta f_{\rm o}) + \,(1-2\phi_{\rm o})\delta^{\rm K}_{ij}\,v^i_{\rm o}\,(n^j+\delta n^j_{\rm o})/c} \,, \\ &= \frac{(1+\psi_{\rm e})(1-\delta f_{\rm e}) + \,(1-2\phi_{\rm e})\,\delta^{\rm K}_{ij}\,v^i_{\rm e}\,(n^j+\delta n^j_{\rm e})/c}{(1+\psi_{\rm o})(1-\delta f_{\rm o}) + \,(1-2\phi_{\rm o})\delta^{\rm K}_{ij}\,v^i_{\rm o}\,(n^j+\delta n^j_{\rm o})/c} \,, \\ &= \frac{1+\psi_{\rm e}-\delta f_{\rm e}+v_{\parallel\,{\rm e}}/c}{1+\psi_{\rm o}-\delta f_{\rm o}+v_{\parallel\,{\rm o}}/c} \,, \end{aligned} \tag{2.149}$$

where we kept only linear order terms at each step, the final answer is

$$\delta\ln a = \frac{(v_{\parallel\,{\rm e}}-v_{\parallel\,{\rm o}})}{c} + (\psi_{\rm e}-\psi_{\rm o}) - \delta f_{\rm e} + \delta f_{\rm o} \,. \tag{2.150}$$

From here on out, the only two missing ingredients are the perturbations δf and δn_i. Using these two quantities, we can deduce δx^0 and δx^i by integrating Eq. (2.141) as well as compute $\delta \ln a$ (and consequently Δx^0) using Eq. (2.150). Finally, we can then deduce δx by plugging Eq. (2.138) in Eq. (2.139) to calculate Δx^i.

In order to compute the specific form of the perturbative quantities in Eq. (2.141). We need to solve the geodesic equation in real space. We can write the geodesic equation (at second order) as (Bertacca et al., 2014a)

$$\frac{{\rm d}k^\mu_{\rm r}}{{\rm d}x} + \left(\Gamma^\mu_{\alpha\,\beta} + \delta x^\gamma_{\rm r}\frac{\partial\Gamma^\mu_{\alpha\,\beta}}{\delta x^\gamma}\right) k^\alpha_{\rm r}\,k^\beta_{\rm r} = 0 \tag{2.151}$$

where $\Gamma^\alpha_{\mu\,\beta}$ are the Christoffel symbols using the metric defined in Eq. (2.134). By solving the previous equation at first order, the frequency change at emission and

observation can be found to be

$$\delta f = -\left[\psi_{\rm o} - v_{\parallel\,{\rm o}}/c\right] + 2\psi + \int_0^x \frac{(\phi' + \psi')}{c}\,{\rm d}\tilde{x} \tag{2.152}$$

while the change in direction is given by

$$\delta n^i = -v^i_{\rm o} - n^i\phi_{\rm o} - \int^x \partial_{\tilde{x}}\,(\psi + \phi)\,{\rm d}\tilde{x}\,, \tag{2.153}$$

where we introduced the $\partial_{\tilde{x}} = \partial/\partial\tilde{x}$.

Using Eq. (2.150), we can immediately derive

$$\boxed{\delta \ln a = \left[\frac{(v_{\parallel\,{\rm e}} - v_{\parallel\,{\rm o}})}{c} - (\psi_{\rm e} - \psi_{\rm o}) - \int_0^x \frac{(\phi' + \psi')}{c}\,{\rm d}\tilde{x}\right]\,,} \tag{2.154}$$

which is the relativistic version of Eq. 2.63.

As for δx^0 and δx^i, they can be computed by integrating Eqs (2.141). They are given by

$$\delta x^0 = -\,x\left[\psi_{\rm o} - v_{\parallel\,{\rm o}}/c\right] + \int_0^x \left[2\psi + (x - \tilde{x})\frac{(\phi' + \psi')}{c}\right]{\rm d}\tilde{x}\,, \tag{2.155}$$

$$\begin{aligned}\delta x^i =& \,n^i\,x\left[\psi_{\rm o} - v_{\parallel\,{\rm o}}/c\right] - x\,v^i{}_{\perp\,{\rm o}}/c \\ &- \int_0^x n^i\left[(\psi - \phi) + (x - \tilde{x})\frac{(\phi' + \psi')}{c}\right]{\rm d}\tilde{x} \\ &- \int_0^x (x - \tilde{x})\left[\partial_{\tilde{x}} - n^i\,(n_j\,\partial^j_{\tilde{x}})\right](\psi + \phi)\,{\rm d}\tilde{x}\,.\end{aligned} \tag{2.156}$$

Combining δx^0, Δx^0 and $\Delta \ln a$, we can derive

$$\begin{aligned}\delta x =& -\left(x + \frac{1}{\mathcal{H}}\right)\left[\psi_{\rm o} - v_{\parallel\,{\rm o}}/c\right] + \frac{1}{\mathcal{H}}\left[\psi_{\rm e} - v_{\parallel\,{\rm e}}/c\right] + \\ &\int_0^x \left[2\psi + (x - \tilde{x})\frac{(\phi' + \psi')}{c}\right]{\rm d}\tilde{x} + \frac{1}{\mathcal{H}}\int_0^x \frac{(\phi' + \psi')}{c}\,{\rm d}\tilde{x}\,.\end{aligned} \tag{2.157}$$

Putting all the results together, the coordinate map defined in Eq. 2.133 is given

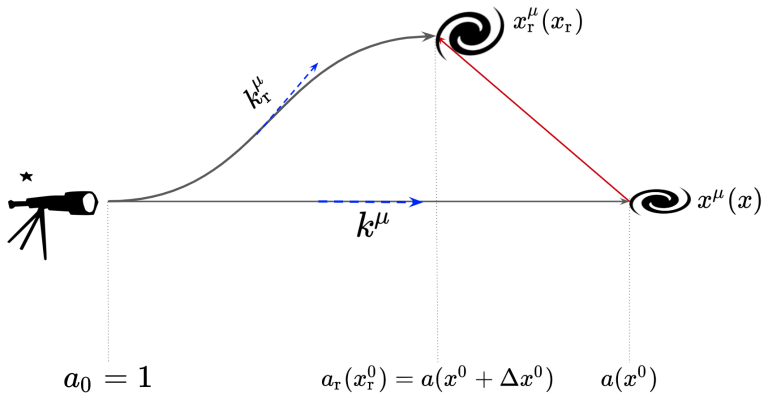


Figure 2.10: A schematic representation of Eqs. (2.158). This figure is based on a similar plot in (Bertacca et al., 2014b).

by

$$\begin{aligned}
\Delta x^0 &= \frac{c}{\mathcal{H}}\,\delta \ln a\,,\\
\Delta x^i &= -n^i\, x\left[\phi_{\rm o} + \psi_{\rm o} + \ v_{\parallel\,{\rm e}}/c\right] - x\,\frac{v^i_{\rm o}}{c} - \frac{c}{\mathcal{H}}\,n^i\,\delta \ln a\\
&\quad + n^i \int_0^x (x-\tilde{x})\,\frac{(\psi' + \phi')}{c}{\rm d}\,\tilde{x} - \int_0^x (x-\tilde{x})\,\delta^i_j \tilde{\partial}^j(\psi+\phi){\rm d}\,\tilde{x}\\
&\quad + 2n^i \int_0^x (\psi+\phi)\,{\rm d}\tilde{x}\,.
\end{aligned} \tag{2.158}$$

We show a schematic representation of the coordinate shift in Fig. 2.10.

2.3.3 Volume Transformation.

The second part of the derivation is to relate the real-space volume element to the observed redshift space volume element. This is achieved through the following relation (Yoo, 2010, Schmidt & Jeong, 2012, Bertacca et al., 2014a)

$${\rm d}\mathcal{V}_{\rm r} = \frac{\sqrt{-g(x^\mu_{\rm r})}}{a^4_{\rm r}}\,\epsilon_{\mu\nu\gamma\rho}\,\frac{u^\mu_{\rm r}}{a_{\rm r}}\,\frac{\partial x^\nu_{\rm r}}{\partial x^1}\,\frac{\partial x^\gamma_{\rm r}}{\partial x^2}\,\frac{\partial x^\rho_{\rm r}}{\partial x^3}\,{\rm d}\mathcal{V}\,. \tag{2.159}$$

In order to compute different linear order contributions of Eq. (2.159), we invoke the following perturbative expansion for the determinant B of a rank two tensor $\mathbb{B}$

$$B = B^{(0)} + B^{(1)}\,{\rm Tr}[\mathbb{B}^{(0)-1}\mathbb{B}^{(1)}] + \mathcal{O}(2)\,, \tag{2.160}$$

where the numbers within the parenthesis are the order of the perturbation and the Tr symbol denotes the trace. Applying that relation to the first part of Eq. (2.159),

i.e., the determinant of the metric, we find that at linear order it is given by

$$\sqrt{-g(x_{\rm r}^{\mu})}/a_{\rm r}^4 = 1 + \frac{g_{\mu}^{\mu}(x^{\mu})}{a_{\rm r}^2}\,. \tag{2.161}$$

Next, we turn to the second part of Eq (2.159) which can be expanded into

$$\epsilon_{\mu\nu\gamma\rho}\,\frac{u_{\rm r}^{\mu}}{a_{\rm r}}\,\frac{\partial x_{\rm r}^{\nu}}{\partial x^1}\,\frac{\partial x_{\rm r}^{\gamma}}{\partial x^2}\frac{\partial x_{\rm r}^{\rho}}{\partial x^3} \;=\; \frac{u_{\rm r}^0}{a_{\rm r}}\left|\frac{{\rm d}^3x_{\rm r}}{{\rm d}^3x}\right| + \frac{u_{\rm r}^i}{a_{\rm r}}\,T_i\;, \tag{2.162}$$

where

$$T_i = n_i - \frac{\partial x_{\rm r}^0}{\partial x^i} + \epsilon_{ijl}\epsilon^{iml}\,\frac{\partial x_{\rm r}^j}{\partial x^m} \tag{2.163}$$

We use Eq. (2.75) to compute the first term on the RHS and then we add the T_i contribution, the final result is given by (Bertacca, 2015)

$$1+\Delta V = \epsilon_{\mu\nu\gamma\rho}\,\frac{u_{\rm r}^{\mu}}{a_{\rm r}}\,\frac{\partial x_{\rm r}^{\nu}}{\partial x^1}\,\frac{\partial x_{\rm r}^{\gamma}}{\partial x^2}\frac{\partial x_{\rm r}^{\rho}}{\partial x^3} \;= 1 + \frac{u_{\rm r}^0 + n_i u_{\rm r}^i}{c\,a} + \left(\frac{2}{x} + \partial_{\parallel}\right)\Delta x_{\parallel} - 2\kappa\,, \tag{2.164}$$

where the gravitational weak-lensing convergence is defined as

$$\kappa = -\frac{1}{2}\partial_{\perp,i}\,\Delta x_{\perp}^i = \frac{1}{2}\int_0^x (x-\tilde{x})\frac{\tilde{x}}{x}\boldsymbol{\nabla}_{\perp}^2(\psi+\phi)\,{\rm d}\tilde{x} - v_{\perp\,{\rm o}}\,. \tag{2.165}$$

We note that the terms derived in the classical analysis (found in Eq. 2.78) are recovered in the previous equation, with additional relativistic contributions.

2.3.4 Galaxy Number Density

Magnification

The next natural step in the derivation is to compute how to relate the quantity $a_{\rm r}^3(x_{\rm r}^0)n_{\rm g}(x_{\rm r}^{\mu}, F_{\rm r})$ to its redshift counterpart. However, this is only possible after we account for the change in the observed flux. The change in the observed flux is usually quantified through the magnification (Schneider et al., 1992, Broadhurst et al., 1995)

$$\mathcal{M} = \frac{F_{\rm r}}{F} = \frac{1}{D_{\rm L\,,r}^2}D_{\rm L}^2 = \left(\frac{D_{\rm L\,,r}^2}{D_{\rm L}^2}\right)^{-2}\,, \tag{2.166}$$

where $D_{\rm L\,,r}$ and $D_{\rm L}$ are the luminosity distances in real and redshift space respectively.

The cosmic rulers prescription provides a method to compute the inverse of the previous relation (Schmidt & Jeong, 2012, Bertacca, 2015)

$$\mathcal{M}^{-1} = \frac{\sqrt{-g(x_{\rm r}^{\alpha})}}{(u_{{\rm r},\rho}\,{\rm d}x_{\rm r}^{\rho}(x_{\rm r})/{\rm d}x\;)\,a^2}\,\epsilon_{\mu\nu\rho\sigma}u^{\mu}\frac{{\rm d}x_{\rm r}^{\nu}(x_{\rm r})}{{\rm d}x}\frac{\partial x_{\rm r}^{\rho}}{\partial x^j}\frac{\partial x_{\rm r}^{\sigma}}{\partial x^k}\alpha^j\beta^k\,, \tag{2.167}$$

where $\{\alpha^i, n^i, \beta^i\}$ form a three dimensional orthonormal basis.

Expanding the previous relation up to linear order, Bertacca (2015) finds

$$\begin{aligned}\mathcal{M}^{-1} =& 1 + \frac{1}{2a_{\rm r}^2} g^\mu_\mu + \frac{u^0_{\rm r} + n_i u^i_{\rm r} + g_{\mu 0} u^\mu_{\rm r} - n^i\, g_{\mu i} u^\mu_{\rm r}}{c} \\ &+ 2\,\delta \ln a + \partial_x \left(\Delta x^0 + \Delta x_\parallel\right) + \frac{2}{x}\Delta x_\parallel - 2\kappa\,. \end{aligned} \tag{2.168}$$

After some algebraic manipulation, we find

$$\begin{aligned}\mathcal{M} =& 1 + \Delta\mathcal{M} = 1 - \frac{1}{2a_{\rm r}^2} g^\mu_\mu - \frac{u^0_{\rm r} + n_i u^i_{\rm r} + g_{\mu} u^\mu_{\rm r} - n^i\, g_{\mu i} u^\mu_{\rm r}}{c} - \\ & 2\,\delta \ln a - \partial_x \left(\Delta x^0 + n_i \Delta x^i\right) - \frac{2}{x} n_i \Delta x^i + 2\kappa\,. \end{aligned} \tag{2.169}$$

Plugging in the quantities computed in the previous sections, we find the magnification in the conformal Newtonian gauge to be

$$\begin{aligned}\mathcal{M} = 1 &+ 2\kappa + 2\psi_{\rm e} - 2\, v_{\parallel\,{\rm o}}/c \\ &- 2\left(1 - \frac{c}{\mathcal{H}x}\right)\delta \ln a - \frac{2}{x}\int_0^x (\psi + \phi)\,{\rm d}\,\tilde{x}\,. \end{aligned} \tag{2.170}$$

Number density perturbation

The final ingredient is then $a_{\rm r}^3(x^0_{\rm r})\, n_{\rm g}(x^\mu_{\rm r}, F_{\rm r})$, where we reparameterised the number density in terms of $\mathcal{M}$. We first look at the $a_{\rm r}^3$ prefactor which can immediately be derived at linear order

$$\left(\frac{a_{\rm r}}{a}\right)^3 = 1 + 3\,\delta \ln a\,. \tag{2.171}$$

As for the change in the number density, we expand $n_{\rm g}(x^\mu_{\rm r}, \mathcal{M})$

$$\begin{aligned} n_{\rm g}(x^\alpha_{\rm r}, F_{\rm r}) &= n_{\rm g}(x^\alpha + \Delta x^\alpha, F + F\Delta\mathcal{M}) \\ &= n_{\rm g}(x^\alpha, F) + \frac{\partial n_{\rm g}(x^0\,, F)}{\partial F} F\Delta\mathcal{M} + \frac{\partial n_{\rm g}(x^0\,, F)}{\partial x^0}\,\Delta x^0 \\ &= n_{\rm g}(x^\alpha, L_{\rm min}) - \frac{\partial n_{\rm g}(x^0\,, L_{\rm min})}{\partial \ln L_{\rm min}}\Delta\mathcal{M} + \frac{\partial n_{\rm g}(x^0\,, L_{\rm min})}{\partial x^0}\,\Delta x^0\,, \end{aligned} \tag{2.172}$$

where we recast the equation in terms of the Luminosity in the last equality. At this point, we define $\bar{n}_{\rm g} = n_{\rm g}(x^0\,, F)$ and the following bias parameters.

First, we define the Evolution bias which quantifies the change in the galaxy population due to redshift evolution

$$\left.\frac{\partial \bar{n}_{\rm g}}{\partial \bar{x}^0}\right|_{\bar{L}=4\pi D^2 F} = \mathcal{H}\left.\frac{\partial \bar{n}_{\rm g}}{\partial \ln \bar{a}}\right|_{\bar{L}} = -\mathcal{H}\left.\frac{\partial \ln \bar{n}_{\rm g}}{\partial \ln(1+z)}\right|_{L_{\rm min}=L_{\rm lim}(z)} = -\mathcal{H}\mathcal{E}\,, \tag{2.173}$$

where

$$\mathcal{E}(z) = -\left.\frac{\partial \ln \bar{n}_{\rm g}}{\partial \ln(1+z)}\right|_{L_{\rm min}=L_{\rm lim}(z)} . \tag{2.174}$$

We also define the magnification-bias parameter which represents the addition or removal of the galaxy population from the observed sample due to (de)magnification

$$\mathcal{Q}(z) = -\left.\frac{\partial \ln \bar{n}_{\rm g}}{\partial \ln L_{\rm min}}\right|_{L_{\rm min}=L_{\rm lim}(z)} . \tag{2.175}$$

We could thus rewrite Eq. (2.172) succinctly as

$$n_{\rm g}(x^\alpha_{\rm r}, F_{\rm r}) = \bar{n}_{\rm g}\left[1 + \delta_{\rm g} + \mathcal{E}\,\delta \ln a - \mathcal{Q}(\mathcal{M}-1)\right] . \tag{2.176}$$

2.3.5 Galaxy Density Contrast

We are finally ready to combine all the previous results to reproduce the observed density contrast

$$\delta_{\rm g,s} = \frac{n_{\rm g}(x^0, \boldsymbol{x}, F) - \bar{n}_{\rm g}(x^0, F)}{\bar{n}_{\rm g}(x^0, F)} \tag{2.177}$$

$$= \frac{n_{\rm g}(x^\alpha_{\rm r}, F_{\rm r}) - \bar{n_{\rm g}}}{\bar{n}_{\rm g}} + 3\delta \ln a^{(1)} + (1 - \sqrt{-g}) + \Delta V^{(1)} . \tag{2.178}$$

Putting all of the previous functions together, we find (Challinor & Lewis, 2011, Bertacca, 2015)

$$\begin{aligned}\delta_{\rm g,s}(\boldsymbol{x}) = \delta_{\rm g} &- \frac{1}{\mathcal{H}}\frac{\partial(\boldsymbol{v}_{\rm e}\cdot\boldsymbol{n})}{\partial x} + 2(1-\mathcal{Q})\kappa + \frac{2(1-\mathcal{Q})}{x}\int_0^x (\psi+\phi)\,{\rm d}\tilde{x} \\ &+ \left[\mathcal{E} - 2\mathcal{Q} - \frac{\mathcal{H}'}{\mathcal{H}^2} - \frac{2(1-\mathcal{Q})\,c}{x\,\mathcal{H}}\right] \times \left[\frac{\boldsymbol{v}_{\rm e}}{c}\cdot\boldsymbol{n} - (\psi_{\rm e} - \psi_{\rm o}) - \int_0^x \frac{(\psi'+\phi')}{c}\,{\rm d}\tilde{x}\right] \\ &- 2(1-\mathcal{Q})\,\psi_{\rm e} + \phi_{\rm e} + \frac{\psi'_{\rm e}}{\mathcal{H}} + \left[2 - \mathcal{E} + \frac{\mathcal{H}'}{\mathcal{H}^2} + \frac{2(1-\mathcal{Q})\,c}{x\,\mathcal{H}}\right]\frac{\boldsymbol{v}_{\rm o}}{c}\cdot\boldsymbol{n} .\end{aligned} \tag{2.179}$$

We note that the method developed in this thesis is simulation-based as outlined in Chapter 3. This implies that our input for the galaxy overdensity in real space is defined in the simulation gauge, which is given by the synchronous-comoving gauge. However, our treatment so far has been in the Poisson gauge. To account for the difference, we apply the following relation (Jeong et al., 2012, Borzyszkowski et al., 2017)

$$\delta_{\rm g} = b\delta_{\rm SC} + (3 - \mathcal{E})\,\mathcal{H}\Phi_{\rm v} , \tag{2.180}$$

where $\Phi_{\rm v}$ is the linear velocity potential and $\delta_{\rm SC}$ is the DM overdensity in the synchronous-comoving gauge (see Sec. 2.3.1). Consequently, Eq. (2.179) is written

as

$$
\begin{aligned}
\delta_{\mathrm{g,s}}(\boldsymbol{x}) = {} & b\,\delta_{\mathrm{SC}} - \frac{1}{\mathcal{H}}\frac{\partial(\boldsymbol{v}_{\mathrm{e}}\cdot\boldsymbol{n})}{\partial x} + 2(1-\mathcal{Q})\kappa \\
& + \left[\mathcal{E} - 2\mathcal{Q} - \frac{\mathcal{H}'}{\mathcal{H}^2} - \frac{2(1-\mathcal{Q})\,c}{x\,\mathcal{H}}\right] \\
& \times \left[\frac{\boldsymbol{v}_{\mathrm{e}}}{c}\cdot\boldsymbol{n} - (\psi_{\mathrm{e}} - \psi_{\mathrm{o}}) - \int_0^x \frac{(\psi' + \phi')}{c}\,\mathrm{d}\tilde{x}\right] \\
& - 2(1-\mathcal{Q})\,\psi_{\mathrm{e}} + \phi_{\mathrm{e}} + \frac{\psi'_{\mathrm{e}}}{\mathcal{H}} + (3-\mathcal{E})\,\mathcal{H}\Phi_{\mathrm{v}} \\
& + \left[2 - \mathcal{E} + \frac{\mathcal{H}'}{\mathcal{H}^2} + \frac{2(1-\mathcal{Q})\,c}{x\,\mathcal{H}}\right]\frac{\boldsymbol{v}_{\mathrm{o}}}{c}\cdot\boldsymbol{n} \\
& + \frac{2(1-\mathcal{Q})}{x}\int_0^x (\psi + \phi)\,\mathrm{d}\tilde{x}\,.
\end{aligned}
\tag{2.181}
$$

Equation (2.181) captures the concept of "relativistic (linear) RSD" and serves as the foundation of our investigation. Simply put, it asserts that discrepancies in galaxy overdensities between real and redshift space stem from various physical phenomena. The second term on the right-hand side corresponds to the classic Kaiser correction, attributed to peculiar velocity variations along the line of sight (Kaiser, 1987). Additionally, the third term accounts for weak-lensing effects arising from volume and magnification corrections, which are anticipated to influence clustering statistics on large scales (e.g., Matsubara, 2000b, Hui et al., 2007, 2008, Yoo et al., 2009, Challinor & Lewis, 2011, Raccanelli et al., 2016b, Borzyszkowski et al., 2017). Further corrections reliant on gravitational potentials and peculiar velocities follow. Previous studies have demonstrated that terms proportional to the peculiar velocity of the observer, particularly the last term in Eq. (2.181), could manifest observable features in the galaxy two-point correlation function at wide angles (Bertacca, 2020). Furthermore, they can superimpose an oscillatory signal onto the power-spectrum monopole at very large scales as we see in Chapter 7.

When are these effects significant?

As seen in Eq. (2.181), most of the additional terms due to relativistic effects are of the order of the Bardeen potentials ($\sim \phi, \psi$). Assuming ϕ and ψ are of the same order as the potential φ/c^2 of the Poisson Equation (see Eq. 2.32).[10] We can gauge the regimes of significance of these effects by looking at the scales where φ is comparable to the real-space overdensity δ. We accomplish that by examining the Poisson equation in Fourier space (Fourier transforming Eq. 2.32)

$$
k^2\tilde{\varphi}(\boldsymbol{k}) = -\frac{3H_0^2\Omega_{\mathrm{m}}}{2a}\,\tilde{\delta}(\boldsymbol{k})\,. \tag{2.182}
$$

[10] A fair assumption to make as they are identical in the ΛCDM model at linear order.

Let us rewrite the previous equation using the physical scale $\lambda_{\rm p} = a/k$ and the Hubble radius $\lambda_{\rm H} = 1/H$, i.e.,

$$\tilde{\varphi}(\boldsymbol{k}) \approx -\left(\frac{\lambda_{\rm p}}{\lambda_{\rm H}}\right)^2 \tilde{\delta}(\boldsymbol{k})\,. \tag{2.183}$$

From the previous relation, we can observe that the gravitational potential is reduced by an amplitude factor of $(\lambda_{\rm p}/\lambda_{\rm H})^2$ compared to the matter density contrast. Consequently, effects that scales as φ are negligible for $(\lambda_{\rm p}/\lambda_{\rm H})^2 \ll 1$. However, that factor is only ~ 1 when the physical scale $\lambda_{\rm p}$ approaches the Hubble radius. We are then able to conclude that the φ related effects become significant at large scales, i.e., small k where $\lambda_{\rm p} \sim \lambda_{\rm H}$.

2.3.6 The Velocity of the Observer Term

We devote a considerable part of this work to the study of the relativistic RSD terms induced by the peculiar velocity of the observer $\boldsymbol{v}_{\rm o}$ (see Chaps. 6 and 7). Let us briefly discuss how the velocity of the observer adjusts the galaxy density contrast in the classical and relativistic treatments, highlighting the differences between the two results.

To that end, let us rewrite Eq. (2.181), i.e.

$$\delta_{\rm g,s}(\boldsymbol{x}) = \delta_{\rm com}(\boldsymbol{x}) + \frac{\alpha_{\rm o}(x)}{x}\,\frac{\boldsymbol{v}_{\rm o}\cdot\hat{\boldsymbol{x}}}{a\,H}\,, \tag{2.184}$$

where

$$\alpha_{\rm o}(x) = 2(1-\mathcal{Q}) \;-\; \frac{x\,H}{c\,(1+z)}\,\mathcal{E} \;+\; \frac{x\,H}{c\,(1+z)}\left[3 - \frac{1+z}{H}\,\frac{{\rm d}H}{{\rm d}z}\right]\,, \tag{2.185}$$

and $\delta_{\rm com}$ collects all the terms that do not depend on $\boldsymbol{v}_{\rm o}$.

Repeating that exercise for the classical result in Eq. (2.80), we find that the equation is the same as Eq. (2.184) except for the function form of $\alpha_{\rm o}$. Analytically, we find [11]

$$\delta^{\rm c}_{\rm g,s}(\boldsymbol{x}) = \delta^{\rm c}_{\rm com}(\boldsymbol{x}) + \frac{\alpha_{\rm c}(x)}{x}\,\frac{\boldsymbol{v}_{\rm o}\cdot\hat{\boldsymbol{x}}}{a\,H}\,, \tag{2.186}$$

where

$$\alpha_{\rm c}(x) = \frac{{\rm d}\,\ln x^2\,\bar{n}_{\rm g,s}(x)}{{\rm d}\ln x} + \frac{Hx}{c(1+z)}\left[1 - \frac{{\rm d}\ln H}{{\rm d}\ln(1+z)}\right]\,. \tag{2.187}$$

To facilitate the comparison between Eqs (2.184) and (2.186), let us rewrite $\alpha_{\rm c}$ in terms of the magnification and evolution biases. To that end, we expand the total

[11] we add the superscript c to differentiate the classical and relativistic results

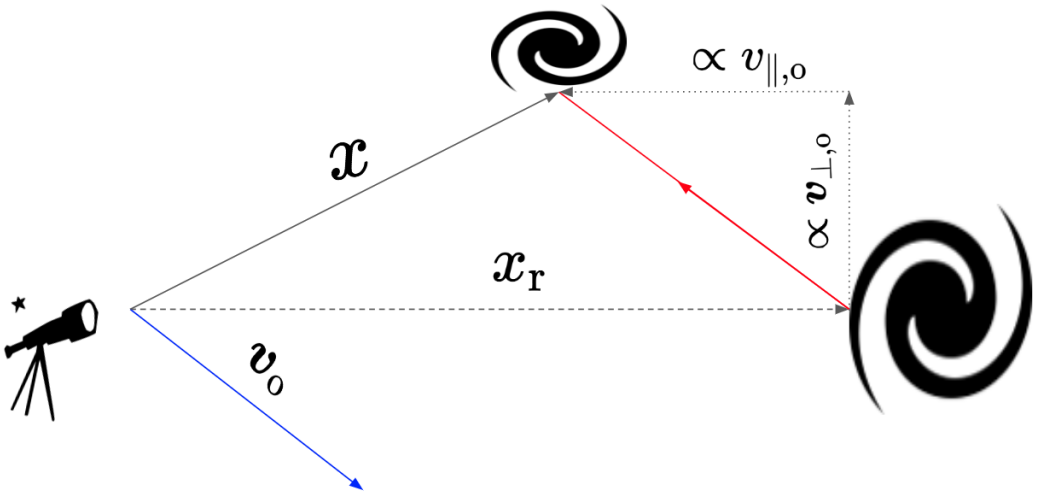


Figure 2.11: A schematic representation of the impact of the velocity of the observer in the relativistic treatment.

logarithmic derivative of the density as follows (see also Paper I)

$$\begin{aligned}
\frac{\mathrm{d}\,\ln x^2\,\bar{n}_{\mathrm{g,s}}(x)}{\mathrm{d}\ln x} &= 2 + \frac{\mathrm{d}\ln \bar{n}_{\mathrm{g,s}}(x)(x)}{\mathrm{d}\ln x} \\
&= 2 + \left.\frac{\partial \ln \bar{n}_{\mathrm{g}}}{\partial \ln L_{\mathrm{min}}}\right|_{L_{\mathrm{min}}=L_{\mathrm{lim}}(z)} \frac{\mathrm{d}\ln L_{\mathrm{lim}}}{\mathrm{d}\ln x} + \frac{\partial \ln \bar{n}_{\mathrm{g}}}{\partial \ln x} \\
&= 2 - Q\,\frac{\mathrm{d}\ln L_{\mathrm{lim}}}{\mathrm{d}\ln x} + \frac{\partial \ln \bar{n}_{\mathrm{g}}}{\partial \ln(1+z)}\,\frac{\mathrm{d}\ln(1+z)}{\mathrm{d}\ln x} \\
&= 2 - Q\,\frac{\mathrm{d}\ln L_{\mathrm{lim}}}{\mathrm{d}\ln x} - \mathcal{E}\,\frac{H\,x}{c\,(1+z)} \\
&= 2\,(1-Q) - (2Q+\mathcal{E})\,\frac{H\,x}{c\,(1+z)}\,.
\end{aligned} \tag{2.188}$$

Plugging the result back in Eq. (2.187), we find

$$\alpha_{\mathrm{c}}(x) = 2\,(1-Q) - \frac{H\,x}{c\,(1+z)}\left[2Q + \mathcal{E} + 1 - \frac{\mathrm{d}\ln H}{\mathrm{d}\ln(1+z)}\right]\,, \tag{2.189}$$

which clearly differs from Eq. (2.185) by

$$\Delta\alpha \equiv \alpha_c - \alpha_{\mathrm{o}} = \frac{rH}{c(1+z)}\,(-2\mathcal{Q} - 2)\,. \tag{2.190}$$

We can understand the difference in the two results by contrasting the impact of the $\boldsymbol{v}_{\mathrm{o}}$ in the coordinate maps of both the classical and relativistic treatments. As shown schematically in Fig. 2.11, in the relativistic prescription, $\boldsymbol{v}_{\mathrm{o}}$ alters i) the ob-

served redshift through the Doppler shift (see Eq. 2.154), *ii*) the observed direction through relativistic aberration (see Eq. 2.153), and *iii*) the observed Flux through the magnification (see Eq. 2.170). On the other hand, in the classical treatment, only the Doppler shift is included in the analysis. We exploit the difference between the two treatments in Chapter 7 in order to measure the velocity of the observer and other cosmological parameters.

Chapter 3

THE LIGER METHOD

3.1 OUTLINE OF THE LIGER METHOD

The LIght cones with GEneral Relativity (LIGER) method (Borzyszkowski et al., 2017, Paper I, Paper II, Paper III) is a numerical recipe for constructing realizations of the galaxy distribution on the past light cone of an observer. The method takes two inputs: *i)* a Newtonian N-body simulation and *ii)* four survey functions that describe the average number density, linear bias, magnification bias and evolution bias of the target galaxy population of the survey of interest. The method outputs the mock galaxy catalogues that take into account relativistic RSD at linear order in cosmological perturbations. The LIGER is particularly suited to estimate galaxy correlations on large spatial scales. Throughout this project, we use the LIGER method to produce mock catalogues for a *Euclid*-like survey (Laureijs et al., 2011) which were subsequently used in Paper I and III. As well as a full-sky SKA02-like survey which is used in Paper II. In the following sections, we outline the specifics of the LIGER method which are summarily described by the schematic in Fig. 3.1.

3.1.1 COORDINATE TRANSFORMATION: THE LIGER CODE

The LIGER code[1] computes both the coordinate maps (Eq. 2.158) and the magnification (Eq. 2.170) starting from the real-space position of each particle in the input N-body simulation. Then the code identifies the snapshots within which the backward light cone of the observer intersects with the galaxy's world line after adding the displacement written Eq. (2.158). Finally, we output the intersection position, the corresponding magnification, and redshift change. In this subsection, we outline the details of the computation, referring to Eqs.(2.158) and (2.170) as the 'shift equations'.

[1]The code is publicly accessible at LIGER .

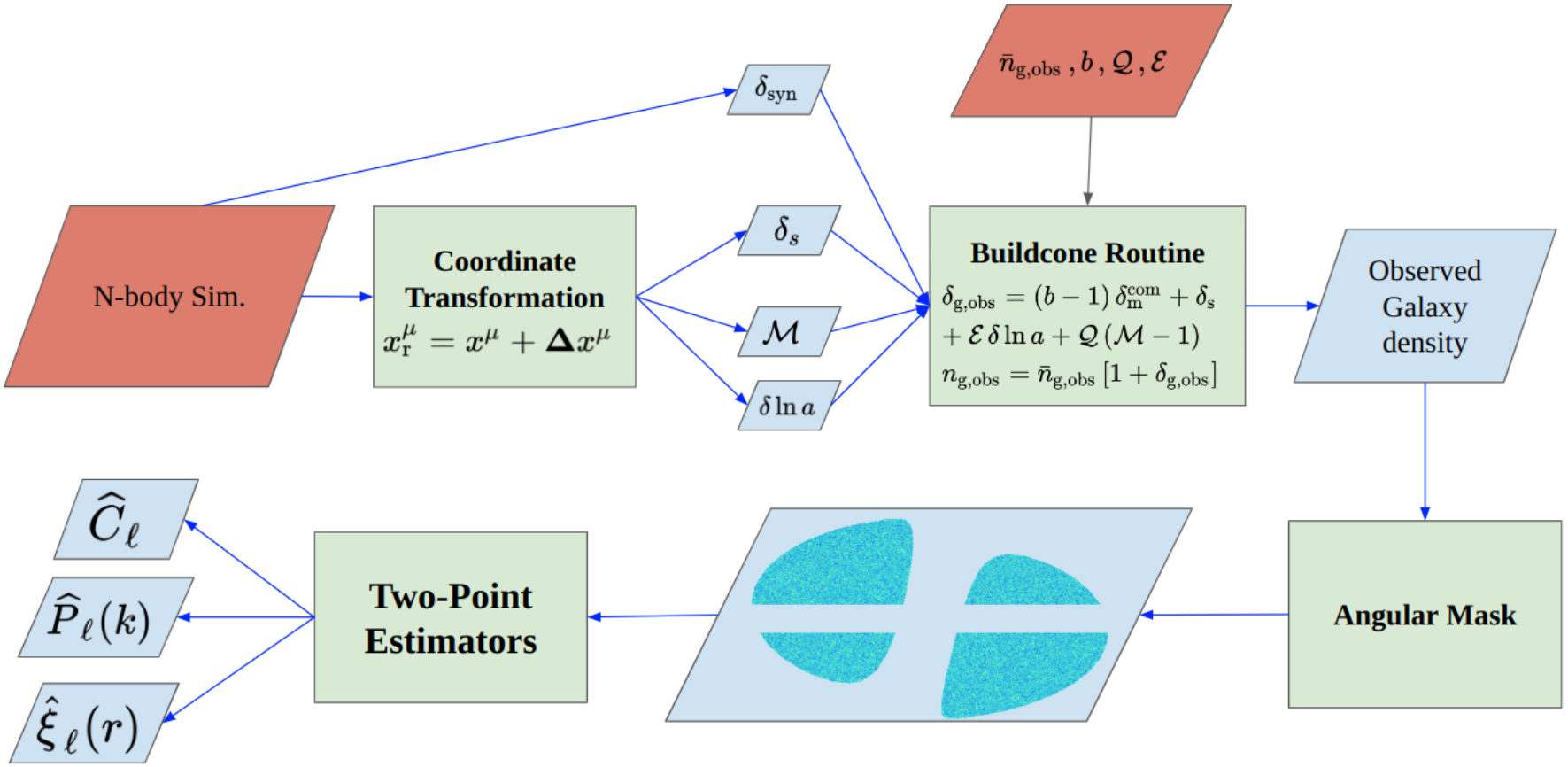


Figure 3.1: Flow chart of the LIGER method in the 'large-box' mode. Inputs are shown as red parallelograms. Outputs are displayed as blue ones and processes are plotted as green rectangles.

Newtonian to Relativistic dictionary

The input to the LIGER code are snapshots of N-body simulations at different redshifts. In order to connect the quantities defined in the N-body simulation to the functions necessary for the computation of the shift equations, the potential functions need to be written in terms of the N-body simulations particle overdensity . The matter density contrast of N-body simulations $\delta_{\rm sim}$ and its source gravitational potential $\varphi_{\rm sim}$ correspond to their synchronous-comoving gauge quantities (see Sec. 2.3.1), i.e., $\delta_{\rm sim} = \delta_{\rm SC}$ and $\varphi_{\rm sim} = \varphi$. It has been shown that at linear order in the ΛCDM model, the Poisson gauge ψ (see Eq. 2.134) can be obtained from the synchronous-comoving density contrast $\delta_{\rm SC}$ (Bardeen, 1980, Chisari & Zaldarriaga, 2011, Green & Wald, 2012, Bruni et al., 2014). The dictionary that translates the N-body output to the Poisson gauge is given by (for more details, see Sec. 2.1 in Borzyszkowski et al., 2017)

$$\psi\, c^2 = \phi\, c^2 = \varphi\,, \qquad \nabla_x^2 \varphi = \frac{3H_0^2\Omega_{\rm m}}{2a}\delta_{\rm sim}\,, \qquad v^i = v^i_{\rm sim}\,, \tag{3.1}$$

where $v^i_{\rm sim}$ is the simulation 3D velocity field. For all our subsequent applications, we compute the potential on a 512^3 grid in a box of length $L_{\rm box} = 12\,h^{-1}\,$Gpc (see Sec. 3.2 for details about the input simulations). We note that the c^2 factor in the first relation in Eq. (3.1) is due to the dimensionless definition of the Bardeen potentials in Eq. (2.134).

Potential Computation

The shift equations require $\phi(\boldsymbol{x}, z)$, its first and second-order spatial derivatives as well as its temporal derivatives. To compute ϕ, we first paint the simulation overdensity $\delta_{\rm sim}(\boldsymbol{x}_j, z_m)$ using a cloud in cell (CIC Hockney & Eastwood, 1988) particle mesh algorithm for each snapshot, m, of the input simulation. Numerically, we compute

$$\delta_{\rm sim}(\boldsymbol{x}_j, z_m) = \frac{1}{\bar{n}_{\rm sim}} \sum_i W_{\rm CIC}(\boldsymbol{x}_j - \boldsymbol{r}_i^m) - 1\,, \tag{3.2}$$

where $\bar{n}_{\rm sim} = N_{\rm sim}/V_{\rm sim}$ is average particle number density of the input simulation, $\boldsymbol{x}_j$ is the grid cell coordinate position, z_m is the redshift of the mth snapshot, $\boldsymbol{r}_i^m$ is the coordinate position of the ith particle in the snapshot m of the input simulation and $W_{\rm CIC}$ denotes the 'cloud-in-cell' kernel.

We then solve the Poisson equation in Fourier space using

$$\tilde{\varphi}(\boldsymbol{k}_j, z_m) = -\frac{3H_0^2\Omega_{\rm m}}{2a\,k_j^2}\,\tilde{\delta}(\boldsymbol{k}_j, z_m)\,, \tag{3.3}$$

where $\boldsymbol{k}_j$ is the wavenumber on the Fourier grid. We perform the discrete Fourier transforms (DFTs) using the Fastest Fourier Transform in the West (FFTW Frigo &

Johnson, 2005) algorithm. The spatial derivatives of φ are also computed via DFTs with different combinations of the $\tilde{\varphi}(\boldsymbol{k}_j, z_m)$ and the elements of the wavenumber $\boldsymbol{k}_j$ (also known as spectral differentiation Canuto et al., 2007). Finally, the temporal derivatives are computed using a Finite-Difference Method (FDM) that combines the potential at various redshifts at position $\boldsymbol{x}_j$ (Borzyszkowski et al., 2017). For all subsequent calculations, the discrete gridded functions are interpolated in the line of sight distance (the adopted affine parameter) to compute the local and integral terms in the shift equations.

Particle Shift

Equipped with all the necessary ingredients, we are then able to compute the shift equations. We invert Eq. (2.133), $x^\mu(x) = x^\mu_{\rm r}(x_{\rm r}) - \Delta x^\mu(x)$, to compute the redshift positions $x^\mu(x)$ from the real-space (simulation) positions $x^\mu_{\rm r}(x_{\rm r})$. Examining the functional form of $\Delta x^\mu(x)$ in Eq. (2.158), we note that there are local terms and integral terms that should be computed along the line of sight from the observed to the source position **in redshift space** where the photon worldline is straight (see Fig. 2.10). As our input (the output of the N-body simulations) is in realspace, we adopt the Born approximation and compute line-of-sight integrals in realspace which is sufficient in the large-scale regimes we target in this work, as the angular deflections are generally small (For extension to beyond the Born approximation, see Bertacca, 2015). We compute the line of sight integrals using the fast voxel traverse algorithm (Amanatides & Woo, 1987) using the gridded discrete values of the potential and its derivatives. The second subtlety in the computation of the

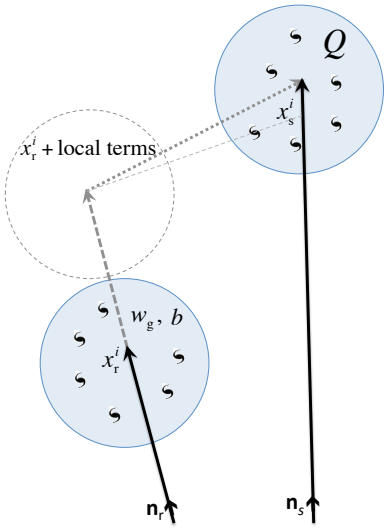


Figure 3.2: Schematic showing the shift calculation in LIGER. The particle is shifted by the local terms first represented by the dashed line. The new position is then used to compute the integral shifts. The integral shifts represented by the dotted arrow are then applied. This figure is borrowed from (Borzyszkowski et al., 2017).

line-of-sight integrals is that the upper limit is the redshift space position x which

is unknown at the start of the procedure. Instead of evaluating the integrals up to the real-space position $x_{\rm r}$, we adopt the following iterative procedure.

i) compute the redshift position using **only** local terms in the shift equations,

ii) evaluate the integral terms in the shift equations up to that local-terms shifted position,

iii) subsequently, apply the integral shifts to the particle position,

iv) finally, repeat the last two steps once.

In principle, we could repeat steps *ii* and *iii* till convergence, but testing the procedure we found that repeating the procedure more than once does not produce any significant change in the particle position. A schematic representation of this process is shown in Fig. 3.2.

Particles are observed when and where their worldline intersects with the observer's backward lightcone. To pinpoint this intersection, we locate several snapshots around it for each particle in the simulations. Subsequently, we compute the redshift position (and magnification) for each of these snapshots, following the methodology outlined in this section. Finally, we employ cubic interpolation to determine the values of the calculated quantities (redshift positions and magnification) precisely at the intersection point of the photon worldline and the backward lightcone. This intersection point satisfies the line element condition for null worldlines, $\mathrm{d}s^2 = 0$.

3.1.2 BUILDCONE ROUTINE

The LIGER method contains two operational modes (cf. Paper I). For our purposes, we choose the one called 'large-box' mode. This version computes the shift equations (as outlined in the previous section) on dark matter particles extracted from large-box low-resolution simulations. It requires a second toolkit, which we name Buildcone that takes in the dark-matter lightcone and produces the galaxy map tailored to a particular survey. The Buildcone procedure can be summarized as follows. We rewrite Eq. (2.181) in the following way

$$\delta_{\rm g,s} = (b-1)\,\delta_{\rm SC} + \delta_{\rm s} + \mathcal{E}\,\delta\ln a + \mathcal{Q}\,(\mathcal{M}-1)\,. \tag{3.4}$$

Equation (3.4) relates the produced redshift space dark matter overdensity to its galaxy counterpart[2]. It constitutes the master equation that the Buildcone toolkit implements. Note that $\delta_{\rm SC}$ is the real-space matter overdensity in the synchronous-

[2]Here, in Eq. (3.4), we follow Borzyszkowski et al. (2017) in assuming that $|\mathcal{H}\Phi_v| \ll |\delta\ln a|$.

comoving gauge and $\delta_{\rm s}$ is the redshift-space density contrast given by

$$\begin{aligned}\delta_{\rm s} =& \delta_{\rm SC} + \left(\frac{\partial_0 \mathcal{H}}{\mathcal{H}^2} + \frac{2c}{x\mathcal{H}}\right) \delta \ln a + \phi_{\rm e} - 2\psi_{\rm e} + \frac{(\partial_0 \phi)_{\rm e}}{\mathcal{H}} + 3\mathcal{H}\phi_v \\ & - \frac{1}{\mathcal{H}} \left[n^i \partial_i \left(n^j v_j\right)\right]_{\rm e} + \frac{2}{x} \int_0^x (\phi + \psi)\, {\rm d}\tilde{x} - 2\kappa \,. \end{aligned} \tag{3.5}$$

Regarding the overdensity $\delta_{\rm SC}$, let us note that in order to incorporate correctly the galaxy bias, it should be defined in the rest frame of cold dark matter, which is assumed to coincide with the rest frame of galaxies on large scales. The synchronous-comoving gauge sets a frame where the spherical collapse model has an exact GR interpretation (Wands & Slosar, 2009). Consequently, the synchronous-comoving gauge is suitable for the peak-background split approach (Press & Schechter, 1974) and for defining the bias parameters. In that frame the fluctuations of galaxy number density are related to the underlying matter density fluctuation on cosmological scales by a local bias (see Sec. 2.3.1, Jeong et al., 2012).

Assuming the linear b, magnification $\mathcal{Q}$ and evolution biases $\mathcal{E}$ are defined by the survey, we can construct the galactic density by building the four functions $\delta_{\rm SC}$, $\delta_{\rm s}$, $\delta \ln a$, and $\mathcal{M}$.

Let us define the resultant intersection (light-cone) position of the LIGER code for the ith particle with and without any RSD effects as r_i and s_i, respectively. We denote the particles' redshift change and magnification by $\delta \ln a_i$ and $\mathcal{M}_i$. We use the particles information to build $\delta_{\rm SC}$ on a finite grid via a mass-weighted 'cloud-in-cell' smoothing to (Hockney & Eastwood, 1988, see also Appendix A2 in Paper I),

$$\delta(\boldsymbol{x}_j) = \frac{1}{\bar{\eta}} \sum_i W_{\rm CIC}(\boldsymbol{x}_j - \boldsymbol{r}_i) - 1 \,, \tag{3.6}$$

where $\bar{\eta}$ are the average particle number density of the lightcone particle distribution, $\boldsymbol{x}_j$ is the grid cell coordinate position and $W_{\rm CIC}$ denotes the 'cloud-in-cell' kernel[3]. We calculate $\delta_{\rm s}$ by using the redshift positions s_i instead in Eq. (3.6) [4].

The determination of the remaining fields $\delta \ln a$ and $\mathcal{M}$ introduces further complexities as we only know of the redshift change and magnification at the locations of the particles. Consequently, using Eq. (3.6) directly results in accurately describing these functions in volumes where there are more particles while poorly predicting them in relatively empty regions. To mitigate that issue, we opt for a mass-weighted average. Taking the magnification as an example, the mass-weighted

[3]We note that Eqs. (3.2) and (3.6) both represent the 'cloud-in-cell' smoothing. The only difference in the input as the former takes the particle positions at each simulation snapshot while the latter takes in the lightcone particle positions.

[4]In using Eq. (3.6), we ignore both the velocity potential terms and the potential terms from the linear perturbation of $\sqrt{-g}$ (see Eq. 2.161) in $\delta_{\rm s}$. However, these terms are negligible for scales smaller than the horizon(for more details, see Borzyszkowski et al., 2017).

average is defined as

$$\mathcal{M}(\boldsymbol{x}_j) = \frac{\sum_i \mathcal{M}_i \, W_{\rm CIC}(\boldsymbol{x}_j - \boldsymbol{r}_i)}{\sum_i W_{\rm CIC}(\boldsymbol{x}_j - \boldsymbol{r}_i)} \,. \tag{3.7}$$

Finally, $\delta \ln a$ is produced similarly.

The mass-weighted average is accurate at linear order but deviates from the true field at higher orders which is sufficient for our purposes. We note that mass-weighted estimators are usually employed to estimate the peculiar velocity field in N-body simulations (e.g. Bernardeau & van de Weygaert, 1996).

Equipped with all four fields, we are able to compute $\delta_{\rm g,s}(\boldsymbol{x}_j)$ on a fixed grid. In all our applications, we compute these fields within a box of length $L_{\rm BC} = 10h^{-1}\,$Gpc with 1024^3 cells. The final step is to enforce the predicted average density on the galaxy density contrast, to that end, we simply use

$$n_{{\rm g},s}(\boldsymbol{x}_j) = \bar{n}_{\rm g,s}(x_j)\,[1 + \delta_{{\rm g},s}(\boldsymbol{x_j})] \,, \tag{3.8}$$

where $\bar{n}_{\rm g,s,obs}(x)$ is the expected average background comoving number density of the survey. From here on, we drop the s subscript when referring to the redshift space functions (the galaxy density and overdensity).

3.1.3 Code Improvements

This thesis uses `LIGER V2` which improves on the initial one (Borzyszkowski et al., 2017, `LIGER V1`) in several ways. In terms of additional physics, we take into account the peculiar velocity and potential of the observer in the calculation of the shift equations. These contributions had been neglected in the original code. A second difference is that we add the option to either set a user-defined observer velocity or the code associates the observer with a simulation particle (the closest one to the user-selected location). From the numerical point of view, we made changes that led to a considerable speed up (especially for the calculation of the line-of-sight integrals) and reduced memory consumption. We also fixed multiple numerical issues regarding the implementation of the shifts, handling particles too close to the observer and the magnification computation. Moreover, I/O operations has been improved with some completely new features. Finally, we designed and implemented the new `Buildcone` toolkit that allows the usage of the `LIGER` method directly on low-res large box DM simulations.

3.2 Simulations

In this section, we discuss the input N-body simulations we use for the projects discussed in this thesis. In principle, the `LIGER` method can be applied directly to

calculate coordinate maps of evolved galaxies in hydrodynamic simulations. However, reproducing the correct clustering for modes comparable to the *Euclid* survey size (see the discussion in Borzyszkowski et al., 2017, Paper I) requires simulating the matter density in a box larger than twice the maximum comoving radial distance of the survey in question. Consequently, the box should fit twice the comoving distance corresponding to $z = 1.8$ in order to produce *Euclid* (the deepest survey we discuss in the work) mock catalogues. The side of the simulation box should be at least $L_{\rm BOX} \geq 2 \times 3.4 h^{-1}$Gpc. However, simulating cosmological volumes of that size using hydrodynamic simulations with an adequate mass and spatial resolution to trace galaxy formation is not computationally feasible with the current computational capabilities, especially considering that we need a sufficiently large sample of simulations in order to numerically compute mock covariances.

To prevent replications and/or fictitious correlations at very large scales (e.g. due to the periodic boundary conditions of the simulation box) we simulate the matter density field in a much larger box with a comoving side length $L_{\rm BOX} = 12\,h^{-1}$Gpc. We use the second-order Lagrangian perturbation theory (2LPT) code `MUSIC` (Hahn & Abel, 2011) to build the dark matter distributions as we are only interested in linear and quasi-linear scales. The simulation box is run with 1024^3 particles, with a mass $m_{\rm par} = 1.4 \times 10^{14}{\rm M}_{\odot}$ each. Initially, the particles form a Cartesian grid with a cell side length $l_{\rm CELL} = 11.7\,h^{-1}$Mpc.

We adopt a cosmology based on the Planck mission results (Planck Collaboration: Aghanim et al., 2020a), setting the background matter density parameter $\Omega_{\rm m} = 0.3158$, the baryonic density parameter $\Omega_{\rm b} = 0.0508$ and the Hubble dimensionless parameter $h = 0.673$. We assume that the primordial power spectrum of power-law shape is characterized by a spectral index $n_{\rm s} = 0.966$ and an amplitude $A_{\rm s} = 2.1 \times 10^{-9}$ defined at the pivot wavenumber $k_* = 0.05\ {\rm Mpc}^{-1}$. We compute the matter transfer function using the `CAMB` code (Lewis & Bridle, 2002). We compute the initial power spectrum needed for the `MUSIC` runs using the Boltzmann solver `CAMB` code, see (Lewis & Bridle, 2002).

We note that the grid size employed, $l_{\rm CELL}$, is of the same order as the 2LPT shifts computed by the `MUSIC` code, consequently, the particle positions produced by `MUSIC` form a slightly perturbed grid, which in turn means that the simulations are unable to accurately capture the shot noise (SN) contribution (Gabrielli, 2004, Joyce & Marcos, 2007). Of course, the SN correction affects both the galaxy catalogues and their summary statistics (e.g., for the two-point statistics and their associated covariances, see Landy & Szalay, 1993, Feldman et al., 1994, Tegmark et al., 1998a). For that reason, we manually add Poissonian SN to our light cones in order to mimic the actual survey as will be outlined in Sec. 3.4.2 (see also Appendix A3 in Elkhashab et al., 2021).

We produce 35 realizations using the `MUSIC` code saving 21 snapshots, linearly spaced in the scale factor a in the range $a \in [1/3, 1]$. Finally, by positioning the observers in the box at the centres of an Hexagonal Close Packing (HCP) lattice, we are able to extract 4 light cones out of each simulation whilst ensuring that none of the observers' lightcones intersect. Consequently, the total number of produced

lightcones is 140.

3.2.1 Simulation Validation

For all the projects discussed in this thesis, it's imperative to produce dark matter realizations that accurately describe the large-scale clustering of galaxies. In this subsection, we discuss the numerical techniques tested before adopting the 2LPT `MUSIC` code.

We tested two numerical techniques to produce relatively fast and large-scale-accurate dark matter realizations that we can use as input to the `LIGER` method. First, we used the `L-PICOLA` code (Howlett et al., 2015), based on the `COLA` method (Tassev et al., 2013). Second, we tested the `2LPT` approach implemented in the `MUSIC` code. For both codes, we ran two simulations using 512^3 and 1024^3 particles, respectively. In the `L-PICOLA` case, the gravity solver was always employed using 1024^3 grid points. All simulations employed the same box size, $L_{\rm BOX} = 12\,h^{-1}$Gpc, and the initial power spectrum was generated using the `CAMB` code.

We present a comparison of the power spectra from both simulations in Fig. 3.3. The `MUSIC` spectra are depicted with solid red lines for the 1024^3 particle case and dashed red lines for the 512^3 particle case, while the `L-PICOLA` spectra are shown in blue. The bottom panel illustrates the fractional difference between the simulation spectra and the theoretical `CAMB` prediction. As shown in the figure, the `MUSIC` spectra accurately recover the theoretical prediction up to $k \leq 0.05\,h$ Mpc^{-1}. For that reason, we opted for the 2LPT approach.

The `COLA` method has been demonstrated to produce accurate results for smaller box sizes. However, the larger box size we adopted is necessary to avoid replications and spurious signals at large scales and to accommodate four non-intersecting light cones with a radius of $R_{\rm LC} = 3.6\,h^{-1}$Gpc.

The number of snapshots, on the other hand, was chosen based on three factors:

- **Smooth evolution of the potential:** We tested the potential evolution by varying the number of snapshots between 15 and 50 and found that 15 snapshots are sufficient for a smooth evolution, provided that the snapshots are linearly spaced in the scale factor a.

- **Memory consumption** Increasing the number of snapshots significantly increases the memory requirements of the `LIGER` code. Computing the potential on a 512^3 grid for 21 snapshots required an average memory usage of 200 GB.

- **Correct clustering amplitude**: Recovering the correct clustering amplitude for redshifts greater than 1 required increasing the number of snapshots for that redshift range.

Finally, the number of simulations produced was mainly constrained by the storage capabilities of the employed computational center.

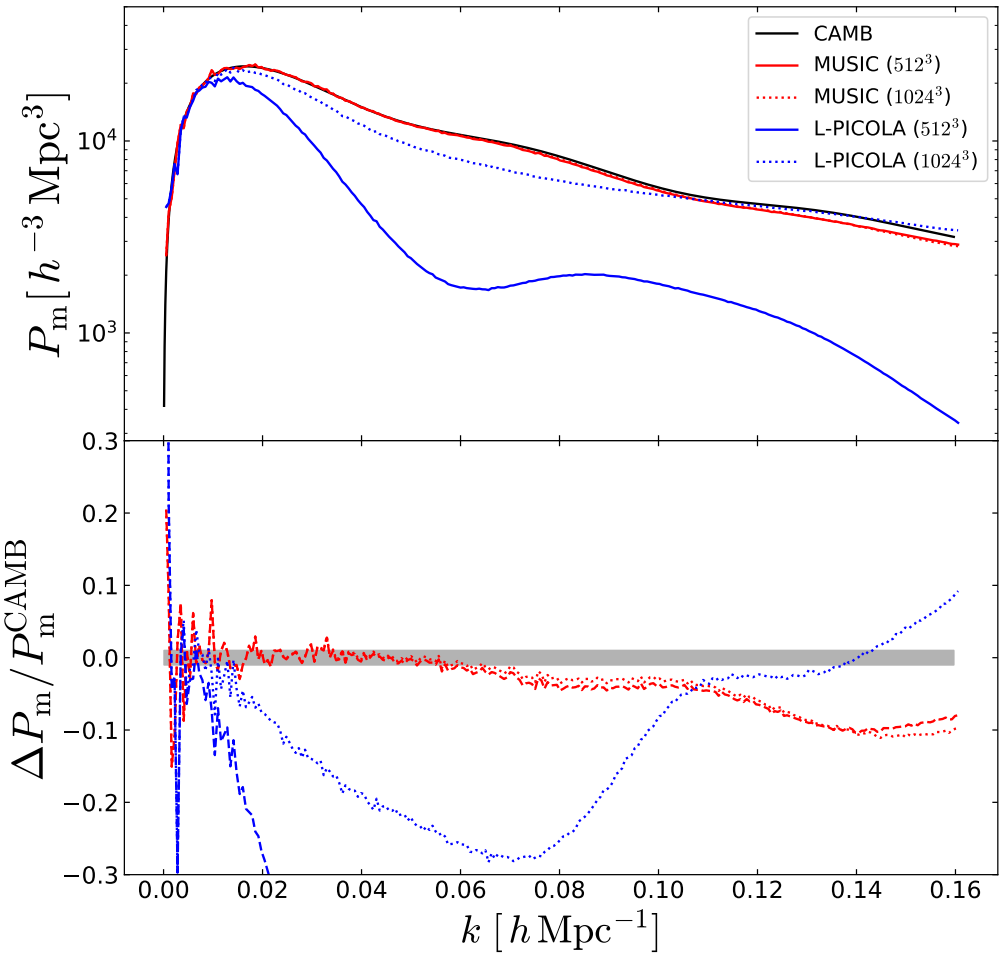


Figure 3.3: Comparison between the different codes tested with the theoretical CAMB spectra at $z = 0$. We show the spectra of the simulations produced using the MUSIC code (adopted in this work) in red, while the L-PICOLA spectra are shown in blue. The dashed lines represent the simulations with 1024^3 particles, while solid lines are computed from the 512^3 simulations. In the bottom panel, we plot the fractional difference with respect to CAMB.

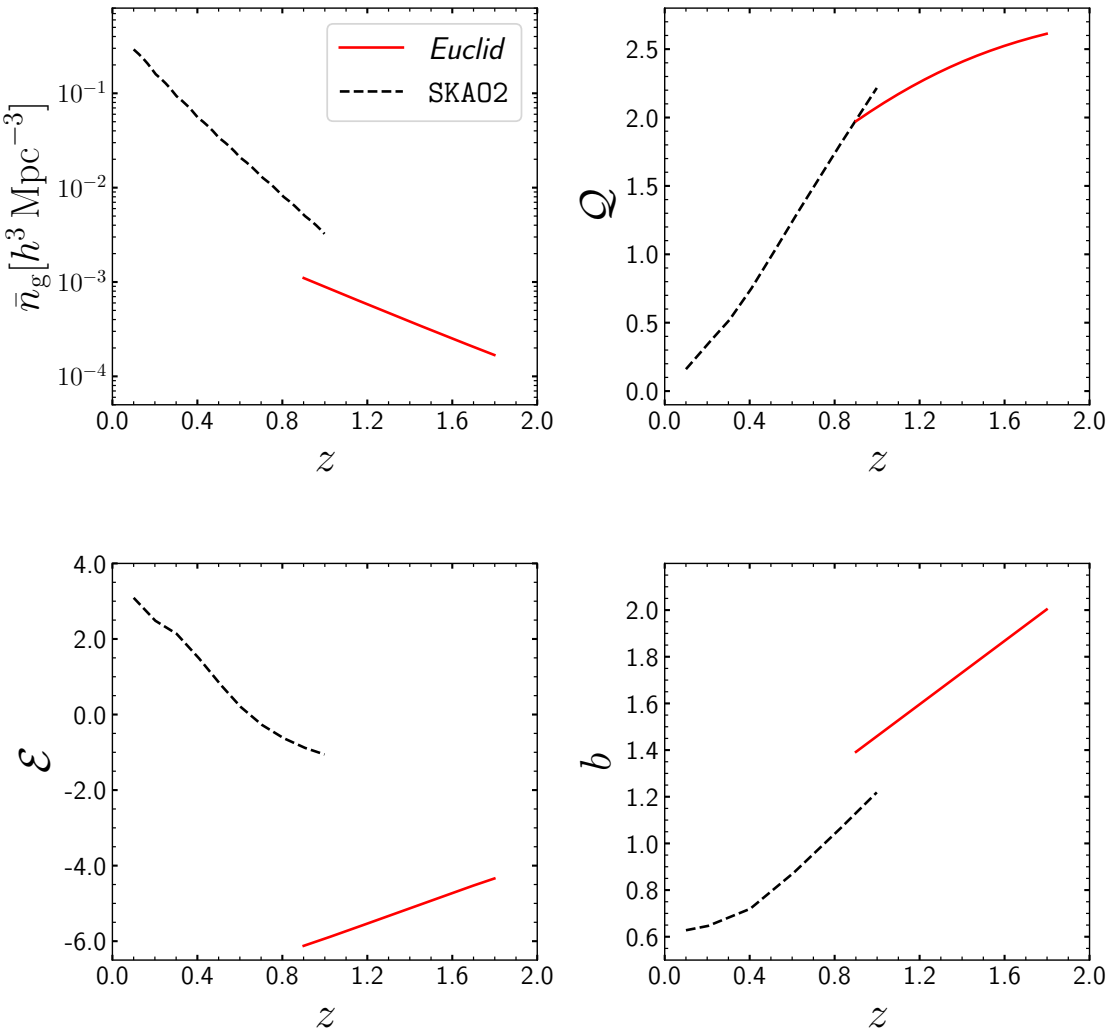


Figure 3.4: The survey functions $\bar{n}_g$ (top-left panel), $\mathcal{Q}$ (top-right), $\mathcal{E}$ (bottom-left), and b (bottom-right) used to build the mock galaxy surveys. Solid and dashed lines refer to the sample *Euclid* and SKA02 surveys, respectively. This figure is taken from Paper III.

3.3 Survey Functions

As explained in Sec. 3.1.1, we require four survey-specific functions – $b(z)$, $\mathcal{Q}(z)$, $\mathcal{E}(z)$ and $\bar{n}_g$, hereafter survey functions – to construct survey-tailored galaxy mock samples using Eq. (3.4). In this project, we build galaxy mocks that mimic two future surveys, the *Euclid* wide spectroscopic survey and the 21cm neutral hydrogen HI survey SKA02.

3.3.1 The *Euclid* Wide Spectroscopic Sample (Hα Surveys).

The primary objective of the recently launched *Euclid* space telescope (Laureijs et al., 2011, Racca et al., 2016) is to shed light on the nature of Dark energy and,

in turn, explain the accelerated expansion of the Universe. To achieve this, *Euclid* conducts an extensive survey spanning nearly 15,000 deg^2 of the extragalactic sky. This mission is designed to optimize the synergy between two cosmological methodologies: weak gravitational lensing and galaxy clustering. It leverages two key instruments: the Visual Imager (VIS, Euclid Collaboration: Borlaff et al., 2022), functioning within the 550 to 900 nm pass-band, captures high-resolution galaxy images crucial for cosmic shear measurements. Additionally, the Near-Infrared Spectrometer and Photometer (NISP, Maciaszek et al., 2022, Euclid Collaboration: Schirmer et al., 2022) performs imaging photometry, vital for estimating photometric redshifts, and slitless spectroscopy to precisely gauge the redshifts of Hα emission lines within the range $0.9 < z < 1.8$. Over a span of six years, the *Euclid* Wide Spectroscopic Survey aims to measure the redshifts and positions of nearly 30 million Hα emission-line galaxies, while its photometric counterpart intends to map the positions and morphologies of approximately 1.5 billion galaxies.

In this project, we build galactic mocks that mimic the radial and angular distribution of the Euclid Wide Spectroscopic Survey (hereafter simply the *Euclid* survey). To that end, we adopt model 3 from Pozzetti et al. (2016) to describe the redshift-dependent luminosity function of the *Euclid* target galaxies. This model is confirmed by a compilation of Hα line galaxies (Bagley et al., 2020).

The model defines the redshift dependent luminosity function $\Theta(L, z)$ for the Hα population to be

$$\Theta(L, z)\,\mathrm{d}L = 10^{-2.7}\left(\frac{L}{L_\star}\right)^{-1.4}\exp\left[\frac{-L}{L_\star(z)}\right]\frac{\mathrm{d}L}{L_\star(z)}\,, \tag{3.9}$$

where

$$\log_{10}\left[\frac{L_{\star,z}(z)}{42.6}\right] = -c\,(z-2.2)^2\,. \tag{3.10}$$

We use luminosity function defined in Eq. (3.9) as input for Eq. (2.72), assuming a flux limit of $F_{\rm lim} = 2\times10^{-16}$ erg cm^{-2} s^{-1}, to compute the average comoving density. As for the magnification bias, combining Eqs. (2.175) and (2.72), it is straightforward to see that

$$\mathcal{Q}(z) = -\frac{L_{\rm min}}{\bar{n}_{\rm g}(z)}\left.\frac{\partial\bar{n}_{\rm g}}{\partial L_{\rm min}}\right|_{L_{\rm min}=L_{\rm lim}(z)} = \frac{L_{\rm lim}(z)\,\Theta(L_{\rm lim}(z), z)}{\bar{n}_{\rm g}(z)}\,. \tag{3.11}$$

Similarly for the Evolution bias,

$$\mathcal{E}(z) = -\left.\frac{\partial\ln\int_{L_{\rm min}}^{\infty}\Theta(L,z)\,\mathrm{d}L}{\partial\ln(1+z)}\right|_{L_{\rm min}=L_{\rm lim}(z)} = -\frac{1}{\bar{n}_{\rm g}(z)}\int_{L_{\rm lim}(z)}^{\infty}\frac{\partial\Theta(L,z)}{\partial\ln(1+z)}\,\mathrm{d}L\,. \tag{3.12}$$

However, we actually compute the evolution bias through the relation

$$\mathcal{E}(z) = -\frac{\mathrm{d}\ln\bar{n}_\mathrm{g}(z)}{\mathrm{d}\ln(1+z)} - 2\,\mathcal{Q}(z)\,\left[1 + \frac{c\,(1+z)}{H(z)\,x(z)}\right] . \tag{3.13}$$

which can be derived as follows (Bertacca, 2015, Maartens et al., 2021, see also App B in Paper I)

$$\begin{aligned}\frac{\mathrm{d}\ln\bar{n}_\mathrm{g}(z)}{\mathrm{d}\ln(1+z)} &= \frac{1}{\bar{n}_\mathrm{g}(z)}\,\frac{\mathrm{d}\,\bar{n}_\mathrm{g}(z)}{\mathrm{d}\ln(1+z)} \\ &= -\frac{1}{\bar{n}_\mathrm{g}(z)}\,\left[-\Theta(L_\mathrm{lim}(z),z)\,\frac{\mathrm{d}L_\mathrm{lim}(z)}{\mathrm{d}\ln(1+z)} + \int_{L_\mathrm{lim}(z)}^{\infty}\frac{\partial\Theta(L,z)}{\partial\ln(1+z)}\,\mathrm{d}L\right] \\ &= 2\,\mathcal{Q}(z)\,\left[1 + \frac{c\,(1+z)}{H(z)\,x(z)}\right] + \mathcal{E}(z) . \end{aligned} \tag{3.14}$$

Finally, for the linear bias, we use the formula

$$b(z) = 1.46 + 0.68\,(z-1) , \tag{3.15}$$

which represents a linear fit using the values presented in Euclid Collaboration: Blanchard et al. (2020). The produced survey functions for the *Euclid* survey are shown as solid red lines in Fig. 3.4.

3.3.2 The SKAO2 Galaxy Sample (HI Surveys)

The primary objective of the SKAO (Square Kilometre Array Observatory) survey is to explore the fundamental nature of gravitational interactions and assess whether modifications to the prevailing theoretical framework, specifically General Relativity, are warranted (Bacon et al., 2020). The inaugural phase of the initiative is slated to execute a 10,000-hour observational campaign targeting HI galaxies within the redshift interval $0 < z < 0.5$, spanning an observational footprint of $5,000$ deg^2. This endeavour aims to detect approximately 3 million galaxies. To accomplish this, the survey will leverage the capabilities of a state-of-the-art mid-frequency array comprising 197 dish antennas.

Subsequently, a prospective "Phase 2" survey, known as SKAO2, is envisaged to extend over a comparable spatial expanse of 30,000 deg^2. This ambitious follow-up campaign is designed to probe a broader redshift range, targeting galaxies within the interval $0.1 < z < 2.0$. The anticipated outcome is the detection of a staggering 1 billion galaxies, heralding a new era of comprehensive cosmological exploration (Bull, 2016).

For our work, we produce full-sky mocks that follow the predicted survey functions of the SKAO2 survey in the redshift range $0.1 < z < 1.0$. We use a cubic spline to interpolate between the tabulated values computed in Maartens et al. (2021) for $n(z)$, $\mathcal{E}$ and $\mathcal{Q}$ which are derived by simulating HI galaxies in the requested redshift range. As for the linear bias, we adopt the values in (Yahya et al., 2015). The

resultant survey functions are shown in Fig. 3.4 as dashed black lines.

3.4 Galaxy Catalogs

3.4.1 Observed Galaxy Densities

In order to properly quantify the impact of distinct relativistic redshift distortions, we produce different versions of the observed galaxy density (using Eqs 3.4 and 3.8) that take into account an increasing number of relativistic effects. In other words, we apply Eqs. (2.158) and (2.170) cumulatively to produce multiple versions of the galactic mocks in redshift space. This cumulative application of the relativistic RSD aims to measure the relative importance of the source velocity term, the potential contributions (which are dominated by the weak lensing term) and the velocity of the observer terms separately. Consequently, we produce different types of mocks that contain varying degrees of relativistic RSD.

We produce five types of mocks for the three project using the *Euclid* survey functions

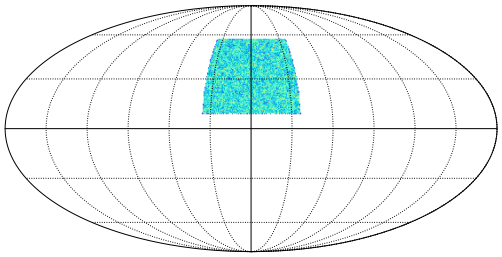
- Real space mocks ($\mathcal{R}$).
- Mocks incorporating all terms proportional to the peculiar velocity of source galaxies ($\boldsymbol{v}_{\rm e}$), denoted as $\mathcal{V}$. However, we set $\mathcal{Q}$ and $\mathcal{E}$ to zero in Eq. (3.4).
- Mocks including all relativistic RSD except observer velocity $\boldsymbol{v}_{\rm o}$ related terms ($\mathcal{G}$).
- Mocks incorporating all relativistic RSD, including the radial shift due to observer peculiar velocities. We produce two versions of these mocks
  - The $\mathcal{O}$ mocks are generated with observer velocity set to the Planck measurement (Planck Collaboration: Aghanim et al., 2020b) given by $v_\odot = 369.82 \pm 0.1\,\mathrm{km\ s^{-1}}$, in the direction of $(264.021 \pm 0.011^\circ, 48.253 \pm 0.005^\circ)$.
  - The $\mathcal{O}^{\rm R}$ mocks assign a random peculiar velocity to each observer.

Note that in $\mathcal{O}$ mocks, we have all relativistic RSD and are the closest representative to the observed galactic overdensity at linear order.

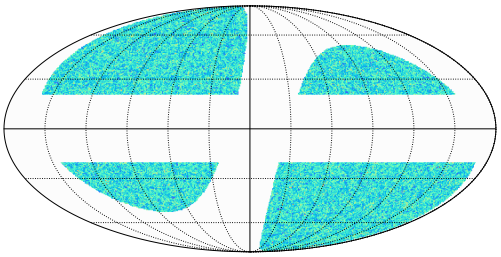
As for the SKAO2 survey, we build two sets of mocks that include the same RSDs as the $\mathcal{G}$ and the $\mathcal{O}$ mocks using, however, the SKAO2 survey functions.

3.4.2 Catalog Production

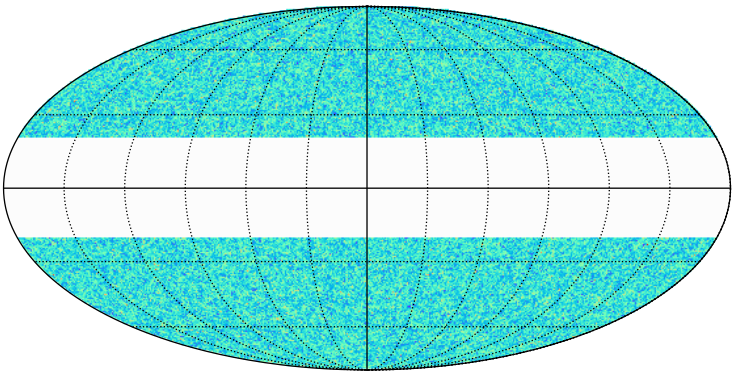
In this section, we outline how to produce galaxy and random catalogues out of the galactic density realizations constructed using Eq. (3.4).



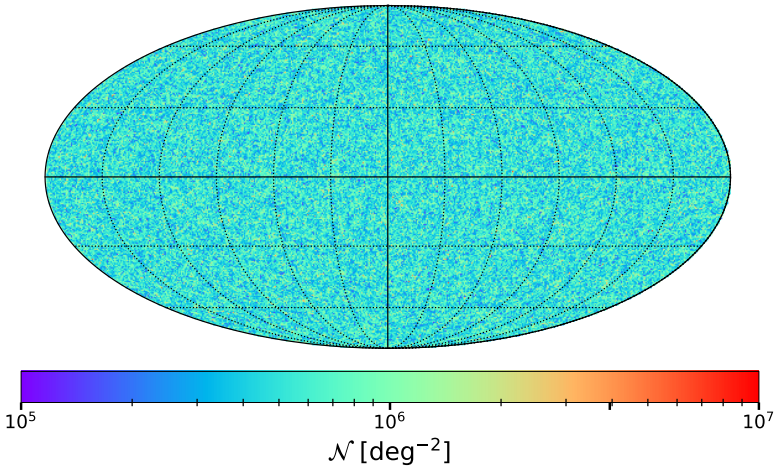
(a) Small patch with area 2565 deg^2.



(b) *Euclid* footprint.



(c) Galactic footprint.



(d) Full-sky footprint.

Figure 3.5: The four survey footprints utilized across various sections of this study shown in galactic coordinates.

Galaxy Catalog

In order to construct the galaxy catalogues from the galactic density of each type of mock, we implement the following procedure[5] :

- We first construct the number counts $N_{\rm g}(\boldsymbol{x}_j) = n_{\rm g}(\boldsymbol{x}_j)V_{\rm cell}$, where $V_{\rm cell} = (11.7)^3\ h^{-3}\,{\rm Mpc}^3$ is the volume of one cell of the grid used to sample the continuous fields (see Sec. 3.1.2).
- We add Poissonian shot noise to the LIGER mocks by replacing the derived number counts $N_{\rm g}(\boldsymbol{x})$ in each cell by an integer number $N_{\rm g}^{\rm S}(\boldsymbol{x})$ sampled from a Poisson distribution with a mean given by $N_{\rm g}(\boldsymbol{x})$ (see Paper I).
- Then, we assign $N_{\rm g}^{\rm S}(\boldsymbol{x})$ particles at each cell. Each particle is assigned a random position inside its respective cell.

In the end, we have a full-sky catalogue in the redshift range $0.9 < z < 1.8$ for the *Euclid* survey and $0.05 < z < 1.0$ for the SKAO2 surveys.

The clustering statistics analyzed in this study are influenced by the survey's geometry, specifically by the fraction of the sky coverage as well as the shape of the mask (e.g., Gibelyou & Huterer, 2012, Wilson et al., 2017, Paviot et al., 2022). We explore the impact of these masks in various sections of this thesis. To this end, we generate galactic catalogues using different survey masks, namely:

a) *Patch footprint*: This refers to a small rectangular patch spanning $2565\,{\rm deg}^2$. Its layout is displayed in Fig. 3.5a.

b) *Euclid footprint*: Here, we exclude regions within 20° of the Galactic and Ecliptic planes to match the expected angular distribution of the *Euclid* survey, illustrated in Fig. 3.5b. We note that this is an idealized version of the final *Euclid* footprint that ignores multiple angular observational systematics (Euclid Collaboration: Scaramella et al., 2022).

c) *Galactic footprint*: In this scenario, galaxies within a specified angular distance θ from the Galactic plane are removed, as visualized in Fig. 3.5c.

d) *Full-sky footprint*: This denotes a survey covering the entire celestial sphere, as depicted in Fig. 3.5d.

All footprints are utilized for the *Euclid* catalogues, while the SKAO2 mocks are exclusively generated using the full-sky footprint.

Random Catalogs

A random catalogue that matches the galaxy catalogue radial distribution without any inherent clustering signal is necessary for the power spectrum and 2PCF estimators (see Secs. 4.2 and 4.3). Let us now focus on the random catalogue

[5] we omit the subscript s from now on for brevity.

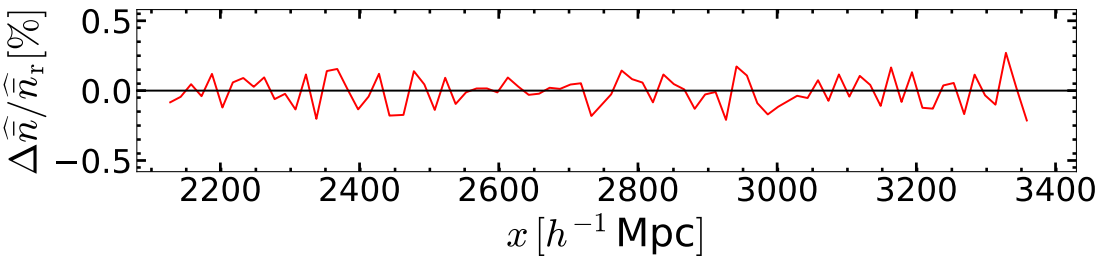


Figure 3.6: Fractional difference between the average density of the galaxy catalogue and the scaled random catalogue for one of the $\mathcal{V}$ mocks.

production. We create a random catalogue for each galaxy catalogue using a spline method (Ross et al., 2012). In this procedure, for each galaxy catalogue, we compute the average comoving number count $\bar{N}_{\rm g}(x)$, by averaging over the number of galaxies in radial shells with width $\delta x = 20\,h^{-1}\,\mathrm{Mpc}$. Then, we use $\bar{N}_{\rm g}(x)$ to get a cumulative probability distribution for the redshifts

$$\mathcal{P}(z) = \frac{\int_0^z \bar{N}_{\rm g}[x(z')]\mathrm{d}z'}{\int_0^\infty \bar{N}_{\rm g}[x(z')]\mathrm{d}z'}\,. \tag{3.16}$$

Finally, in order to produce the redshifts of the galaxies in the random catalogue, we sample from a cubic spline fit of the computed $\mathcal{P}(z)$ using an inverse transform method. As for the angular coordinates, we sample two random angles from uniform angular distributions, we only enforce that the galaxies lie within the footprint of the galaxy catalogue.

We repeat this procedure till the total number of galaxies in the random catalogue $\widehat{N}_{\rm r,T}$ is five times larger than the total number of galaxies, i.e., $\widehat{N}_{\rm r,T} = 5\widehat{N}_{\rm g,T}$. We limit the ratio between the two catalogues to five for three reasons. Firstly, a larger random catalogue would be computationally expensive in the estimation of the two-point correlation function multipoles. Secondly, our results are obtained by comparing the numerical estimates of the two-point statistics to each other, consequently, the size of the random catalogue has little impact on the significance we report. Thirdly, increasing the size of the random catalogue is mostly beneficial in pushing the shot-noise dominated regimes to smaller scales which is unimportant for our large-scale centred studies.

At the end of the catalogue production procedure, for each realization of each type of mock, we produce a galaxy catalogue consisting of $\widehat{N}_{\rm g,T}$ particles with a 3D density $\widehat{n}_{\rm g}(\boldsymbol{x})$, an angular averaged density $\bar{n}_{\rm g}(x)$ computed by averaging the galaxy catalogues in radial shells of size $\delta x = 20\,h^{-1}\,\mathrm{Mpc}$.[6] We also produce a corresponding random catalogue consisting of $\widehat{N}_{\rm r,T} = 5\,\widehat{N}_{\rm g,T}$ particles with a 3D density $\widehat{n}_{\rm r}(\boldsymbol{x})$. We show the fractional difference, $(5\bar{n}_{\rm g}/\bar{n}_{\rm r}) - 1$, between the average

[6]Computing the average density from the catalogue leads to the so-called Integral Constraint effect as the true average density does not match the estimated one (See Sec. 5.1, Peacock, 1991, Bahr-Kalus et al., 2021).

density of the galaxy catalogue and the scaled random catalogue for one of our mocks in Fig. 3.6. It is clear in the figure that we are able to reproduce the redshift behaviour of the galaxy catalogue using the spline method.

3.5 Validation of the LIGER Method

In this section, we discuss the validation tests carried out in Paper I, Paper III to validate the LIGER mocks by confirming their ability to capture the expected clustering properties derived from theoretical predictions. Our validation process involves comparing the angular power spectra, denoted as C_ℓ, estimated from our mocks with those computed using two independent Boltzmann solvers: CAMB (Lewis & Bridle, 2002) and CLASS (Blas et al., 2011). To measure the LIGER C_ℓ, we employ a pseudo angular power spectrum estimator, the details of which are discussed in Sec. 4.1. It's crucial to note that while CLASS and CAMB numerically integrate the perturbative expression for the power spectrum at leading order, the LIGER approach involves generating mock catalogues and subsequently utilizing them to estimate the angular power spectrum. Consequently, the LIGER spectra are affected by sample variance.

3.5.1 Validation with CAMB

In this section, we validate the LIGER method by comparing the angular power spectra estimated from the LIGER mocks to the theoretical prediction computed using the CAMB code. For simplicity, we perform the comparison for a full-sky survey and we average the LIGER angular power spectra over 140 full-sky mocks. We conduct the comparative analysis in two cases. First, we compare the angular power spectrum assuming only the intrinsic galaxy density fluctuations $\delta_{\rm g}^{\rm com}$ (the real-space case) to the $\mathcal{R}$ mocks. For the second case, we compare CAMB, adding all corrections, to the spectra obtained using the $\mathcal{G}$ mocks, which include all relativistic RSD except the observer velocity. In both cases, the mocks used follow the *Euclid* survey functions shown in Sec. 3.3. In that comparison, we input a density-weighted average of the b and $\mathcal{Q}$ functions into the CAMB code as the code takes in constant values for these quantities for each redshift bin.

We show the comparison in Fig. 3.7. In the top panel of the figure, we plot the mean and standard deviation of the $\mathcal{R}$ (empty red squares) and the $\mathcal{G}$ spectra (empty blue squares). The error bars denote the standard deviation computed from all 140 mocks. The corresponding CAMB theoretical predictions are marked by Xs (with the same colour scheme). In the bottom panel, we plot the mean and standard error (error on the mean) of the percent difference between the mocks' spectra and their theoretical counterpart. The red dots indicate the percentage difference in real space while the blue dots represent the percent difference in redshift space. As shown in the figure, the mock spectra remarkably agree within 1% with the CAMB results up to $\ell \approx 90$. Several factors contribute to this disparity. Firstly, at these scales and redshifts, nonlinear physics induces a slight power reduction

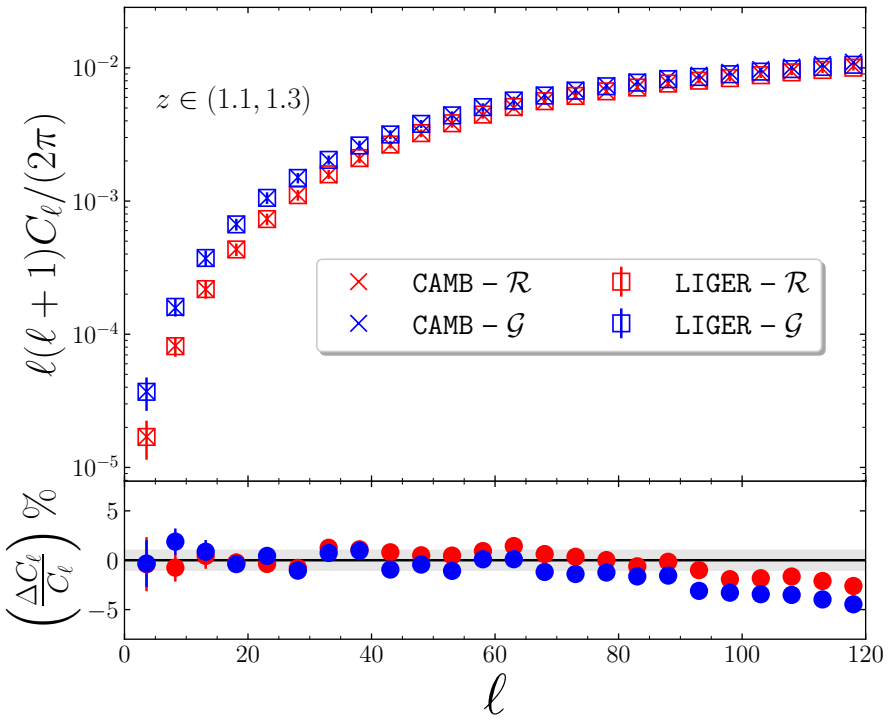


Figure 3.7: Top: the angular power spectra in $\Delta\ell = 5$ bins for both the LIGER mocks and the CAMB prediction. The red (blue) squares denote the mean of the spectra of the $\mathcal{R}$ ($\mathcal{G}$) mocks, while the error bars denote the standard deviation computed from all realizations. The CAMB predictions are shown using X markers with the same colour scheme as their mock counterparts.
Bottom: the fractional differences between the mock estimated angular power spectra and the CAMB produced spectra. The error bars in the bottom plot denote the error on the mean. The colour scheme matches the top plot. This figure is taken from Paper III.

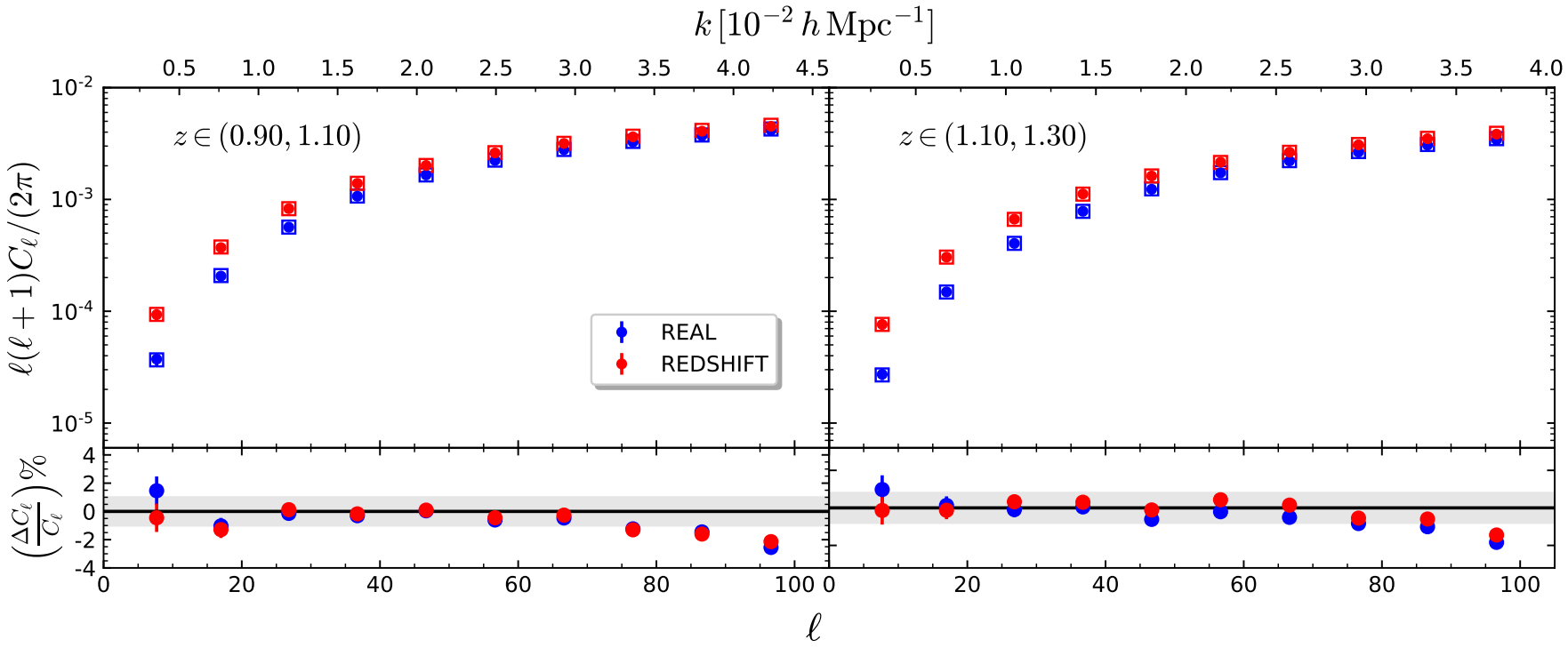


Figure 3.8: The angular power spectra, binned with $\Delta\ell = 10$, are plotted using `CLASS` (represented by empty squares) and derived from 140 mock catalogues produced with `liger` (indicated by solid stars). The spectra are presented for both real space (shown in blue) and redshift space incorporating all light-cone effects (displayed in red). The lower panels illustrate the fractional differences. Additionally, the top axis provides an approximate conversion of the multipole order ℓ to a comoving wavenumber k at the central redshift. This figure is taken from Paper I

relative to linear predictions (see e.g. Fig. 7 in Werner & Porciani, 2020). However, 2LPT fails to adequately capture this phenomenon, resulting in power spectra that deviate slightly from those obtained in actual N-body simulations (see e.g. figure 6 in Taruya et al., 2018). Lastly, the relatively coarse mesh employed in our extensive computational volumes leads to a spurious loss of power as the wavenumber k approaches the Nyquist limit $k_{\rm N} = \pi/\Delta x \simeq 0.268, h\,{\rm Mpc}^{-1}$ (see section 3 in Hahn & Abel, 2011). It is important to note that this limitation stems from the simulations utilized, rather than from `LIGER` itself.

Nevertheless, this study concentrates on the largest scales probed by forthcoming galaxy redshift surveys ($k < 0.02, h\,{\rm Mpc}^{-1}$), wherein `class` and `LIGER` exhibit remarkable accord. In other words, since we are interested in the impact of large-scale relativistic effects, the discrepancy at higher multipoles does not affect our results.

3.5.2 Validation with CLASS

We also compare our results against those obtained with the `CLASS` code. Similarly to the `CAMB` test, we compare the $\mathcal{R}$ and $\mathcal{G}$ mocks, however, we set the linear bias to 1 and the magnification bias to zero while creating the mocks for this test. We still consider a full-sky survey.

The outcomes pertaining to the redshift intervals $0.9 < z < 1.1$ and $1.1 < z < 1.3$ are depicted in Fig. 3.8, illustrating the angular power spectrum $\ell(\ell+1)C_\ell/(2\pi)$ against the multipole order $\ell < 100$. The top axis serves to approximate the comoving wavenumber k using the relation $k \simeq (\ell+1/2)/x(\bar{z})$, where $x(\bar{z})$ represents the comoving radial distance in the background cosmology and $\bar{z}$ denotes the central redshift within each bin.

Comparing the results from `CLASS` and `LIGER` reveals a convergence within a one percent margin for $\ell \lesssim 85$, while the `LIGER` spectra underestimate the power for higher multipoles. These differences can be attributed to the same reasons discussed in the previous subsection.

Nevertheless, this study concentrates on the largest scales probed by forthcoming galaxy redshift surveys ($k < 0.04, h\,{\rm Mpc}^{-1}$), where `CLASS`, `CAMB`, and `LIGER` exhibit remarkable agreement.

3.6 Summary

In this Chapter, we have discussed the production of 140 galactic mocks of the *Euclid* and `SKAO2` surveys accounting for relativistic RSD using the `LIGER` method. `LIGER` is a computational algorithm that takes in N-body simulations and produces galactic catalogues that match the observed distributions of a user-defined survey accounting for relativistic RSD at first order in cosmological perturbations. We have outlined both the theoretical details and computational implementation in Sec. 3.1. We then delineated the type of simulations employed in this work in Sec. 3.2. Subsequently, in Sec. 3.3, we showcased the two example surveys used

for the creation of mock catalogues; the *Euclid* and SKAO2 surveys. We discussed the types of galaxy catalogues we produce as well as their random counterpart in Sec. 3.4.2. Finally, we presented two validation tests that prove that the produced catalogues are able to recover the required two-point clustering for the purposes of this study in Sec. 3.5. In these tests, we found that the LIGER mocks are able to reproduce the galaxy clustering accounting for relativistic clustering with percent level accuracy up to the $\ell < 80$ in comparison to two different Boltzmann solvers.

Chapter 4

Estimators and Statistical Analysis

In this section, we detail the various estimators utilized in both this thesis and its accompanying paper. Across the different projects, we investigated the impact of relativistic RSD on three key two-point statistics: the angular power spectrum, the power spectrum multipoles, and the two-point correlation function. Here, we provide an overview of the estimators employed for each of these statistical analyses. Finally, in the last two sections, we discuss the statistical methods employed to assess the significance of the various relativistic RSD discussed in this thesis and estimate parameter constraints.

4.1 The Angular Power Spectrum

We estimate the angular power spectra for the `LIGER` validation tests (see Sec. 3.5) as well as the analysis of the angular power spectra of the *Euclid* survey (see Sec. 6.3). We compute the angular power spectrum in all these cases as follows. We first divide our galaxy catalogues (see Sec. 3.4.2) into multiple redshift bins, each marked by the index i. We then make use of the Hierarchical Equal Area isoLatitude Pixelation (`HEALPix` Górski et al., 2005) algorithm to partition the sky into $N_{\rm pix} = 12\,(1024)^2$ pixels. Following that step, we then simply count how many galaxies – with redshifts inside the bin i – lie within each map pixel `healpix` to create a count of angular numbers $N^i_{\rm g}(\Omega)$. The projected plots shown in Fig. 3.5 are built following that procedure. We are then able to compute the projected galaxy number density contrast on the sky $\Sigma^i_{\rm g}$ via

$$\Sigma^i_{\rm g} = \frac{\mathcal{N}^i_{\rm g}(\Omega)}{\bar{\mathcal{N}}^i_{\rm g}} - 1\,, \tag{4.1}$$

where $\bar{\mathcal{N}}^i_{\rm g}$ is the average galaxy number count.

After creating the projected density `healpix` maps, we decompose the projected

densities into spherical harmonics

$$a^i_{\ell m} = \int \Sigma^i_{\rm g}(\Omega) Y^*_{\ell m}(\Omega) {\rm d}\Omega \,, \tag{4.2}$$

and we are then able to compute the angular power spectrum using the pseudo C_ℓ estimator (PCL) which is defined as (Peebles, 1980)

$$\hat{C}_\ell^{\,ij} = \frac{1}{w_p^2(2\ell+1)} \sum_{m=-\ell}^{\ell} a^i_{\ell m} {a^j_{\ell m}}^* - \frac{1}{\bar{N}^i_{\rm g}} \,, \tag{4.3}$$

where $w_{\rm p}$ is a correction factor due to the finite pixelation of the sphere (see the `HEALPix` documentation for more details) and $f_{\rm sky}$ is defined as the fraction of the sky included in the survey footprint. In computing the previous two equations, we make use of the `anafast` algorithm within the `healpy` library (Zonca et al., 2019). We use that procedure in both Paper I and III in both cases using the *Euclid* survey mock catalogues. In particular for the *Euclid* footprint case (see Fig. 3.5b), the estimated $\hat{C}_\ell$ do not match the full-sky case as spherical-harmonics are not orthogonal for cut-skies. Therefore, the angular footprint mixes different modes (Peebles, 1973, Wandelt et al., 2001). It is possible to construct an estimate of the full-sky $\hat{C}_\ell$ from the cut-sky measurement using the mode-mode survey coupling matrix (Hivon et al., 2002). However, we do not adopt this approach as we only cross-compare the results of the different types of mocks (see Sec. 6.3).

4.2 The Two-point Correlation Function

In this subsection, we introduce the estimator utilised in evaluating the 2PCF function multipoles from the mock galaxy and random catalogues described in Sec. 3.4.2. The estimated 2PCF multipoles and their related results were used in Paper III as discussed in Sec. 6.4.

For the 2PCF estimator, we use the midpoint convention to define the 2PCF (see Sec. 2.2.2 for a discussion of the different conventions), i.e., the one adopted in the Euclid collaboration (Euclid Collaboration: De la Torre et al. in prep.). In that convention, we use the MP (see Fig. 2.8c) to estimate $\xi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ which defines the separation $\boldsymbol{r}$ and midpoint $\boldsymbol{x}_{\rm m}$ as

$$\boldsymbol{r} = \boldsymbol{x_2} - \boldsymbol{x_1} \quad , \quad \boldsymbol{x}_{\rm m} = \frac{\boldsymbol{x_2} + \boldsymbol{x_1}}{2} \,. \tag{4.4}$$

While the pair orientation angle φ is given by

$$\mu = \cos\varphi = \hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{x}}_{\boldsymbol{m}} \,. \tag{4.5}$$

With this choice of coordinate system, it is natural to redefine the 2PCF function as $\xi(r, x_{\rm m}, \mu) = \langle \delta_{\rm g}(\boldsymbol{x}_{\rm m} + \boldsymbol{r}/2) \delta_{\rm g}(\boldsymbol{x}_{\rm m} - \boldsymbol{r}/2) \rangle$.

We use the Euclid implementation (Euclid Collaboration: De la Torre et al. in prep.) of the Landy-Szalay (LS, Landy & Szalay, 1993) estimator to compute the 2PCF multipoles. The Landy-Szalay (LS) estimator is based on counting three sets of pairs: galaxy-galaxy DD, random-random RR and galaxy-random DR pair counts, at a given comoving separation r and angle φ (through μ) and in a redshift bin. From these three sets, we can obtain the normalized galaxy-galaxy ones

$$\mathrm{DD}(r,\mu) = \frac{DD(r,\mu)}{N_{\mathrm{g,T}}(N_{\mathrm{g,T}}-1)}\,, \tag{4.6}$$

the normalized random-random pair counts

$$\mathrm{RR}(r,\mu) = \frac{RR(r,\mu)}{N_{\mathrm{r,T}}(N_{\mathrm{r,T}}-1)}\,, \tag{4.7}$$

and the normalized galaxy-random pair counts

$$\mathrm{DR}(r,\mu) = \frac{DR(r,\mu)}{N_{\mathrm{r,T}}\,N_{\mathrm{g,T}}}\,. \tag{4.8}$$

Here we have denoted the normalized pair counts {DD, DR, RR} by Roman letters and the raw pair counts by italic letters $\{DD,\ DR,\ RR\}$. As the pair counts presented above are computed within the redshift bin, they implicitly average over the distance x_{m}. Then the LS estimator for the redshift bin can be constructed as

$$\hat{\xi}(r,\mu) = \frac{\mathrm{DD}(\mathrm{r},\mu) - 2\mathrm{DR}(\mathrm{r},\mu) + \mathrm{RR}(\mathrm{r},\mu)}{\mathrm{RR}(\mathrm{r},\mu)}. \tag{4.9}$$

Finally, we calculate its multipole expansion using

$$\hat{\xi}_\ell(r) = \frac{2\ell+1}{2}\int \hat{\xi}(r,\mu)\,\mathcal{L}_\ell(\mu)\mathrm{d}\mu\,. \tag{4.10}$$

4.3 The Power Spectrum Multipoles

The last estimator we discuss is the one employed to measure the power spectrum multipoles. We employ two slightly different estimators.

The Yamamoto–Bianchi Estimator.

We use the Yamamoto–Bianchi (YB) estimator (Bianchi et al., 2015, Scoccimarro, 2015) to estimate the power spectrum multipoles from the *Euclid* survey mocks in Paper III. The YB estimator computes the Fourier space correlation function of the Feldman–Kaiser–Peacock (FKP, Feldman et al., 1994) field defined as

$$F(\boldsymbol{x}) = \frac{w(\boldsymbol{x})}{\sqrt{A}}\left[\,\widehat{n}_{\mathrm{g}}(\boldsymbol{x}) - \alpha\,\widehat{n}_{\mathrm{r}}(\boldsymbol{x})\,\right], \tag{4.11}$$

where, in a given redshift bin, $\widehat{n}_{\rm g}(\boldsymbol{x})$ is the observed galaxy catalogue 3D density and $\widehat{n}_{\rm r}(\boldsymbol{x})$ is the corresponding random catalogue. For the weighting function $w(\boldsymbol{x})$, we use the standard FKP weights in order to minimize the variance in the measurements (Feldman et al., 1994)

$$w(\boldsymbol{x}) = \mathcal{I}(\boldsymbol{x})[1 + \overline{n}_{\rm g}(x)\mathcal{P}_0]^{-1} \,, \tag{4.12}$$

where $\mathcal{P}_0 = 2 \times 10^4 h^{-3}$ Mpc^3 and $\mathcal{I}(\boldsymbol{x})$ is an indicator function that is one inside the redshift bin and zero elsewhere. The α parameter accounts for the difference between the average number densities of the random and galaxy catalogues (see Sec. 3.4.2)

$$\alpha = \frac{\int w(\boldsymbol{x})\, \widehat{n}_{\rm g}(\boldsymbol{x}) \mathrm{d}^3x}{\int w(\boldsymbol{x})\, \widehat{n}_{\rm r}(\boldsymbol{x}) \mathrm{d}^3x} \,. \tag{4.13}$$

Finally, the normalization parameter A is defined as

$$A = \int w^2(\boldsymbol{x})\, \bar{n}^{\,2}{}_{\rm g}(x)\, \mathrm{d}^3x \,. \tag{4.14}$$

After constructing the field F, the power spectrum multipoles can be directly estimated using the Yamamoto estimator (Yamamoto et al., 2000)

$$\hat{P}_\ell(k) = (2\ell+1) \iiint \left[F(\boldsymbol{x}_1) F(\boldsymbol{x}_2) \mathcal{L}_\ell\left(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{x}}_{\rm m}\right) \right. \\ \left. \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot(\mathbf{x_1}-\mathbf{x_2})} \mathrm{d}^3x_1\, \mathrm{d}^3x_2\, \frac{\mathrm{d}\Omega_k}{4\pi} \right] - \hat{P}^{\rm SN}_\ell(k) \,, \tag{4.15}$$

where the shot noise contribution $\hat{P}^{\rm SN}_\ell(k)$ is given by

$$\hat{P}^{\rm SN}_\ell(k) = \frac{(1+\alpha)}{A} \int w^2(\boldsymbol{x})\, \widehat{n}_{\rm g}(\boldsymbol{x})\, \mathcal{L}_\ell(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{x}}_{\rm m})\, \mathrm{d}^3x \,. \tag{4.16}$$

However, computing Eq. (4.15) requires directly counting pairs which can be computationally demanding. A significant speed-up in the estimation of the power spectrum can be obtained by writing the estimator as a product of two Fourier Transforms. In that case, the Fast Fourier Transform (Fast Fourier Transform (FFT)) algorithm can be implemented, significantly reducing the computational overhead. However, this is only possible if the two spatial integrals can be separated in Eq. (4.15). The separation of the two integrals is prohibited by the $\mathcal{L}_\ell(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{x}}_{\rm m})$ term. In order to circumvent this issue, Yamamoto et al. (2006) replace the line of sight $\boldsymbol{x}_{\rm m}$ with the line of sight of one of the two galaxies within the integral, precisely, inside the Legendre polynomial (See also Beutler et al., 2014, Scoccimarro, 2015).

After replacing the argument of the Legendre polynomial, the estimator can be

recast as

$$\hat{P}_\ell(k) = (2\ell+1)\int F_\ell(\boldsymbol{k})F_0(-\boldsymbol{k})\frac{\mathrm{d}\Omega_k}{4\pi} - \hat{P}_\ell^{\mathrm{SN}}(k) \tag{4.17}$$

where

$$F_\ell(\boldsymbol{k}) \equiv \int F(\boldsymbol{x})\,\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\mathcal{L}_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}})\mathrm{d}^3x\,. \tag{4.18}$$

We note that Eqs. (4.15) and (4.17) define two different statistics which generate distinct outputs that only match at small scales or for relatively small surveys (e.g., Samushia et al., 2015, Castorina & White, 2018).

For Paper III, we use the implementation of the *Euclid* collaboration of Eq. (4.17). For specifics on the implementation of the FFTs and testing of the code, please refer to Euclid Collaboration: Salvalaggio et al. in preparation.

FKP ESTIMATOR

In Paper I and II, we use a different variation of the estimator given by Eq. (4.17) to measure the monopole of the power spectrum. Specifically, we implement and use Eq. (4.17) to compute the power spectrum monopole without using a random catalogue. In other words, the definition of the FKP field is given by

$$F_I\left(\boldsymbol{x}\right) = \frac{w\left(\boldsymbol{x}\right)}{\sqrt{A}}\left[\,\widehat{n}_{\mathrm{g}}\left(\boldsymbol{x}\right) - \,\bar{n}_{\mathrm{g}}(x)\,\right]. \tag{4.19}$$

We denote that variation by the FKP estimator to differentiate it from the definition which we call the YB estimator (see Eq. 4.11). We implement the FKP estimator and use it to measure the monopole of the power spectra of the *Euclid* mocks with multiple survey footprints and the `SKAO2` mocks with full-sky footprints.

In all three projects and for both estimators, we estimate the average comoving density $\bar{n}_{\mathrm{g}}$ by computing the angular average of the galaxy catalogue in shells of width $\delta x = 20h^{-1}\,\mathrm{Mpc}$. Similarly, we estimate the power spectra multipoles in an FFT box with a side length of $L_{\mathrm{FFT}} = 16015\,h^{-1}\,\mathrm{Mpc}$ on a grid with $N_{\mathrm{grid}} = 1024^3$ points with the only exception being the spectra employed in Paper III which employs a $N_{\mathrm{grid}} = 512^3$ grid. The size of the FFT box allows us to measure the power spectrum multipoles down to the fundamental frequency $k_{\mathrm{FFT}} = 4\times10^{-4}\,h\,\mathrm{Mpc}^{-1}$ which is imperative to this as relativistic effects are non-negligible at these scales (see Sec. 2.3).

4.4 Statistical Methods

4.4.1 Hypothesis Testing

One of the main objectives of this thesis is to measure the relative significance of relativistic RSD effects in the two-point statistics that will be measured in upcoming surveys. In this section, we outline the method used to quantitatively measure the signal-to-noise of each relativistic RSD for each clustering statistic S across all

projects. We outline the method using the mocks built to match the *Euclid* survey as an example.

Our objective is to ascertain the detectability of specific relativistic RSD within the *Euclid* survey. For instance, we can gauge the influence of the observer's peculiar velocity by contrasting the clustering statistic derived from the $\mathcal{O}$ and $\mathcal{G}$ mock catalogues. Similarly, by comparing the $\mathcal{O}$ and $\mathcal{G}$ light cones, we can quantify the significance of weak lensing contributions.

Let $\boldsymbol{D}_i^a$ denote the n-dimensional column data vector whose elements are the binned measurements of the clustering statistics estimated from the ith mock of type $a \in \{ \mathcal{R}, \mathcal{V}, \mathcal{G}, \mathcal{O} \}$. By taking the expectation value over all 140 realisations, we compute the mean signal and the covariance matrix of the noise in the measurements

$$\boldsymbol{\mu}_a = \mathrm{E}[\boldsymbol{D}_i^a] \,, \tag{4.20}$$

$$\mathsf{C}_a = \mathrm{E}[(\boldsymbol{D}_i^a - \boldsymbol{\mu}_a)(\boldsymbol{D}_i^a - \boldsymbol{\mu}_a)^{\mathrm{T}}] \,. \tag{4.21}$$

We can directly estimate the signal-to-noise (S/N) ratio for the detection of the specific RSD based on the Fisher information. Assuming that we have a b suite of realizations that include a specific RSD and an a suite that excludes the effect, the S/N is given by (e.g., Sec. 3.3.3 in Borzyszkowski et al., 2017)

$$(\mathrm{S/N})^2 = (\boldsymbol{\mu}_b - \boldsymbol{\mu}_a)^{\mathrm{T}} \, \mathsf{C}_b^{-1} \, (\boldsymbol{\mu}_b - \boldsymbol{\mu}_a) \equiv m_0 \,, \tag{4.22}$$

where we correct for the bias of the inverse covariance matrix due to the finite number of realisations used to estimate it by dividing the Precision matrix C_b^{-1} by $(N_{\mathrm{real}}-1)/(N_{\mathrm{real}}-n-2)$ where $N_{\mathrm{real}} = 140$ is the number of realizations (Kaufman, 1967, Hartlap et al., 2007).

Another possible avenue to quantify the detectability of RSDs is through classical hypothesis testing, in particular using the Likelihood ratio. Assuming Gaussian errors, the likelihood that the dataset $\boldsymbol{D}_i^b$ is drawn from a model[1] $\mathcal{M}_a$ with signal $\boldsymbol{\mu_a}$ is

$$\mathcal{L}(\mathcal{M}_a|\boldsymbol{D}_i^b) \propto \frac{\exp\left[-(\boldsymbol{D}_i^b - \boldsymbol{\mu_a})^{\mathrm{T}} \, \mathsf{C}_a^{-1} \, (\boldsymbol{D}_i^b - \boldsymbol{\mu_a})/2\right]}{(2\pi)^{n/2} \det \mathsf{C}_a} \,. \tag{4.23}$$

We can now define the null hypothesis $\mathcal{H}_0$ that the data vector $\boldsymbol{D}_i^b$ is a realisation of model $\mathcal{M}_a$ that excludes some RSD terms present in $\mathcal{M}_b$. We test this hypothesis against the alternative hypothesis $\mathcal{H}_1$ that the data are drawn from $\mathcal{M}_b$. The Neyman-Pearson lemma posits that the likelihood-ratio test statistic,

$$\lambda_i = 2 \ln \left[\frac{\mathcal{L}(\mathcal{M}_a|\boldsymbol{D}_i^b)}{\mathcal{L}(\mathcal{M}_b|\boldsymbol{D}_i^b)} \right] \,, \tag{4.24}$$

[1]The term 'model' is employed to signify a forecast for S based on predetermined parameter values representing the galaxy population and cosmological conditions. We do not vary these parameters at any point. Essentially, a model equates to an infinite suite of simulations, all incorporating the same RSD terms (such as the $\mathcal{G}$ simulations), and is characterized by its associated signal and noise covariance.

is the most powerful test for two hypotheses with fixed model parameters (in other words, simple hypotheses).

We can define the rejection criteria of the null hypothesis by building the probability distribution function (PDF) of λ given the null hypothesis which we denote by $P(\lambda_i|\mathcal{H}_0)$ as well as its alternate hypothesis counterpart, $P(\lambda_i|\mathcal{H}_1)$. The null hypothesis PDF, $P(\lambda_i|\mathcal{H}_0)$, should be a Gaussian distribution with mean $m_0 \geq 0$ and variance $s_0^2 = 4m_0$ (as proven in Appendix A of Borzyszkowski et al., 2017). On the other hand, under $\mathcal{H}_1$, $P(\lambda_i|\mathcal{H}_1)$ is the Gaussian distribution with mean $m = -(\mu_b - \mu_a)^{\mathrm{T}}\mathsf{C}_a^{-1}(\mu_b - \mu_a) \leq 0$ and variance $s^2 = 4|m|$. Assuming RSDs have little impact on the Covariances, it is straightforward to see that $m = -m_0$ and $s = s_0 = 4m_0$. Consequently, the distance between the two means m and m_0 of PDF distributions in units of the standard deviation is $(m_0 - m)/(2\sqrt{m_0}) = \sqrt{m_0}$ which matches exactly the Fisher information (see Eq. 4.22).

Equipped with the two distributions, we can define the rejection criteria as follows. For a confidence level α, we reject the null hypothesis if $\lambda_i < \omega_\alpha$ where ω_α is a real number in the range $\alpha = [0, 1]$ that satisfies $P(\lambda_i < \omega_\alpha|\mathcal{H}_0) = 1 - \alpha$. As $P(\lambda|\mathcal{H}_0)$ is a Gaussian, its percentiles can be computed using the error function (erf). Analytically, $\mathcal{H}_0$ is rejected for the fraction

$$f_\alpha = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\omega_\alpha - m}{2\sqrt{2|m|}}\right)\right] \tag{4.25}$$

of the realisations.

Using the previous relation, we can understand what different values of the S/N computed using Eq. (4.22) mean in terms of the rejected fraction of realizations. For example, if we set $\alpha = 0.95$, we find that $\omega_{95} = m_0 - 3.29\sqrt{m_0}$ and then we directly find

$$f_{95} = f_\alpha = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\sqrt{m_0} - 1.645}{\sqrt{2}}\right)\right]. \tag{4.26}$$

This means that higher values of the SNR correspond to a higher fraction of realizations being rejected. For instance:

- An S/N = 1 corresponds to around a quarter of the realizations being rejected ($f_{95} = 0.259$),
- An S/N = 1.645 corresponds to around half of the realizations being rejected ($f_{95} = 0.5$),
- An S/N = 2.93 results in around 90% of the realizations being rejected.

In order to perform these tests, we need to derive the PDFs s of λ under $\mathcal{H}_0$ and $\mathcal{H}_1$. For the projects reported in this thesis, we use two variations of the likelihood ratio test. Firstly, the '*Double Likelihood ratio*' (DLR) test. In that version, we build $P(\lambda) = P(\lambda|\mathcal{H}_1)$ by[2] directly computing the normalized histogram of λ_i using the

[2]Note that the equality is because the data vector $\boldsymbol{D}_i^a$ is a realization of the model a.

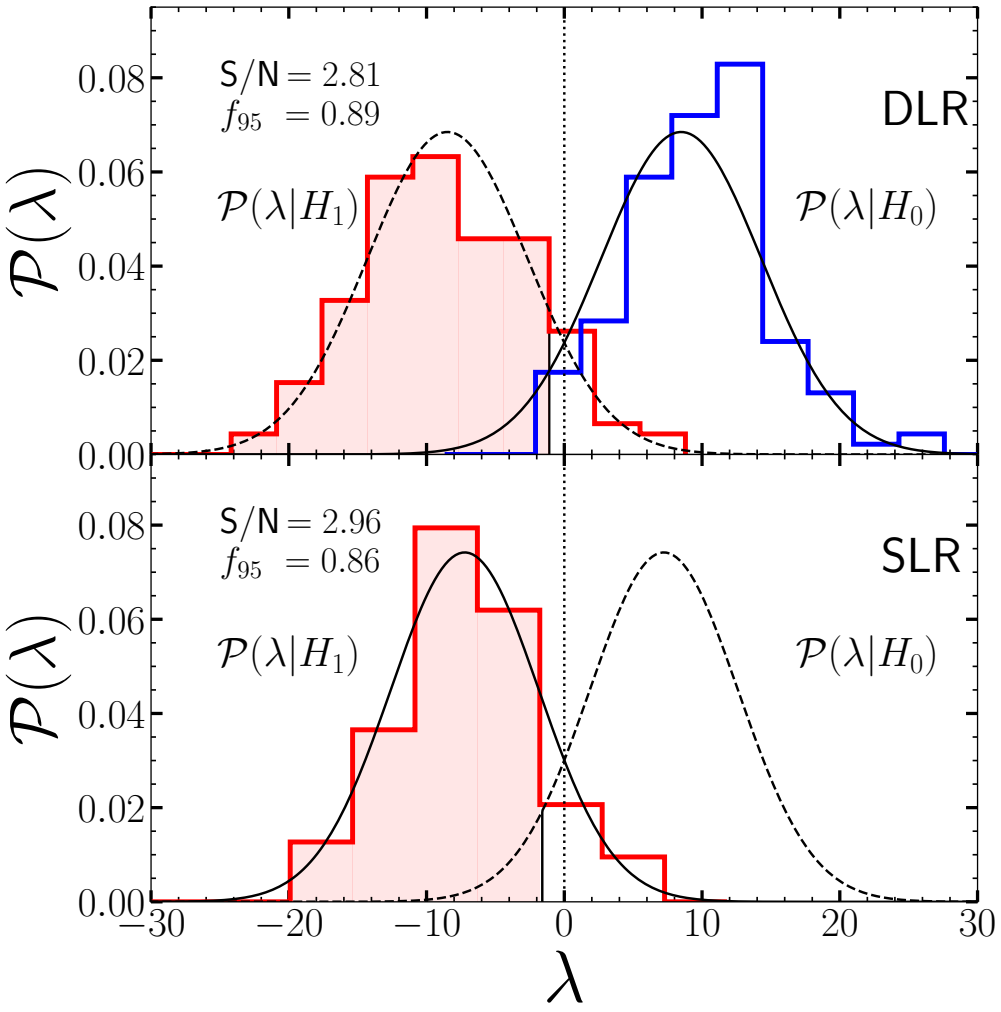


Figure 4.1: An illustrative example of the statistical tests outlined in Sec. 4.4.1. In the top panel, we show the example of the '*Double Likelihood ratio*' (DLR) test. The blue histogram on the right-hand side displays the distribution of the likelihood-ratio test statistic $\tilde{\lambda}$ evaluated from the a mocks that exclude some relativistic RSDs (i.e., under the null hypothesis $\mathcal{H}_0$). The red histogram on the left-hand side, instead, displays the distribution of λ in the b mock catalogues that include the aforementioned effects (i.e. under the alternative hypothesis $\mathcal{H}_1$). The solid and dashed curves display the Gaussian models for the histograms. The S/N is computed via Eq. (4.28) and it is a measure of the weighted separation between the two histograms in units of their standard deviations. The shaded region shows the realisations in which the null hypothesis is rejected at the 95% confidence level. The bottom panel shows the '*Single Likelihood ratio*' (SLR) test where the S/N is computed using only one suite of mocks (see Eq. 4.29). And only the covariance matrix C_b of the more complex model is used. This plot is based on a similar figure in Paper III.

data set that includes the specific RSD effect (the complex model), i.e we compute Eq. (4.24) for each realization. As for the simple hypothesis PDF , $P(\lambda|\mathcal{H}_0)$, we compute the histogram of the $\tilde{\lambda}_i$

$$\tilde{\lambda}_i = 2\ln\left[\frac{\mathcal{L}(\mathcal{M}_a|\boldsymbol{D}_i^a)}{\mathcal{L}(\mathcal{M}_b|\boldsymbol{D}_i^a)}\right] , \tag{4.27}$$

for all realizations. Where by construction, $P(\tilde{\lambda}) \equiv P(\lambda|\mathcal{H}_0)$. We then numerically evaluate the means of the PDFs s, $\hat{m}$ and $\hat{m}_0$ for $P(\lambda|\mathcal{H}_1)$ and $P(\lambda|\mathcal{H}_0)$ respectively. We then define the signal-to-noise ratio as

$$S/N = (\hat{m}_0 - \hat{m}_1)/\hat{s}_0 . \tag{4.28}$$

We also use the PDF $P(\tilde{\lambda})$ to numerically compute ω_{95} and then directly count the fraction of realizations with $\lambda_i < \omega_{95}$ to compute f_{95}. An example of that procedure is shown in the bottom panel of Fig. 4.1. We adopt the following notation for all comparisons using this test: (b-a) DLR. For example, if we apply the DLR to test to compare the $\mathcal{O}$ and $\mathcal{G}$ mocks, we write that we perform the ($\mathcal{O}$-$\mathcal{G}$) DLR test to measure the S/N and f_{95}.

The second test we apply, the '*Single Likelihood ratio*' (SLR) avoids the computation of $\tilde{\lambda}$ altogether. We only compute histograms of the values of λ_i calculated from the 140 b mock catalogues. In that case, we only use the covariance of the complex model b, C_b. This operation is similar to what is done in the analysis of actual surveys, where the covariance is estimated from suites of mock catalogues and kept the same for different theoretical models.

In the SLR test, we compute the signal-to-noise ratio as

$$S/N = 2\hat{m}_0/\hat{s}_0 , \tag{4.29}$$

and determine f_{95} by counting the number of realisations in which $\lambda_i < \omega_{95}$, where ω_{95} is computed from the theoretical Gaussian distribution . We show an example of both likelihood ratio tests in Fig. 4.1. For example, if we apply the SLR to compare the $\mathcal{O}$ and $\mathcal{G}$ mocks, the test is denoted by ($\mathcal{O}$-$\mathcal{G}$) SLR.

4.4.2 MCMC Analysis

In our pursuit of the second objective of this thesis, the measurement of cosmological parameters using relativistic RSD, we use the Markov chain Monte Carlo (MCMC) (Metropolis et al., 1953) algorithm to numerically sample the probability distribution of the parameters of interest, Σ, given a data vector,$\boldsymbol{d}$, denoted by $\mathcal{P}(\Sigma|\boldsymbol{d})$.

The method can be described as follows. According to Bayes's theorem, the PDF of interest is probational to the PDF of the data given the model (the likelihood) multiplied by the prior on the cosmological parameters, i.e.,

$$\mathcal{P}(\Sigma|\boldsymbol{d}) \propto \mathcal{L}(\boldsymbol{m}(\Sigma))\,\Pi(\Sigma) , \tag{4.30}$$

where $\boldsymbol{m}$ is the model of the data vector $\boldsymbol{d}$.

The MCMC algorithm works as follows. It first picks random values for the parameters within the ranges allowed by the Prior PDF. It then proceeds by iteratively exploring the parameter space, proposing new values based on a chosen proposal distribution determined by both the prior and the current state of the chain, and evaluating their likelihood given the observed data. It then computes the ratio of the likelihood of the proposed values to the likelihood of the current values and compares it to a randomly generated number between $[0, 1]$ to determine an acceptance probability. This probability informs the algorithm whether to accept the proposed values and update the chain or reject them and retain the current state. The algorithm iterates until convergence is achieved, at which point the samples form a representative sample of the posterior distribution $P(\Sigma|\boldsymbol{d})$. This distribution provides estimates of parameter uncertainty and enables parameter inference.

The procedure delineated above requires three inputs.

- The data vector $\boldsymbol{d}$.
- The parameters of interest Σ.
- The model m, its likelihood $\mathcal{L}$ and their related functions (e.g. the model covariance).

Numerous techniques and methods have been developed to optimize the efficiency and effectiveness of MCMC algorithms for various purposes. In our work, we employ a MCMC method known for its robustness against affine transformations of the parameter space. Initially introduced in Goodman & Weare (2010), this method has been implemented in the `emcee` code (Foreman-Mackey et al., 2013), which we utilize for our analyses.

Convergence

We adhere to the approach outlined by Foreman-Mackey et al. (2013) in utilizing the integrated autocorrelation time (τ) to assess the convergence of our MCMC chains. Essentially, τ provides an estimate of the number of steps required for the chain to forget its initial state. In all our analyses, we terminate the chains once the number of steps exceeds $200 \times \tau$. For further insights into computing the integrated autocorrelation time and its implications for chain convergence, we recommend referring to the emcee documentation.

4.5 Summary

In this Chapter, we summarize the survey estimators employed throughout this work. Starting with the angular power spectrum, we described the numerical details of the PCL estimator in Sec. 4.1. In Sec. 4.2, we then discuss the pair-counting method implemented in the LS estimator, which we employ to measure

the 2PCF multipoles. We then turn to the two estimators employed to measure the power spectrum multipoles in Sec. 4.3; Namely the YB and Feldman–Kaiser–Peacock (FKP) estimators. Finally, in Sec. 4.4, we discuss the statistical test we employ to quantitatively evaluate the significance of relativistic RSD effects and to extract parameter constraints from the measured summary statistics.

Chapter 5

Wide-angle Effects in the *Euclid* Survey

In this chapter, we present the findings from various tests conducted using the $\mathcal{V}$ mocks. Firstly, we examine the comparison between the power spectrum monopole and the Kaiser model within the Global Plane-Parallel approximation, as detailed in Paper I. Secondly, we contrast the multipoles of the *Euclid* power spectrum with those of the local plane parallel approximation, drawing from insights in Paper III. Our approach begins with a description of each model, followed by a comprehensive discussion of the comparisons.

5.1 The Observed Power Spectrum

We are only able to observe a limited volume of the Universe. Consequently, the power spectrum multipoles estimated from Eq. (4.17) are impacted by the survey window function which defines the physical volume of the survey. In this section, we study that impact and we relate the results of the estimator defined in Eq. (4.17) to the theoretical quantity defined in Eq. (2.98).

To that end, we write the observed field $F(\boldsymbol{x})$ in Eq. (4.11) as the product of the true density and a selection function that defines the properties of the survey window (Peacock, 1991, Beutler et al., 2014, de Mattia & Ruhlmann-Kleider, 2019):

$$F(\boldsymbol{x}) = W(\boldsymbol{x})\,\delta_{\rm g}(\boldsymbol{x}) - W(\boldsymbol{x})\frac{\int W(\boldsymbol{x}')\,\delta_{\rm g}(\boldsymbol{x}')\,{\rm d}^3x'}{\int W(\boldsymbol{x}')\,{\rm d}^3x'}\,, \tag{5.1}$$

where the window function $W(\boldsymbol{x})$ is given by

$$W(\boldsymbol{x}) = \frac{w(\boldsymbol{x})\;\bar{n}_{\rm g}(x)}{\sqrt{A}}\,, \tag{5.2}$$

where w and A are the FKP weights and normalization factor defined in Sec. 4.3.

The first term in Eq. (5.1) leads to the dependence of the observed power spec-

trum on the window function. The second term corresponds to the integral constraint. It ensures that the $F(\boldsymbol{x})$ averages to zero over the survey volume. As seen in Sec. 4.3, we compute the average density from the observed catalogue as is usually done in realistic surveys. That procedure implicitly assumes that the average density within the survey is equal to the 'true' average density of the Universe. However, in a realistic survey, the true value of the mean density $\bar{n}_{\rm g}$ is unknown and the geometry of the survey can be irregular.

We are then able to rewrite the estimated power spectra (Eq. 4.15) in terms of the true galaxy field as

$$P_\ell(k) = \int W(\boldsymbol{x}_1)W(\boldsymbol{x}_2)\,\langle\delta_{\rm g}(\boldsymbol{x}_1)\,\delta_{\rm g}(\boldsymbol{x}_2)\rangle\,e^{-{\rm i}\boldsymbol{k}\cdot(\boldsymbol{x}_2-\boldsymbol{x}_1)} \times \mathcal{L}_\ell\left(\hat{\boldsymbol{k}}\cdot\boldsymbol{x}_{\rm m}\right)\,{\rm d}^3x_1\,{\rm d}^3x_2\,\frac{{\rm d}\Omega_k}{4\pi} + P_{\rm IC}\,, \tag{5.3}$$

where $P_{\rm IC}$ is the contribution of correlating the second term of Eq. (5.1) and the brackets denotes the ensemble average. Let us articulate that the line-of-sight dependence in the previous relation is not only found in the $\mathcal{L}_\ell\left(\hat{\boldsymbol{k}}\cdot\boldsymbol{x}_{\rm m}\right)$ but is also present in $\delta_{\rm g}$ due to RSD as seen in Eq. (2.181).

There are multiple approximations used in the literature in computing Eq. 5.3. They are

- **The distant observer approximation**: It assumes that the radial width of the volume Δx is much smaller than the radial distance x, i.e. $\Delta x << x$. That assumption mostly pertains to the modelling of the theoretical power spectrum and is the basis of Eq. (2.95).

- **Global Plane-Parallel (GPP)**: It states that the galaxies are so far away from the observer that their lines of sight are constant and parallel, i.e. $\hat{\boldsymbol{x}}_{\rm m} \approx \hat{\boldsymbol{x}}_{\rm c}$ where $\hat{\boldsymbol{x}}_{\rm c}$ is a constant direction. The GPP is assumed in the derivation of (2.94).

- **Local Plane Parallel (LPP)**: It replaces the line of sight $\hat{\boldsymbol{x}}_{\rm m}$ by the line of sight to one of the position vectors, i.e. $\hat{\boldsymbol{x}}_{\rm m} \approx \hat{\boldsymbol{x}}_1$. The YB estimator employs that approximation as seen in Eq. (4.17).

In the following sections, we present how the theoretical model of the observed power spectrum is affected by the different approximations as well as how each of the GPP and LPP models compare to the spectra of the $\mathcal{V}$ mocks. Note that we only include the Kaiser velocity gradient term (see Eq. 2.181) in both the $\mathcal{V}$ mocks and the theoretical models discussed in this section.

5.2 The Global Plane-parallel Approximation

Applying the GPP approximation both with the $\delta_{\rm g}$ and $\mathcal{L}_\ell\left(\hat{\boldsymbol{k}}\cdot\boldsymbol{x}_{\rm m}\right)$ terms in Eq. (5.3), it can be rewritten as (Peacock, 1991, Beutler et al., 2014)

$$P^{\rm obs}(\boldsymbol{k}) = \int P_{\rm g}(\boldsymbol{k}')|W(\boldsymbol{k}-\boldsymbol{k}')|^2\,{\rm d}^3k' - \frac{|W(\boldsymbol{k})|^2}{|W(0)|^2}\int P_{\rm g}(\boldsymbol{k}')|W(\boldsymbol{k}')|^2\,{\rm d}^3k'\,, \tag{5.4}$$

where $P_{\rm g}(\boldsymbol{k}')$ is the theoretical power spectrum defined in Eq. (2.95) and $|W(\boldsymbol{k})|^2$ is the power spectrum of the window function defined in Eq. (5.2), which we denote by $P^W(\boldsymbol{k})$ hereafter.

The second term of Eq. (5.4) is the integral constraint $P_{\rm IC}$, it enforces that the observed $P^{\rm obs}(\boldsymbol{k})$ goes to zero at $k=0$. A vanishing monopole is implemented in the power spectrum estimator (see Eq. 4.17) by construction. However, the integral constraint leads to a negative bias in the estimated power spectrum due to the window function mixing between the different k modes. In Paper I, we concentrate on the monopole of the power spectrum. To that end, we write Eq. (5.4) for the monopole:

$$P_0^{\rm obs}(\boldsymbol{k}) = \int P_0(k')|W(\boldsymbol{k}-\boldsymbol{k}')|^2\,{\rm d}^3k' - \frac{|W(\boldsymbol{k})|^2}{|W(0)|^2}\int P_0(k')|W(\boldsymbol{k}')|^2\,{\rm d}^3k'\,, \tag{5.5}$$

where P_0 is the monopole of the theoretical power spectrum (see Eq. 2.97). We denote the result of the previous Equation by the GPP Kaiser model.

We can understand the impact of the survey window through Eq. (5.5). Starting with the first term on the RHS, it is clear that the estimator mixes distinct Fourier modes and computes a weighted average of the true power spectrum. That mixing is most prominent at the maxima of the weight function $|W(\mathbf{k})|^2$. Typically, the magnitude squared of the weight function $|W(\mathbf{k})|^2$ exhibits a sequence of peaks that diminish in amplitude as the wave vector k increases. The primary peak is centred around $k=0$ and possesses a width of approximately $k_W \simeq V^{-1/3}$, where V is the volume under consideration.

5.2.1 Systematic and Statistical Errors in the FKP Estimator

Before comparing the monopole spectra measured from the $\mathcal{V}$ mocks and the results of Eq. (5.5). Let us briefly discuss the systematic errors (the integral constraint) and the statistical errors (the standard deviation) of the power spectrum measurements using the FKP estimator (see Sec. 4.3).

Variance

For Gaussian density fluctuations at scales $\Delta k \gg k_W$, the covariance matrix of FKP estimates is diagonal with the variance given by (e.g. Feldman et al., 1994,

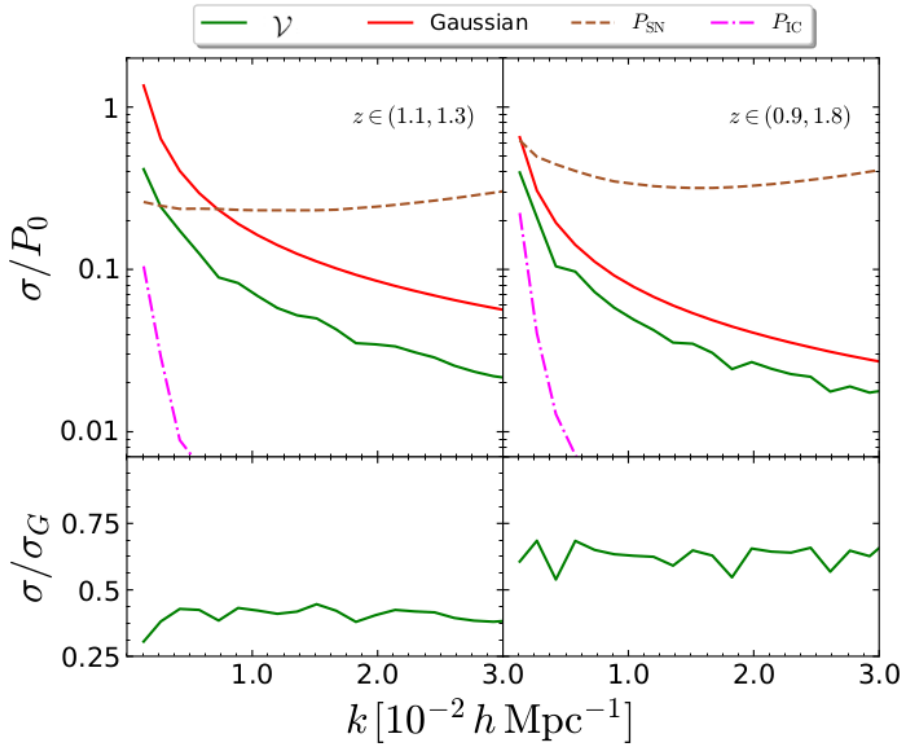


Figure 5.1: Top: the green solid line shows the standard deviation (RMS scatter) of the FKP estimator for the *Euclid* footprint in two redshift bins. We compute the scatter by averaging over all 140 estimated power spectra. We also plot in red the Gaussian prediction computed using Eq. (5.6). The systematic shot noise contribution, $P_{\rm SN}$, is shown as a brown dashed line. Finally, we show a numerical estimate of the systematic error generated by the integral constraint, $P_{\rm IC}$. Bottom: the ratios between the numerical RMS scatter and the Gaussian prediction (the green and red lines in the top plot) are shown in green solid lines. This Figure is based on a similar plot in Paper I

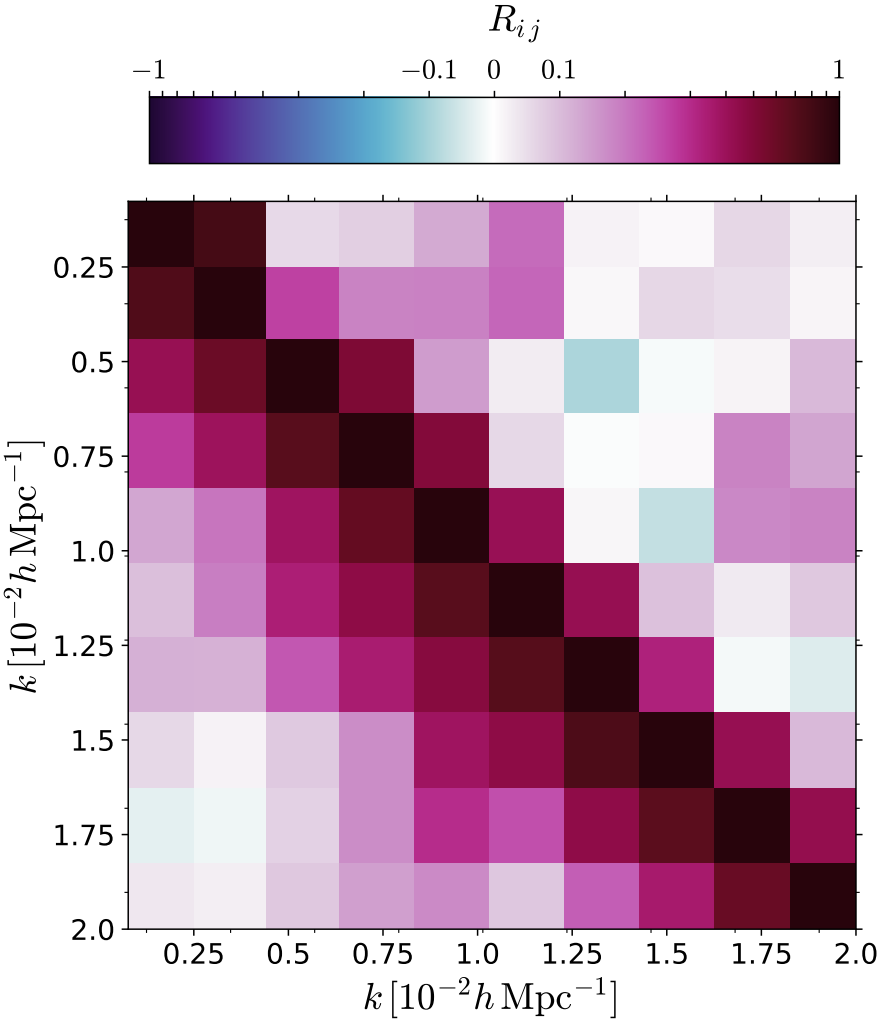


Figure 5.2: We plot the Correlation matrix **R** computed from the monopole of the $\mathcal{V}$ spectra in the redshift bin $(1.1, 1.3)$ (lower triangular part) and $(0.9, 1.8)$ (upper triangular part). Note that the latter has smaller off-diagonal elements due to the lower k_W.

Tegmark et al., 1998b)

$$\frac{\sigma_{\rm G}^2(k)}{P_0^2(k)} \simeq \frac{2\,(2\pi)^3}{V_{\rm eff}(k)\,V_k}, \tag{5.6}$$

where

$$V_{\rm eff}(k) = \int_{\mathcal{S}} \left[\frac{\bar{n}(\boldsymbol{x})\,P_0(k)}{1+\bar{n}(\boldsymbol{x})\,P_0(k)}\right]^2 {\rm d}^3x \tag{5.7}$$

denotes the effective volume and V_k is the volume of the k-shell with width Δk (usually computed through $V_k \simeq 4\pi k^2 \Delta k$). The previous relation ignores contributions to the covariance matrix that arise from the connected trispectrum of the galaxy distribution convolved with the survey window function as they are significant at smaller scales (e.g. Chan & Blot, 2017).

We compare the relative RMS scatter $\sigma/\langle\hat{P}_0\rangle$ computed from the $\mathcal{V}$ mocks to the Gaussian result in Fig. 5.1. In the figure, we show the numerical estimate from the $\mathcal{V}$ spectra in green and the theoretical Gaussian prediction in red for two redshift bins; a relatively thin bin and a bin that encompasses the *Euclid* redshift range. It is clear from Fig. 5.1 that the variance is significantly reduced in the case of the thin redshift bin.

That effect can be understood by comparing the two panels of the Figure. For small redshift bins, k and Δk are comparable with k_W, the window function thus mixes the power between the nearby wavenumbers, effectively reducing the variance of the FKP estimates with respect to the Gaussian approximation and simultaneously increasing the covariance among nearby k-bins. This impact of the window function is reduced as we employ thicker redshift bins as apparent in the right panel of Fig. 5.1.

To visualize the impact on the covariance, we show the correlation matrix of $\hat{P}_0(k)$ for the $\mathcal{V}$ mocks in Fig. 5.2. We plot the correlation matrix for both redshift bins in the lower and upper triangular parts, respectively. We immediately see that the off-diagonal elements are larger in the thinner redshift bin.

In Fig. 5.1, we also plot the shot noise contribution $P_{\rm SN}$ (brown dashed line) to show that in the k-ranges considered in this work, the shot-noise contribution is either larger or comparable to the statistical errors, however it is always smaller than the signal, i.e. $\sigma \leq P_{\rm SN} < P_0$.

Integral Constraint

To estimate the intergal constraint contribution to the power spectrum $P_{\rm IC}(k)$ for the *Euclid* catalogues, we compute the power spectra for the $\mathcal{V}$ mocks by substituting each mock's $\bar{n}_g$ with the average comoving number density across all 140 mocks. Our findings for $P_{\rm IC}(k)$ are depicted with a magenta dash-dotted line in Fig. 5.1, revealing a bias that is insignificantly small compared to the statistical uncertainties at wavenumbers $k \gtrsim$ a few $\times\, 10^{-3} h$ Mpc^{-1}. In absolute terms, utilizing $\bar{n}_g$ results in an underestimation of $P_0(k)$ by 5 to 20 % (depending on the redshift bin) on the largest scales probed, where the random errors are at least several factors larger.

5.2.2 Comparison with the $\mathcal{V}$ Mocks

After discussing the bias and variance of the FKP estimates, we turn to the comparison of the spectra themselves. In this section, we compare the monopole spectra of the $\mathcal{V}$ mocks built with the *Euclid* survey functions in 4 redshift bins bordered by $z \in \{0.9, 1.1, 1.3, 1.5, 1.8\}$ to the GPP model computed using Eq. (5.5). We compute the spectra for the mocks employing three different angular masks, specifically the full-sky, patch and *Euclid* footprints shown in Fig. 3.5. In order to produce the window-convolved theoretical prediction, we compute the theoretical power spectrum monopole using the GPP model (see Eq. 2.95). After which, we are able to compute Eq. (5.5) numerically for the full-sky footprint.

The results of this comparison are presented in Fig. 5.3. The figure displays the monopole of the power spectrum obtained from the $\mathcal{V}$ mock catalogues without shot noise (depicted by a black solid line surrounded by a grey-shaded region). Notably, the clustering signal in redshift space appears significantly amplified compared to real space. To facilitate the comparison between the simulations and the GPP Kaiser prediction, we also plot the Kaiser-boosted linear power spectrum for the light cone before (dash-dotted black curve) and after (dashed black curve) convolution with the survey window function. Notably, the window-convolved linear prediction closely aligns with our simulations for wavenumbers $k \gtrsim 0.02, h$ Mpc^{-1} across all redshift bins. Conversely, the simulations exhibit a distinct increase in clustering amplitude on larger scales. The vertical line in the plots highlights the wavenumber at which this systematic deviation surpasses the RMS statistical error.

The lower panels illustrate the mean values across 140 realizations of the ratio $P_0(k)/P(k)$ as a function of k. Different colours represent the various angular masks: the full-sky footprint (green, top), the *Euclid* footprint (blue, middle), and the patch footprint (magenta, bottom). The light-blue shaded area denotes the uncertainty (central 68% range) in the ratio based on the mocks, while the horizontal dashed line corresponds to the Kaiser boost outlined in Eq. (2.85). Notably, the deviation from the Kaiser formula becomes larger as the survey area increases, as expected for wide-angle effects. We note that we set the magnification and evolution bias to zero in creating the $\mathcal{V}$ mocks. Consequently, we conclude that the excess power observed in Fig. 5.3 is primarily attributable to wide-angle effects and the variation in the line of sight neglected in the GPP model.

5.3 Local Plane Parallel Approximation

In this section, we extend the theoretical modelling by using the LPP approximation analogue of Eq. (5.4). As Eq. (5.4) adopts the GPP approximation, it neglects contributions arising from the variable line of sight that are inherent in all multipoles estimated using the YB estimator. As demonstrated in Beutler et al. (2014) and Wilson et al. (2017), the impact of the variable line of sight can be partially mitigated by improving the window convolution prescription.

Under the LPP approximation, the impact of the survey window is represented

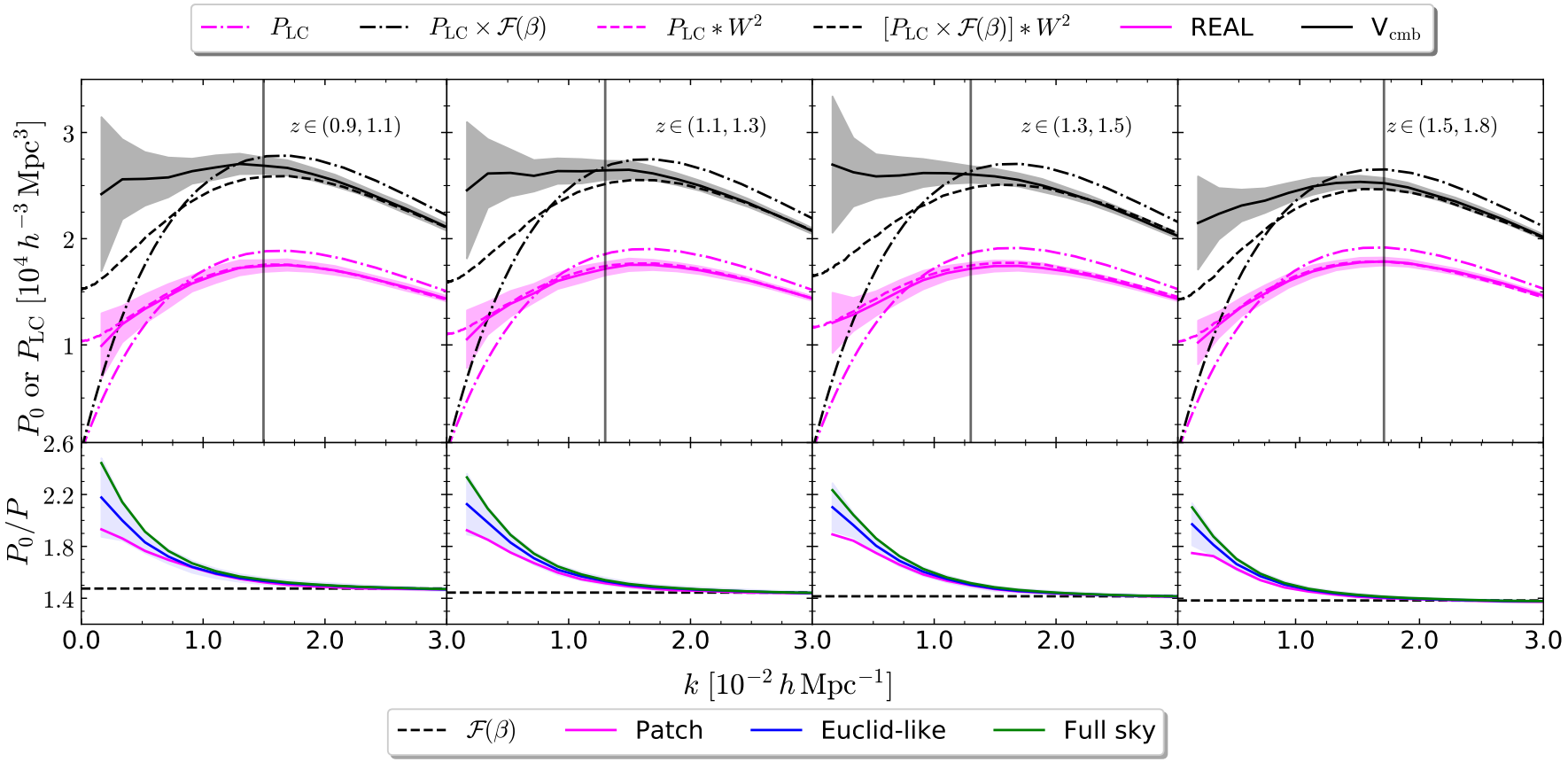


Figure 5.3: Top: Power spectra for a full-sky survey in different redshift bins. The mean (solid) and the scatter (central 68% range, shaded region) of the power spectrum monopole obtained from the 140 $\mathcal{R}$ catalogues (magenta) and of the corresponding the $\mathcal{V}$ mocks (black) are compared with the corresponding GPP linear-theory predictions without (dash-dotted) and with (dashed) accounting for the survey window function. The vertical lines highlight the wavenumber at which the $\mathcal{V}$ spectra deviate by more than one standard deviation from the Kaiser-boosted linear power spectrum. Bottom: The mean ratio between the monopole moment of the power spectrum in redshift space and the real space power spectrum for a full-sky survey, an *Euclid* survey, and the patch of the sky (from top to bottom) all with the same galaxy population. The horizontal dashed line indicates the Kaiser factor while the light-shaded area denotes the central 68% range for the Euclid case

as a matrix multiplication that intermingles the various theoretical multipoles. Mathematically, this yields: (Beutler & McDonald, 2021)

$$P_\ell^{\rm obs}(k) = \sum_{\ell'} \int k'^2\, \mathcal{W}_{\ell\ell'}(k,k')\, P_{\ell'}(k')\, {\rm d}k' - \frac{P_\ell^W(k)}{P_0^W(0)} \sum_{\ell'} \int k'^2\, \mathcal{W}_{\ell\ell'}(0,k')\, P_{\ell'}(k')\, {\rm d}k' \,, \tag{5.8}$$

where $P_{\ell'}(k')$ are the multipoles of the theoretical power spectrum defined in Eq. (2.95). Eq. (5.8, hereafter referenced as the LPP Kaiser model) still contains the mode-mode mixing in the same multipole that is observed in Eq. (5.5). Additionally, the summation over ℓ' in Eq. (5.8) also results in mixing between different multipoles of the power spectrum. In other words, the observed power spectrum monopole is affected by all multipoles of the true power spectrum.

It is important to note that in Eq. (5.8), we employ the LPP approximation only partially, as we continue to use the GPP power spectrum model (see Eq. 2.95). Consequently, the summation in Eq. (5.8) is only non-zero for $\ell' = \{0, 2, 4\}$. Higher-order corrections to the previous model can be derived by expanding the power spectrum perturbatively in the parameter $(k\,x)^{-1}$ (Castorina & White, 2018). However, these perturbative corrections have been shown to be subdominant (e.g., see Castorina et al., 2019, Beutler & McDonald, 2021) compared to the model seen in Eq. (5.8) .

The mixing matrix of Eq. (5.8) $\mathcal{W}_{\ell\ell'}(k,k')$ can be constructed using

$$\mathcal{W}_{\ell\ell'}(k,k') = \frac{{\rm i}^{2\ell'-\ell}}{2\pi^2}(2\ell+1) \sum_L \begin{pmatrix} \ell & \ell' & L \\ 0 & 0 & 0 \end{pmatrix}^2 \int r^2 j_\ell(k\,r)\, j_{\ell'}(k'r)\, Q_L(r)\, {\rm d}r \,, \tag{5.9}$$

where the terms in brackets are the Wigner 3j symbols, j_ℓ denotes the spherical Bessel function of order ℓ and $Q_L(r)$ are the 2PCF multipoles of the window function. The 2PCF of the window function is defined as

$$Q(\boldsymbol{r}) = \langle W(\boldsymbol{x}')W(\boldsymbol{x}'+\boldsymbol{r})\rangle. \tag{5.10}$$

Its multipoles $Q_L(\boldsymbol{r})$ can be directly computed with the 2PCF estimator detailed in Eq. (4.10) by counting pairs in the random catalogue (Wilson et al., 2017). However, it is computationally faster to estimate it as follows. First, we redefine the window function in Eq. (5.2) as

$$\hat{W}(\boldsymbol{x}) = \frac{\alpha\, w(\boldsymbol{x})\, \hat{n}_{\rm r}(\boldsymbol{x})}{\sqrt{A}} \,. \tag{5.11}$$

We then estimate the power spectrum multipoles of the window function, $\hat{W}(\boldsymbol{x})$, (hereafter denoted by P_ℓ^W) using the YB estimator. Finally, we calculate the 2PCF

multipoles via a Hankel transform

$$Q_\ell(r) = \frac{i^\ell}{2\pi^2} \int k^2 P_\ell^W(k) j_\ell(k\,r)\,\mathrm{d}k\,. \tag{5.12}$$

In the following sections, we compare the spectra of the $\mathcal{V}$ mocks to the theoretical prediction obtained through the LPP model shown in Eq. (5.8). Since only the first three even multipoles are non-zero in the GPP theoretical model (see Eq. 2.95), we consider the $\ell' \in \{0, 2, 4\}$ multipoles in Eq. (5.8). Consequently, we only need to compute the 2PCF multipoles of the window function up to $L = 8$ based on the triangle rules of the Wigner 3j symbols in Eq. (5.9). All Hankel transforms are performed using the `FFTLog` algorithm (Hamilton, 2000).

5.3.1 Connected and Disconnected regions

The procedure outlined in the previous section has been thoroughly tested on simply connected regions such as the patch footprint shown in Fig. 3.5a (Beutler & McDonald, 2021), while it has not been used frequently to analyse disconnected patches. In other words, disconnected patches in the sky (e.g. the northern and southern galactic caps) are analysed independently. The *Euclid* footprint avoids the Ecliptic and Galactic planes and is thus composed of four disconnected regions (see Fig. 3.5b). To access the largest scales possible in that footprint, we need to analyse all disconnected regions simultaneously as presented in Sec. 5.2.2 instead of estimating four separate spectra. In this section, we compare the random 2PCF multipoles and the mixing matrices of the two approaches; *i)* analysing the single patch shown in Fig. 3.5a, *ii)* studying the full *Euclid* footprint. To compare these two methodologies, Fig. 5.4 presents a comparison of the functions Q_0, Q_2, and Q_4 obtained for the entire *Euclid* footprint (solid lines) and the rectangular patch covering $2565\,\mathrm{deg}^2$ (dashed lines). Both scenarios are evaluated within the redshift bin $z \in (1.1, 1.3)$. For spatial separations of a few $h^{-1}\,\mathrm{Mpc}$, the monopole moment converges to unity, while Q_2 and Q_4 approach zero, indicating the nearly constant and isotropic nature of $Q(\mathbf{r})$ within spatial lags contained within the survey window function. Conversely, $Q(\mathbf{r})$ diminishes as r exceeds the maximum distance between two galaxies in the patch. In the intermediate region between these asymptotic behaviours, the two-point correlation function (2PCF) of the random catalogue exhibits high anisotropy, with $|Q_2|$ and Q_4 assuming values larger than Q_0. Notably, disparities between the two survey footprints manifest prominently on the largest spatial scales. For example, in the *Euclid* case, Q_4 exhibits negativity for $1600\,h^{-1}\,\mathrm{Mpc} \lesssim r \lesssim 4400\,h^{-1}\,\mathrm{Mpc}$ and demonstrates a strong peak at $r \simeq 5000\,h^{-1}\,\mathrm{Mpc}$, attributed to the presence of disconnected patches.

In Fig. 5.5, we plot elements of the mixing matrices $\mathcal{W}_{\ell,\ell'}$ for the two geometries. We plot the *Euclid* footprint in black while the connected patch is shown in red. We show the behaviour of the multipoles of the mixing matrix around $k = 0.03\,h\,\mathrm{Mpc}^{-1}$ as a function of k' for the second redshift bin $z \in (1.1, 1.3)$. We can see that for all combinations of the multipoles, there are high-frequency oscillations in the *Eu-*

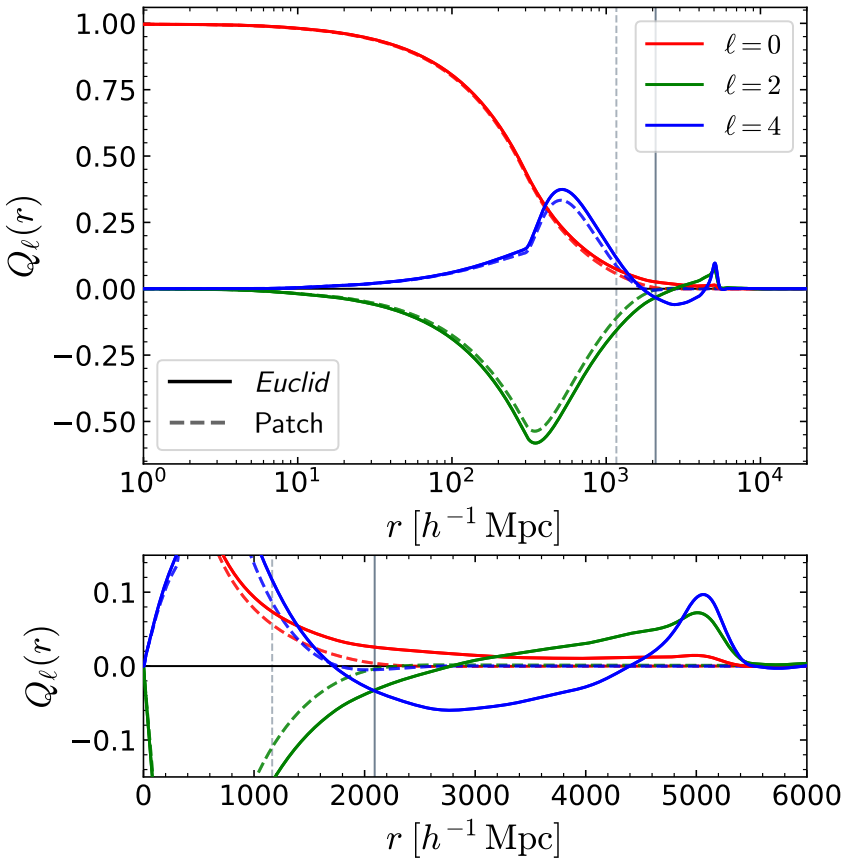


Figure 5.4: We plot the 2PCF multipoles of the survey window computed from the *Euclid* whole footprint (solid lines) and from a single rectangular patch (dashed) in the second redshift bin $z \in (1.1, 1.3)$. The two vertical lines represent the characteristic lengths $V^{1/3}$ of the two volumes. In the case of a single patch, the Q_ℓ functions tend to zero beyond this scale. However, the non-trivial geometry of the *Euclid* footprint sources an extra signal at higher values of r. This Figure is borrowed from Paper III.

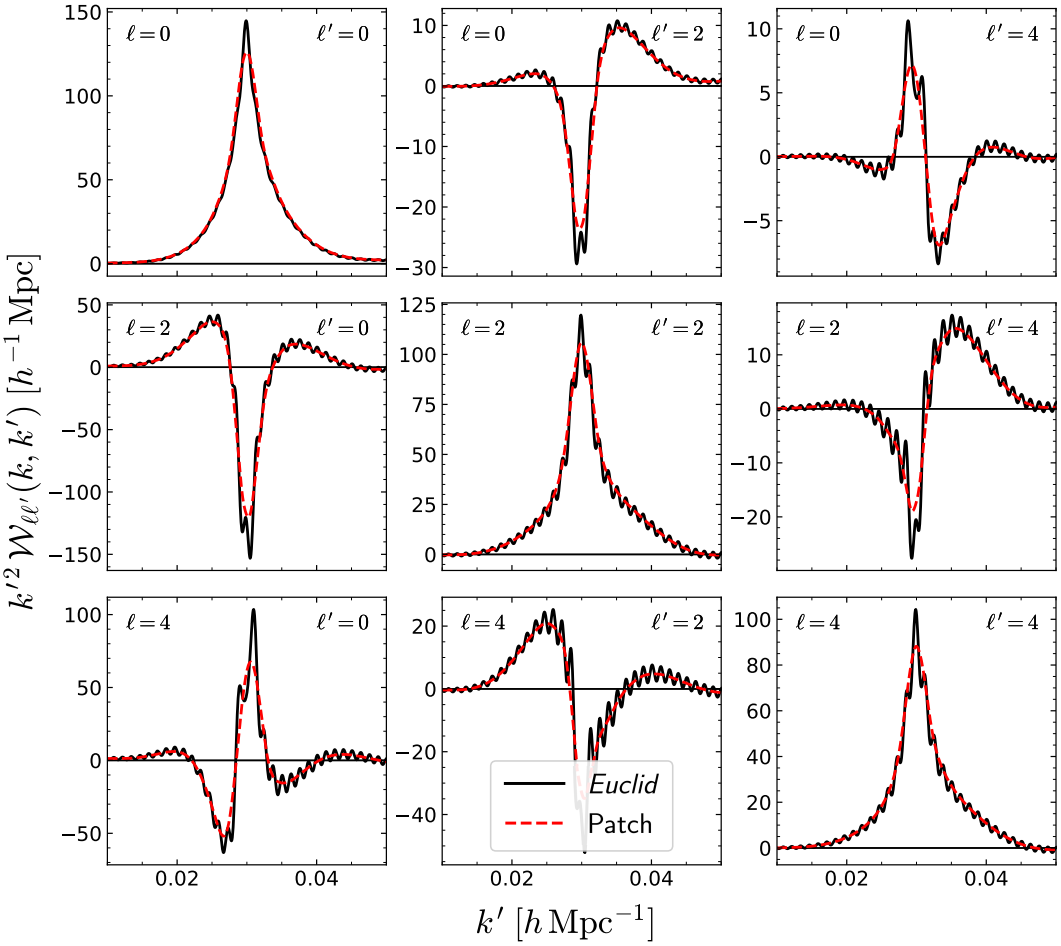
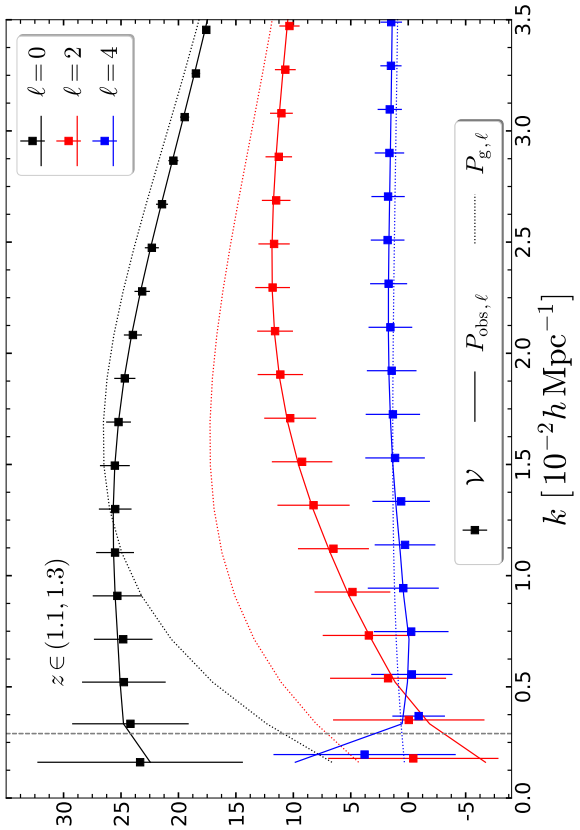
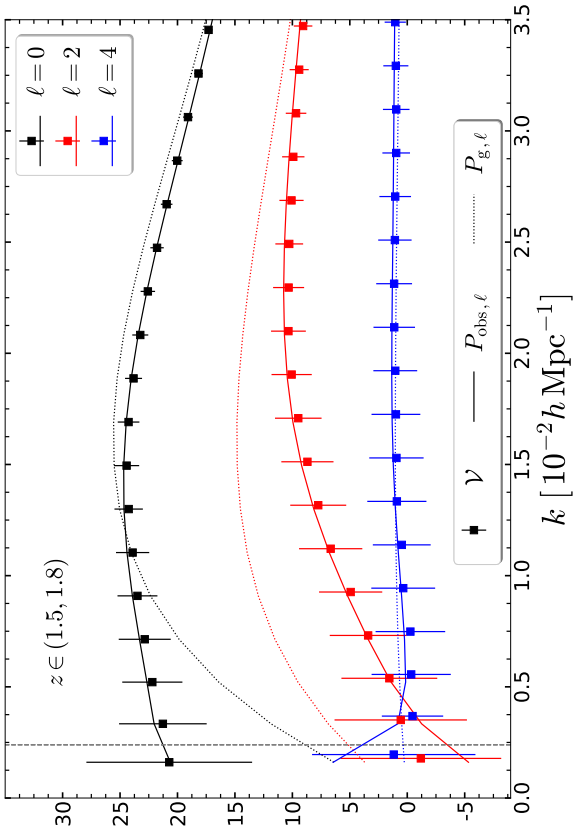


Figure 5.5: Elements of the mixing matrix for the two surveys discussed in Fig. 5.4 evaluated at $k = 0.03\,h\,\mathrm{Mpc}^{-1}$. For $\ell = \ell'$, the resulting kernel quickly drops to zero when $|k - k'|$ is larger than the inverse of the size of the window. For $\ell \neq \ell'$ the power spectrum multipoles mix together in non-trivial ways, switching signs across modes. High-frequency wiggles are noticeable in the *Euclid* mixing matrix, sourced by the extra spike seen in the random 2PCF multipoles $Q_\ell(x)$. This Figure is taken from Paper III.



(a) $z \in [0.9, 1.1]$.

(b) $z \in [1.1, 1.3]$.



(c) $z \in [1.3, 1.5]$

(d) $z \in [1.5, 1.8]$.

Figure 5.6: The power spectrum multipoles are shown for all the redshift bins for the $\mathcal{V}$ mocks. The theoretical prediction given by the Kaiser model before and after the window function convolution is shown in solid and dashed lines respectively. The top black line represents the monopole, the red line below denotes the quadropole while the blue bottom line represents the hexdecapole.

clid footprint that are absent in the single patch case. This can be traced back to the extra spike in the random catalogue 2PCF (see bottom panel of Fig. 5.4). Secondly, from the figure, we can understand the non-trivial impact of the survey window function. Intuitively, we expect that the convolved power spectrum at a specific scale k is dominated by the original unconvolved power spectrum at the same scale. This behavior is confirmed in the $\ell = \ell'$ diagonal cells of Fig 5.5 as $\mathcal{W}_{\ell,\ell'}$ peaks around $k' = k$ and drops on both sides. However, it is clear that the convolved power spectra include contributions from nearby modes as well. Correlation across different multipoles is less trivial but qualitatively looks like the diagonal case multiplied by powers of the cosine function.

5.3.2 Comparison with $\mathcal{V}$ mocks.

We then turn to the comparison between the estimated mock $\mathcal{V}$ spectra and the theoretical model that accounts for the survey window and the integral constraint (given by Eq. 5.8). The primary objective of the first comparison is to validate the theoretical model, given that it has been previously tested only on connected survey masks. However, our analysis encompasses several disconnected patches, hence the need for this validation. For clarity, let us first summarise how the theoretical prediction is computed: *i*) We obtain the matter power spectrum at the redshift $z = 0$ using the `CAMB` code, *ii)* we compute the galaxy power spectrum averaged over the redshift using Eq. 2.95, *iii)* finally, we account for the survey window function and the integral constraint via Eq. (5.8) using the `FFTLog`.

We show the results of the comparison in Fig. 5.6. We plot the means and standard deviations of the multipoles of the $\mathcal{V}$ spectra as squares with error bars, while the theoretical expectations are plotted with lines. Each panel denotes one of the redshift bins of the *Euclid* survey. The vertical dashed line denotes the fundamental frequency $k_{\rm F,S} = 2\pi/V_{\rm S}^{1/3}$ (Tegmark et al., 1998a). It is clear in the Figure that the theoretical model accurately reproduces the mock results for most of the frequencies. Slight discrepancies are observed for $k \sim k_{\rm F,S}$, which are caused by the residual wide-angle effects. Let us stress that setting the magnification and evolution bias in the $\mathcal{V}$ mocks to zero removes part of the peculiar velocity effects. Consequently, the success of the comparison in Fig. 5.6 indicates that the model is able to reproduce the impact of the Kaiser gradient term and its related wide-angle terms, not the full peculiar velocity sourced terms.

5.4 Summary

In this Chapter, we discussed how the limited volume observed in surveys impacts the observed power spectrum. Using the spectra measured from the *Euclid*-like $\mathcal{V}$ mocks as a test case, we found that the GPP Kaiser model underestimates the power spectrum monopole at large scales (see Sec. 5.2). In Sec. 5.3, we turned to the comparison with the LPP model which accounts partially for the variation of the line of sight. We first confirmed that it is possible to use the LPP approximation to

analyse disconnected patches simultaneously in contrast to the standard analysis. Comparing the LPP model to the $\mathcal{V}$ spectra, we found excellent agreement for all redshifts, except at extremely large scales $k \sim k_{,\mathrm{S}}$. A natural extension of this chapter is to compare the $\mathcal{V}$ spectra to the power spectrum multipoles computed via the spherical Fourier-Bessel Basis presented in (Wen et al., 2024), as this method is exact at linear order.

Chapter 6

Relativistic RSD in the *Euclid* Survey.

In this chapter, we delve into a pivotal objective of this thesis: assessing the detectability of relativistic RSD using upcoming galaxy redshift surveys, focusing primarily on the *Euclid* Wide Spectroscopic Sample. Specifically, we scrutinize two key effects: *i)* the gravitational weak-lensing RSD, and *ii)* the signal induced by the peculiar velocity of the observer. We gauge the significance of these effects through the comparison of various summary statistics. The gravitational weak-lensing RSD's impact is evaluated by contrasting the estimated statistics of the $\mathcal{G}$ mocks with those of $\mathcal{V}$ mocks, while the observer's peculiar velocity effect is examined by analyzing the discrepancy between the $\mathcal{O}$ and $\mathcal{G}$ statistics. Ultimately, we assess the importance of all relativistic RSD at linear order by juxtaposing the statistics derived from $\mathcal{O}$ and $\mathcal{V}$ mocks. Through these analyses, we aim to shed light on the discernibility and implications of relativistic RSD in the context of upcoming galaxy redshift surveys, particularly within the *Euclid* spectroscopic sample.

This chapter is based on results from Paper I where we focus on the importance of relativistic RSD in the power spectrum monopole and Paper III where we extend these results to all even power spectrum multipoles, 2PCF multipoles and the angular power spectrum.

6.1 Power Spectrum Monopole

In this section, we present the findings from the tests conducted in Paper I. That study employed the FKP estimator (see Sec. 4.3) to measure the monopole of the power spectrum for both $\mathcal{V}$ and $\mathcal{G}$ mocks in 4 tomographic redshift bins bordered by $z \in \{0.9, 1.1, 1.3, 1.5, 1.8\}$ and a bin that encompasses the entire survey, i.e. $z \in (0.9, 1.8)$. These computations are conducted using two distinct footprints: the *Euclid* footprint and a full-sky footprint, illustrated in Fig. 3.5. Furthermore, we compute the spectrum of each mock with and without incorporating the shot-noise addition, as outlined in Sec. 3.4.2.

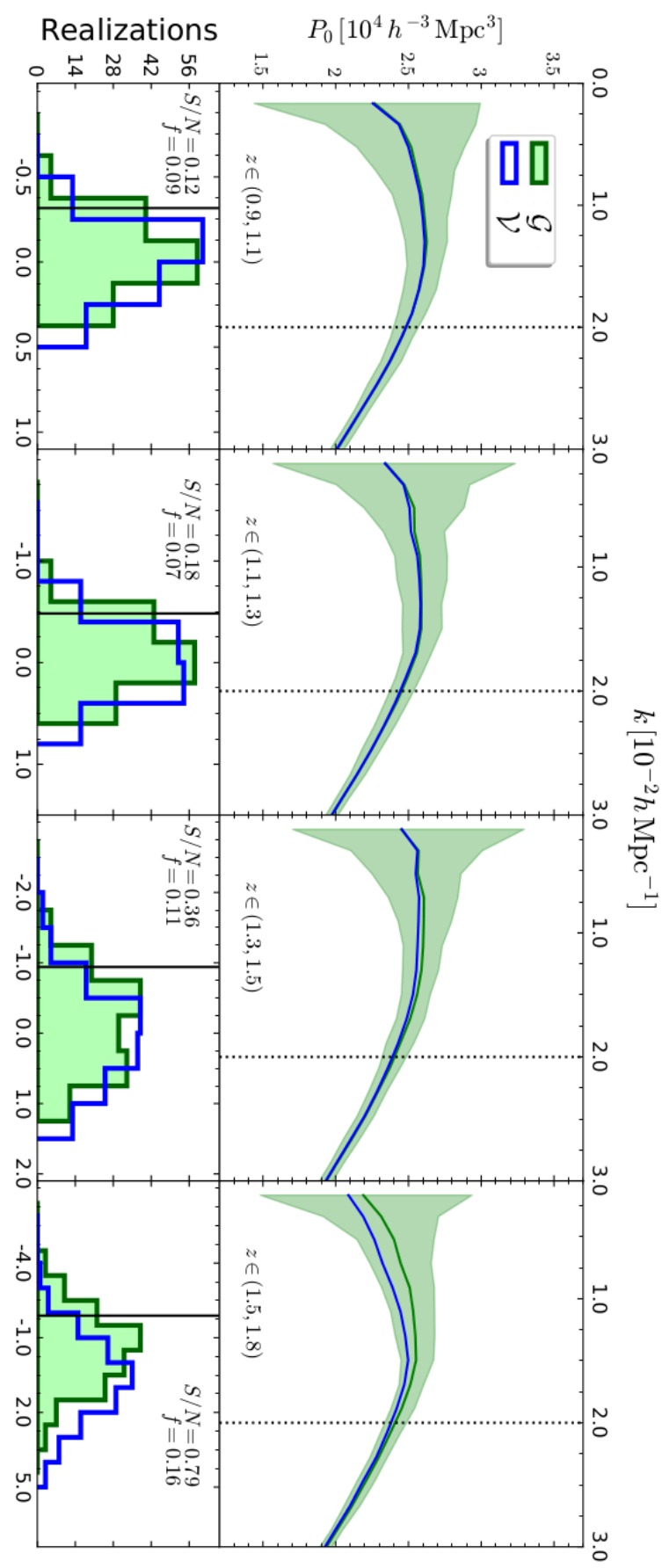


Figure 6.1: Top: Monopole moment of the power spectrum in different redshift bins for a Euclid-like survey. The mean power spectrum monopole obtained with the $\mathcal{G}$ mocks (green solid line) is compared with its counterpart computed from the $\mathcal{V}$ mocks (blue solid line). The shaded region denotes the RMS scatter among the $\mathcal{G}$ mocks (central 68% region). The dotted vertical line marks the maximum k used in the statistical test. Bottom: DLR test between the models shown in the top panel (see Sec. 4.4.1 for details). The two histograms represent $\mathcal{P}(\lambda)$ for the $\mathcal{G}$ mocks (left, light green) and the $\mathcal{V}$ mocks (right, blue). The vertical line marks the decision threshold ω_{95}. Each panel reports the power of the test f and the separation between the histograms in terms of the S/N parameter. This figure is taken (with minor adjustments) from Paper I.

Subsequently, we employ the DLR test (see Sec 4.4.1) to measure the significance of the weak lensing signal. Finally, we compare the $\mathcal{G}$ model against the GPP Kaiser power spectrum model to measure the (S/N) by which the GPP model is rejected.

6.1.1 Weak Gravitational Lensing

We first compare the monopole spectra obtained from the $\mathcal{V}$ to the ones estimated from the $\mathcal{G}$ mocks. Both mocks include the SN contribution using the *Euclid* footprint. We show the results of the comparison in the top panel of Fig. 6.1. In the Figure, we display the mean spectra obtained by averaging over the 140 light cones of the $\mathcal{V}$ mocks in blue while the mean of the $\mathcal{G}$ mocks is shown in green. The shaded region indicates the RMS scatter (central 68-per-cent range) for the $\mathcal{G}$ mocks. We observe a slight increase in the amplitude of the power-spectrum monopole on large scales due to gravitational lensing. However, this systematic shift is smaller than the scatter among the mocks in the case of the low redshift bins or comparable to the RMS scatter at higher redshifts. The small difference in the power spectra extracted from the $\mathcal{V}$ and $\mathcal{G}$ mocks can be traced back to the relatively uniform luminosity function of the Hα emitters detected at the Euclid flux threshold, corresponding to a magnification bias range of $2.0 < Q(z) < 2.6$ (see red solid line in Fig. 3.4). Notably, the difference between the two models would be more significant if the faint-end slope of the luminosity function was substantially steeper. We find a relatively higher signal when we consider the large redshift bin as shown in Fig. 6.2. As clear from the top panel in the Figure, the $\mathcal{G}$ signal is of the order of σ different from the $\mathcal{V}$ mean.

We then use the ($\mathcal{G}$-$\mathcal{V}$) DLR test to measure the significance of the weak-lensing signal for the different redshift bins. For the data vector, we use $n = 10$ bins in the range $k \in [0.4, 20] \times 10^{-3}\, h\,\mathrm{Mpc}^{-1}$. We show the results for the tomographic redshift bins in the bottom panel Fig. 6.1. The blue histograms show $\mathcal{P}(\lambda|\mathcal{V})$ while the vertical line highlights w_{95}. The green-filled histograms on the left-hand side, instead, represents $\mathcal{P}(\lambda|\mathcal{G})$. However, the histograms strongly overlap in most of the redshift bins. In most scenarios, distinguishing between the models based on the data from a *Euclid* survey proves challenging. The likelihood-ratio test exhibits relatively low power, with a marginal increase observed solely in the highest-redshift bin. These findings imply that incorporating radial redshift-space distortions arising from peculiar velocities (i.e., departing from the distant-observer approximation in the Kaiser model) should offer adequate accuracy in describing $P_0(k)$, particularly for $z < 1.5$. However, at higher redshifts and for $k < 0.02, h$ Mpc^{-1}, this model may exhibit a systematic shift roughly equivalent to the sampling uncertainty of $P_0(k)$ (excluding shot noise). Such bias introduced by simplified models could impede cosmological inference based on the clustering signal at extremely large scales, emphasizing the necessity to assess the significance of this shift due to relativistic effects on a case-by-case basis. This outcome reflects the low magnification bias observed in *Euclid* galaxies and could change if the faint-end slope of their luminosity function was steeper than currently estimated. The conclusions outlined above pertain to

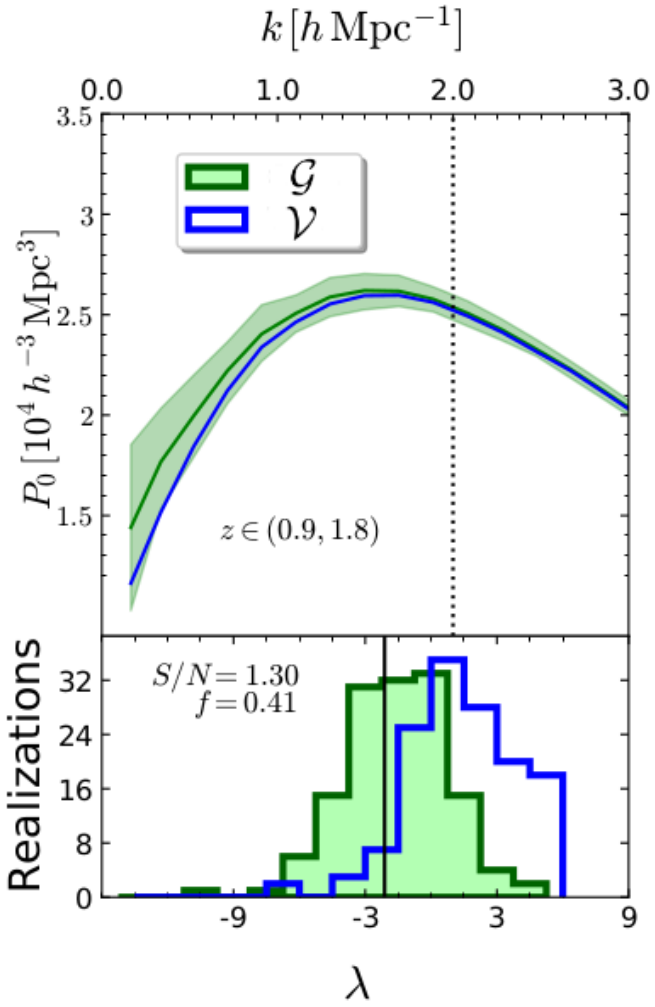


Figure 6.2: As in Fig. 6.1 but for the redshift bin that encompasses the entire survey. This figure is taken from Paper I.

the relatively narrow redshift bins, typically with a width of $\Delta z \simeq 0.2$-0.3. In the bottom panel of Fig. 6.2, we extend the analysis to encompass the entire redshift range covered by the Euclid spectroscopic survey, spanning $0.9 < z < 1.8$. Here, we observe that the impact of light-cone effects, particularly gravitational lensing, is discernible in 41% of the realizations. This finding suggests that the significance of lensing effects becomes more pronounced as the size of the redshift bin under examination increases.

6.1.2 Comparison to the GPP Kaiser Model

We then compare the $\mathcal{G}$ mocks to the GPP Kaiser model presented in Sec. 5.2.2. It is important to note that when computing the Kaiser power spectrum multipoles in both the GPP and LPP predictions, we account for the lightcone (or redshift) evolution by using $F_{\rm eff,\ell}(z_i, z_f)$ which is defined in Eq. (2.96). However, this relation is only valid at linear scales, while it is not possible to define one single relation that evolves simultaneously both the linear and non-linear parts of the power spectrum (Pryer et al., 2022). We show the results of the comparison for the small tomographic bins in Fig. 6.3. The green solid line and shaded regions are the same as in Fig. 6.1. As for the Kaiser prediction plotted in dashed blue, we simply multiply the mean of the spectra computed from the $\mathcal{R}$ mocks by the Kaiser monopole boost factor defined in Eq. (2.85), we denote the result by $\mathcal{F}_0\,\mathcal{R}$. The bottom panel shows the results for the ($\mathcal{G}$-$\mathcal{F}_0\,\mathcal{R}$) DLR test. We use the same 10 bins similar to the ($\mathcal{G}$-$\mathcal{V}$) DLR test. The histogram on the right-hand side shows $\mathcal{P}(\lambda|M_1)$ (which is obtained using the $\mathcal{F}_0\,\mathcal{R}$ mocks) while the vertical line highlights w_{95}. The histogram on the left-hand side, instead, represents $\mathcal{P}(\lambda|M_2)$ which is negative for most of the distribution as the $\mathcal{G}$ model is preferred by most of the $\mathcal{G}$ mocks. From the figure, we surmise that the ($\mathcal{G}$-$\mathcal{F}_0\,\mathcal{R}$) DLR test is able to reject the Kaiser model in approximately three-quarters of the realisations. The power of the test is reported in each panel together with the S/N. Based on these results, we conclude that, for $k < 0.02\,h$ Mpc^{-1}, the GPP Kaiser model for RSD does not provide an accurate description of the power spectrum monopole in the *Euclid* survey and can be statistically ruled out in about 80% of our mock realizations.

6.1.3 Full-sky Footprint and Shot-Noise Contributions

Finally, we repeat both tests carried out in the previous sections for a full-sky footprint, excluding the SN contribution to the statistical errors. Both signal-to-noise ratios and fractions are shown in Table 6.1. Starting with the full-sky ($\mathcal{G}$-$\mathcal{F}_0\,\mathcal{R}$) DLR test, we note that S/N is higher in all bins compared to the *Euclid* footprint and assumes values up to 4.49 (see the left half of Tab. 6.1). In order to estimate the S/N of surveys with higher number densities, we also provide the results obtained neglecting the SN contribution (right half of Table 6.1). These results clearly indicate that the GPP Kaiser model is unable to describe the large-scale power-spectrum monopole of these denser surveys.

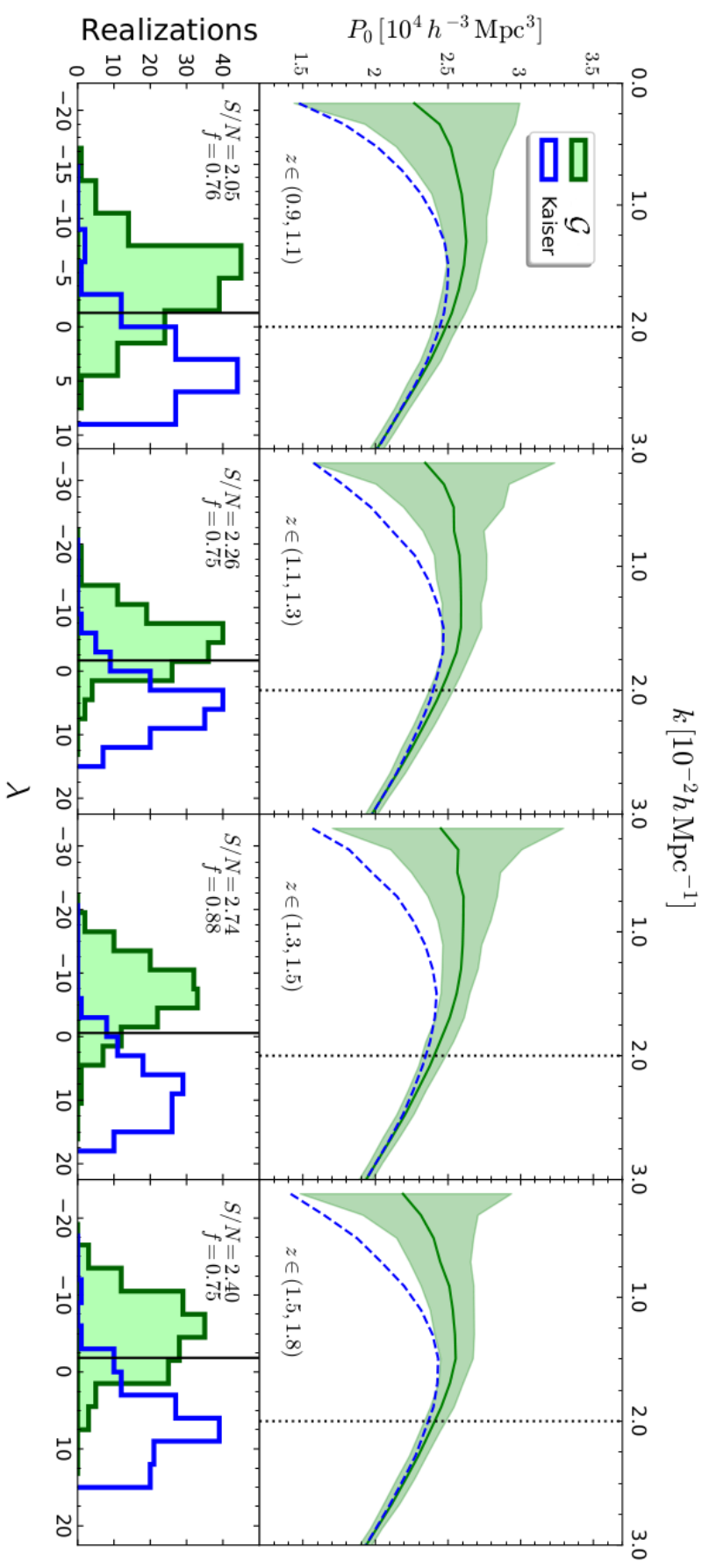


Figure 6.3: As in Fig. 6.1 but for the $\mathcal{G}$ mocks (green solid line) and the Kaiser model applied to the real-space data (blue dashed line). This figure is taken from Paper I.

However, Tab. 6.1 also shows that the GPP Kaiser model performs better in the thickest redshift bin. We can explain this phenomenon by observing that for thick redshift bins, the power spectrum monopole is likely dominated by pairs with radial separations as opposed to narrow redshift bins where there is a large contribution from pairs with wide-angular separations. We test this phenomenon by measuring the S/N in broader redshift bins with the same mean redshift, and indeed we find that the S/N to noise of the ($\mathcal{G}$-$\mathcal{F}_0\,\mathcal{R}$) DLR test decreases with the thickness of the redshift bin.

As for the ($\mathcal{G}$-$\mathcal{V}$) DLR test, we note that predictably the S/N is enhanced in the full-sky case reaching a maximum of 2.3 for the broadest redshift bin. In summary, we derive three main conclusions from the data provided in Table 6.1.

i) We observe a general trend where the signal-to-noise ratio (S/N) tends to increase with the survey size.

ii) We note that the lensing magnification notably influences $P_0(k)$ in broad redshift bins with a width of approximately $\Delta z \simeq 1$, particularly for the full-sky footprint.

iii) We find that the contribution of shot noise to the error budget is relatively minor compared to the sampling errors. This discrepancy is primarily due to the high galaxy number density measured by the *Euclid* survey.

6.2 Power Spectrum Multipoles

In Paper III, we extend the analysis carried out in the previous section in three ways:

- We use the YB estimator (see Sec. 4.3) which makes use of random catalogues and consequently provides a more realistic treatment of the integral constraint.
- We extend the analysis to all even multipoles of the *Euclid* survey.
- We measure the significance of the velocity of the observer term (last line in Eq. 2.181).

We compute the even power spectrum multipoles of the $\mathcal{V}$, $\mathcal{G}$, and $\mathcal{O}$ mocks in the tomographic redshift bins of the *Euclid* survey. We show the mean of the measured power-spectrum multipoles for the different mocks in Fig. 6.4. The shaded regions denote the RMS scatter computed from the $\mathcal{O}$ multipoles while the green, blue and red solid lines denote the means of the $\mathcal{V}$, $\mathcal{G}$, and $\mathcal{O}$ spectra respectively. We rebin the spectra into $\Delta k = 5k_{\rm F}$ bins in order to produce these plots. It is clear that the power spectrum monopole (top row) contains the most information, i.e. has the highest signal-to-noise (P_ℓ/σ_ℓ). Conversely, the hexadecapole moment is diminished by a factor of ten compared to P_0 and fairly low signal-to-noise. The quadrupole moment exhibits properties between those of the monopole and

Table 6.1: The S/N and the fraction f with which the Kaiser and $\mathcal{V}$ models are ruled out by using the $\mathcal{G}$ mocks for a Euclid-like and a full-sky survey. The values listed in the left and right halves of the table are computed using the mock catalogues with and without shot noise, respectively. The right half, therefore, only accounts for the variations due to sampling different skies and consequently measures the gain in S/N due to increasing the number of objects included in the analysis.

	Sampling errors and shot noise								Sampling errors only							
	Euclid-like				Full sky				*Euclid*				Full sky			
	Kaiser		$\mathcal{V}$		Kaiser		$\mathcal{V}$		Kaiser		$\mathcal{V}$		Kaiser		$\mathcal{V}$	
$(z_{\min}, z_{\max})$	S/N	f	S/N	f	S/N	f	S/N	f	S/N	f	S/N	f	S/N	f	S/N	f
$(0.9, 1.1)$	2.05	0.76	0.12	0.09	3.72	0.99	0.15	0.06	2.21	0.74	0.12	0.06	4.11	1.00	0.17	0.06
$(1.1, 1.3)$	2.26	0.75	0.18	0.07	3.82	0.99	0.22	0.07	2.31	0.85	0.22	0.07	4.24	1.00	0.34	0.09
$(1.3, 1.5)$	2.74	0.88	0.36	0.11	4.20	1.00	0.59	0.13	3.15	0.97	0.48	0.19	4.55	1.00	0.70	0.16
$(1.5, 1.8)$	2.40	0.75	0.79	0.16	4.49	1.00	1.48	0.41	2.93	0.91	0.98	0.23	5.54	1.00	1.84	0.62
$(0.9, 1.8)$	1.91	0.56	1.30	0.41	3.48	0.99	2.05	0.69	1.99	0.74	1.47	0.41	3.92	1.00	2.30	0.78

Table 6.2: The measured S/N using the different SLR test of the power spectrum multipoles. The first three columns are based on the relative comparison of the different mocks. However, the last two columns denote the S/N by which the $\mathcal{O}$ mocks rejects the theoretical model given by Eq. (5.8) using 11 and 12 bins k-bins for each multipole, respectively.

$(z_{\rm min}, z_{\rm max})$	$\mathcal{O}-\mathcal{G}$	$\mathcal{G}-\mathcal{V}$	$\mathcal{O}-\mathcal{V}$	$\mathcal{O}-P_{\rm conv}$	$*(\mathcal{O}-P_{\rm conv})$
$(0.9, 1.1)$	1.1	0.4	1.1	1.3	2.5
$(1.1, 1.3)$	1.0	0.6	1.0	1.2	2.5
$(1.3, 1.5)$	1.0	1.0	1.3	1.5	2.7
$(1.5, 1.8)$	1.4	2.1	2.4	2.9	3.7

hexadecapole. For the statistical tests presented in this section, we bin each of the power spectrum multipoles into 12 bins in the range $k \in [3.2, 36] \times 10^{-3}\,h\,{\rm Mpc}^{-1}$. We then concatenate the different multipoles to form a $n = 36$ long data vector. We note that we account for the cross-correlation of the different multipoles in estimating the covariance matrix using Eq. (4.21).

6.2.1 Velocity of the Observer

By comparing the $\mathcal{O}$ and $\mathcal{G}$ spectra in Fig. 6.4, we observe that the peculiar velocity of the observer modifies all multipoles at extremely large scales. We dedicate Chapter 7 to explain this signal, but in short, the velocity of the Observer induces strong oscillations in all power spectrum multipoles at extremely large scales. The coarse k-binning adopted in Fig. 6.4 does not reveal the details of the oscillations and consequently, the signal simply appears as a large-scale boost (or reduction for $\ell = 4$) of the clustering amplitude.

We quantify the importance of that signal using the ($\mathcal{O}$-$\mathcal{G}$) SLR test in Fig. 6.5a. In contrast to the DLR tests shown in the bottom panels of Figs 6.1 and 6.2, the SLR test produces only one histogram $P(\lambda)$ shown in red in Fig. 6.5a. We overplot the expected Gaussian distribution in blue. The significance of the test can be gauged by observing the distance between the PDF peak and $\lambda = 0$ point. Finally, we shade the fraction of the PDF that rejects the null hypothesis (given by the $\mathcal{G}$ model in this case) in light blue. The ($\mathcal{O}$-$\mathcal{G}$) SLR test shows that the chances to detect these signatures with the *Euclid*-wide survey are small as all of the histograms peak fairly close to zero. The obtained S/N is 1.3 at best (see Tab. 6.2 and Fig. 6.5a) since the corrections due to $\boldsymbol{v}_{\rm o}$ are localised on very large scales where the measurement noise is large. However, the localization at large scales clustering could bias measurements of hyper-scale parameters such as the local-non-Gaussianity parameter, $f_{\rm NL}$ (Bahr-Kalus et al., 2021).

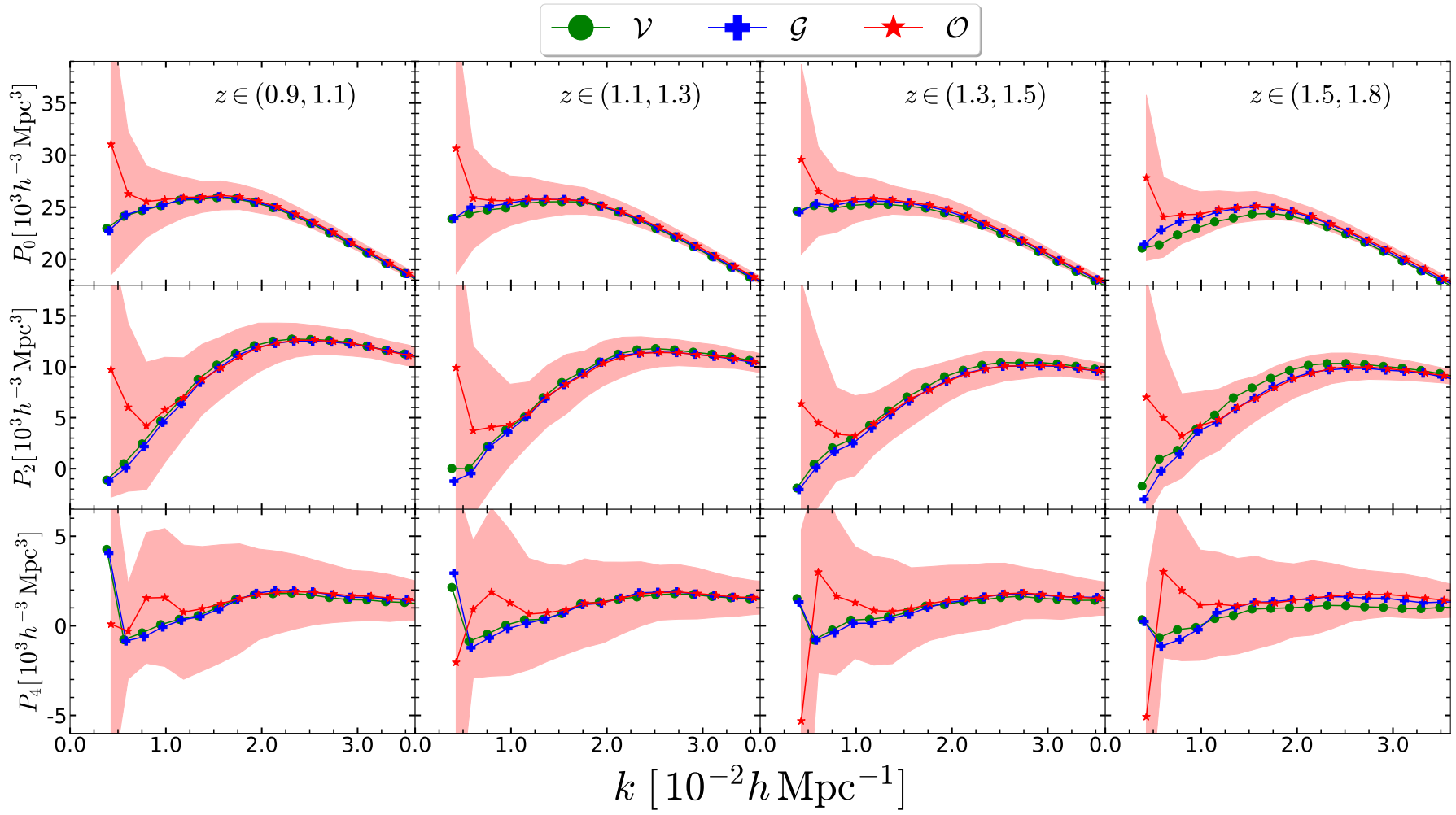


Figure 6.4: The mean $\ell = 0, 2$ and 4 multipoles are shown for the four redshift bins (left to right) that the *Euclid* survey will cover. The solid lines correspond to the mean of the spectra of the $\mathcal{V}$ (green circles), $\mathcal{G}$ (blue crosses) and $\mathcal{O}$ (red stars) mocks. The salmon-shaded regions denote the 68% scatter of the $\mathcal{O}$ spectra. This figure is taken from Paper III.

6.2.2 Weak Gravitational Lensing

Comparing the $\mathcal{G}$ and $\mathcal{V}$ spectra in Fig. 6.4, we observe that while the difference induced by weak-lensing effects is minimal in the low-redshift bins, it becomes more pronounced at deeper redshifts, particularly in the higher multipoles $\ell = 2$, 4. This aligns with the expectation that weak gravitational lensing would exert a greater influence on the clustering of galaxies at higher redshifts.

Quantifying the significance of the Weak-lensing signal through the ($\mathcal{G}$-$\mathcal{V}$) SLR test in Fig. 6.5b, we find that $\mathcal{P}(\lambda)$ peaks around $\lambda = 0$ in the first three redshift bins while it is more distant in the deepest bin. Compressing the results, we see a gradual increase in the S/N until it peaks at S/N = 2.4 in the final bin where 65% of the realizations reject the $\mathcal{V}$ model at 95% confidence level.

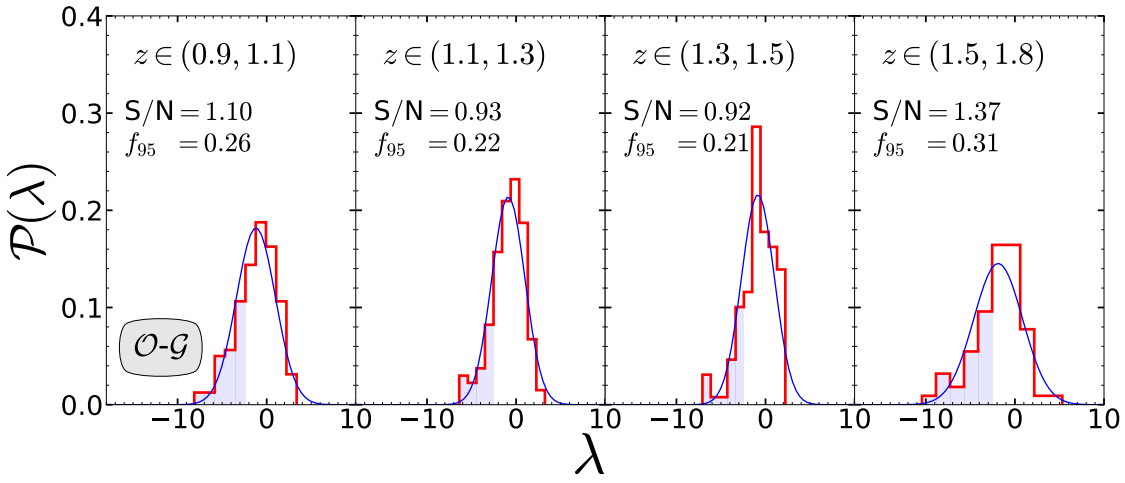
6.2.3 All Relativistic RSD

Finally, we measure the impact of both effects simultaneously by comparing the $\mathcal{V}$ and the $\mathcal{O}$ mocks. The results of the ($\mathcal{O}$-$\mathcal{V}$) SLR test are shown in Fig. 6.5b. The S/N for the detection of the all RSD is always around one in the first three bins (mostly sourced by the observer velocity) and jumps to 2.4 in the last one which is dominated by the Weak-lensing signal. This follows from the signal-to-noise ratios found in the individual tests. In summary, it is clear that the velocity-only model for RSD is ruled out with 95% confidence in 78% of the realisations.

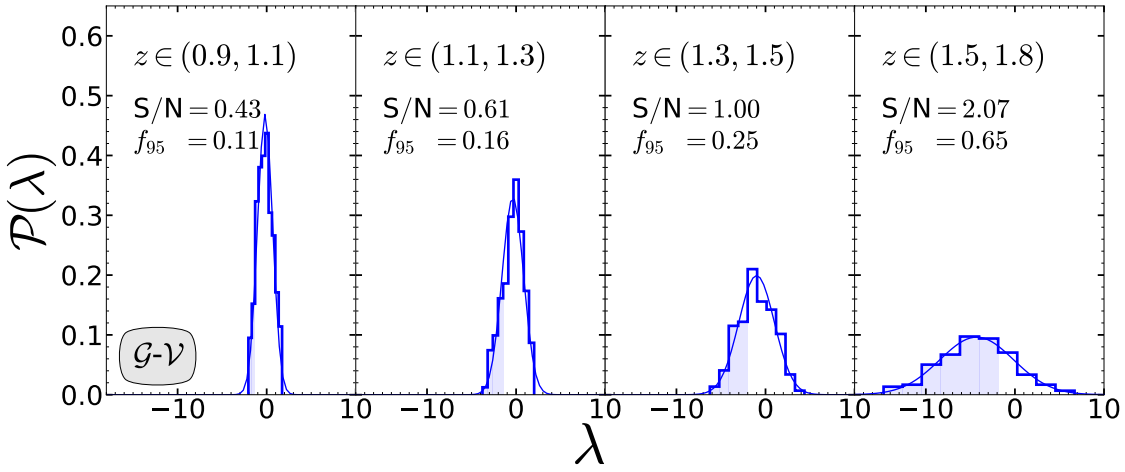
6.2.4 Comparison to the LPP Kaiser Model

Up to this point, we have evaluated the relative significance of various RSD contributions within the forthcoming *Euclid* survey by comparing the outcomes of different mock datasets. While this method has proven effective, it deviates from the customary practice typically employed in interpreting clustering measurements derived from redshift surveys. Traditionally, this involves comparing observed summary statistics with analytical models, often grounded in perturbation theory, while also accounting for the survey's window function. In this section, we pivot to this conventional approach to assess the accuracy of the LPP Kaiser model, as detailed in Section 5.3, in capturing the large-scale characteristics of the power-spectrum multipoles expected to be observed by *Euclid*. This is a direct extension of the treatment carried out in Sec. 6.1.2.

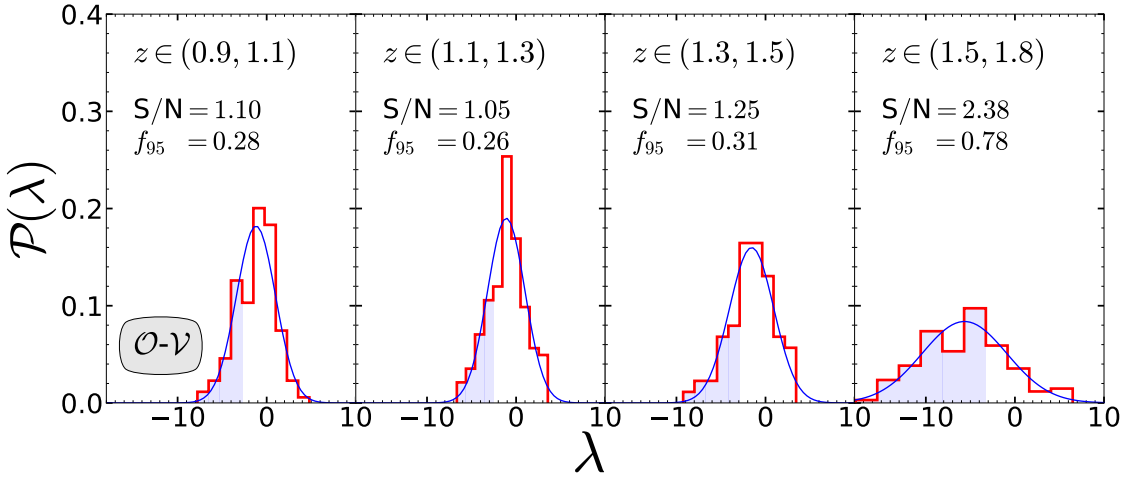
The comparison results are illustrated in Fig. 6.6. As depicted in the Figure, discrepancies between the first few multipoles and the LPP Kaiser model persist consistently, primarily attributable to the signal induced by the observer's velocity. Furthermore, we observe these discrepancies shifting towards smaller scales, particularly noticeable in the higher multipoles, as the mean redshift rises. Notably, the additional signal observed at deeper redshifts is predominantly sourced from the weak gravitational lensing effect. This comparison highlights a discernible difference in the scale range affected by weak lensing compared to the observer's velocity effect, with the former exhibiting influence over a much broader range of scales.



(a) ($\mathcal{O}$-$\mathcal{G}$) SLR test.

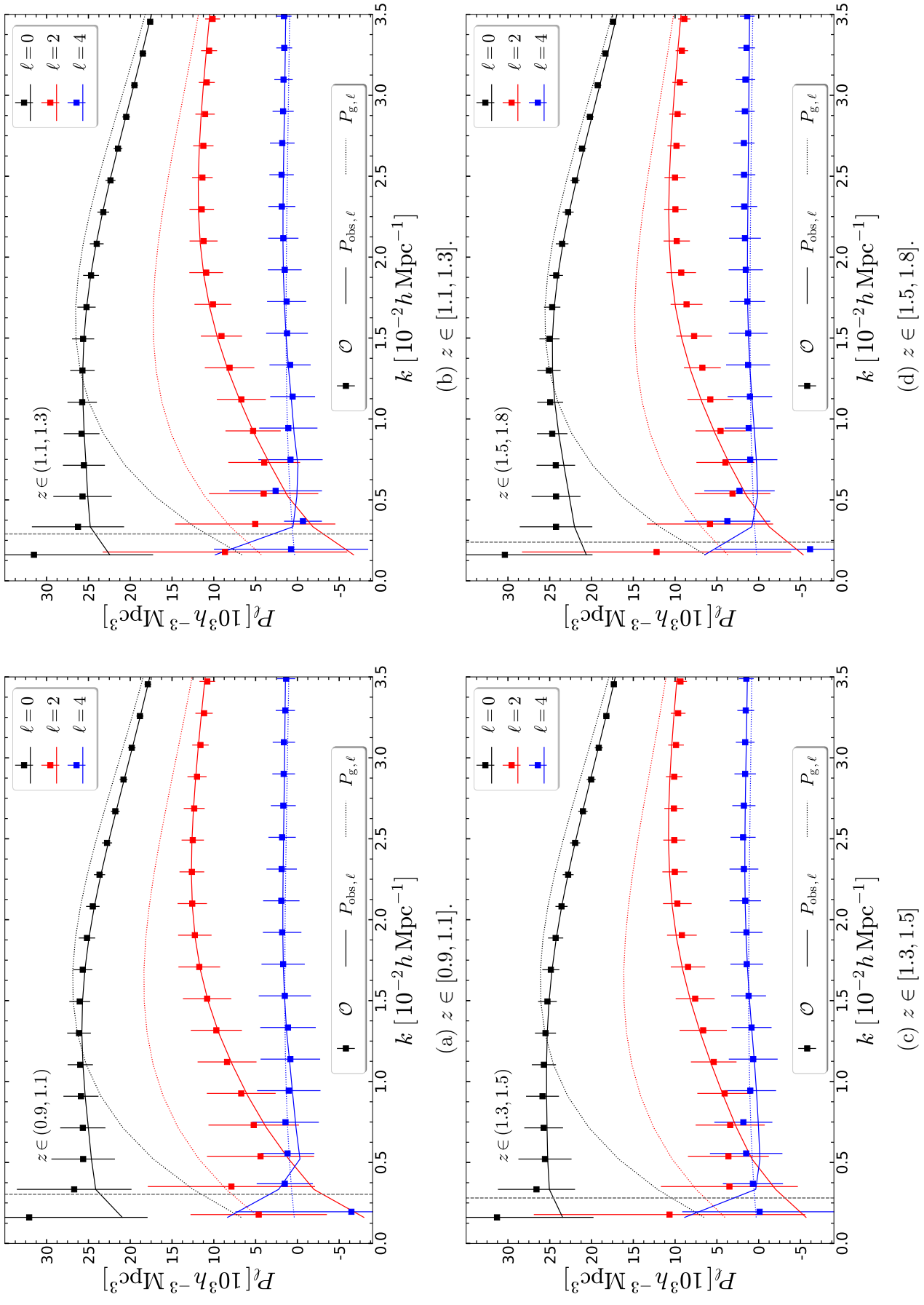


(b) ($\mathcal{G}$-$\mathcal{V}$) SLR test.



(c) ($\mathcal{O}$-$\mathcal{V}$) SLR test.

Figure 6.5: PDF of the likelihood-ratio test statistic for the SLR tests applied to the multipoles of the power spectrum.



(a) $z \in [0.9, 1.1]$. (b) $z \in [1.1, 1.3]$. (c) $z \in [1.3, 1.5]$ (d) $z \in [1.5, 1.8]$.

Figure 6.6: Similar Figure to Fig. 5.6. However, we show the mean of the spectra of the $\mathcal{O}$ mocks as opposed to the $\mathcal{V}$ mocks.

We employ the ($\mathcal{O}$-$P_\ell^{\rm obs}$) SLR test in two cases to measure the significance of the rejection of the LPP model by the $\mathcal{O}$ mocks. Firstly, we use the same binning as the mock-based tests in the previous Section to compute the S/N, we denote this result by ($\mathcal{O}$-$P_\ell^{\rm obs}$) in Tab. 6.2. In that case (see Fig. 6.7a), we find that the PDF of the likelihood ratio peaks far away from $\lambda = 0$ in all redshift bins reaching a maximum separation in the deepest one where the model is rejected in 82% of the realizations. However, as clear from Fig. 5.6, the model is unable to recover the $\mathcal{V}$ mock power spectra for modes smaller than the characteristic frequency. Consequently, we need to pinpoint the contribution of these modes in the S/N values reported in Fig. 6.7a.

To that end, we employ the ($\mathcal{O}$-$P_\ell^{\rm obs}$) SLR test in another case. Where, we deliberately exclude the first k-bin and only consider wavenumbers in the range $k \in [3.2, 36] \times 10^{-3}\,h\,{\rm Mpc}^{-1}$, which can be seen under *($\mathcal{O}$-$P_\ell^{\rm obs}$) in Tab. 6.2 as well as in Fig. 6.7b. We find that excluding these large-scale modes lowers the S/N values by a factor of ~ 1 on average. However, the resulting S/N still reaches a maximum of 2.9 in the deepest bin.

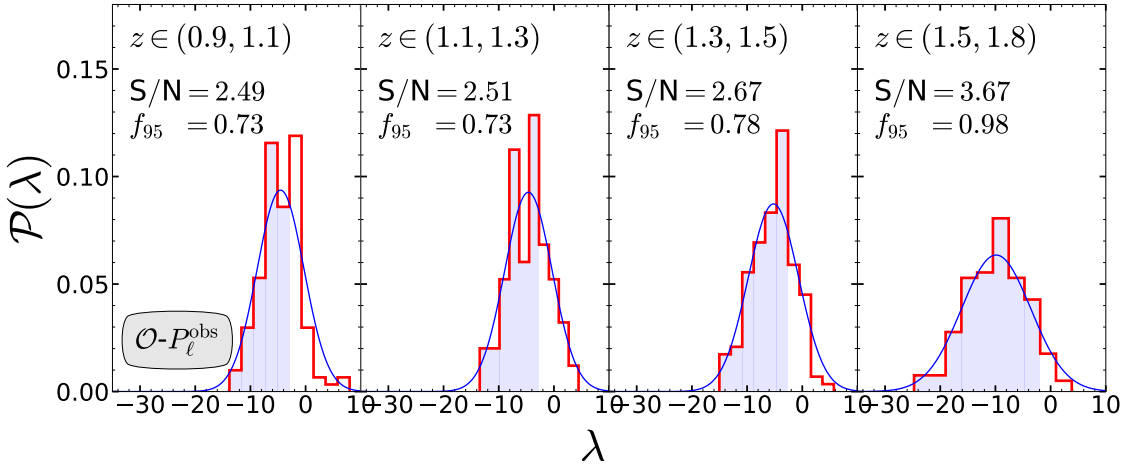
In conclusion, the results from both LPP Kaiser model tests are in good agreement with those deriving from the ($\mathcal{O}$-$\mathcal{V}$) SLR test. Which clearly outlines the importance of including relativistic RSD effects -especially weak-lensing- in the modelling of the *Euclid* power-spectrum multipoles on scales $k < 0.03\,h\,{\rm Mpc}^{-1}$.

6.3 Angular Power Spectrum

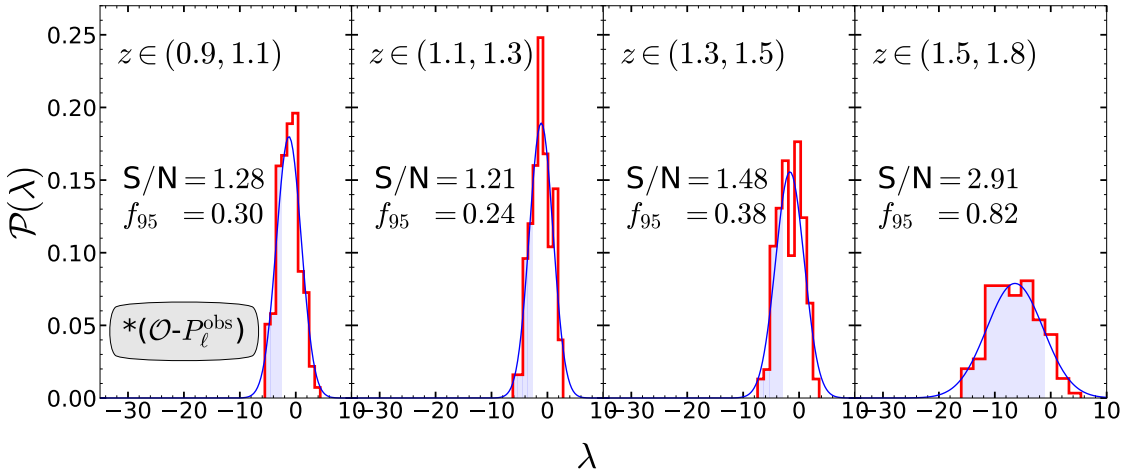
We then tackle the impact of relativistic RSD on the angular power spectrum. We use the PCL estimator (see Sec. 4.1) to compute the auto and cross-angular power spectra for all realizations of the $\mathcal{V}$, $\mathcal{G}$ and $\mathcal{O}$ mocks in the four tomographic redshift bins of the *Euclid* survey as well as the redshift bin that encompasses the entire survey. We estimate the spectra for both the *Euclid* and full-sky footprints shown in Fig. 3.5.

6.3.1 Velocity of the Observer

First of all, we analyze the velocity of the observer term. As seen in Eq. (2.181), the velocity of the observer leaves a dipolar signature on the galaxy density contrast (see also Gibelyou & Huterer, 2012, Maartens et al., 2018, Nadolny et al., 2021). We can measure the impact of this kinematic dipole by looking at the difference between the $\mathcal{O}$ and $\mathcal{G}$ angular spectra, i.e. $\Delta\hat{C}_\ell \equiv \hat{C}_{\ell,\mathcal{O}} - \hat{C}_{\ell,\mathcal{G}}$. In Fig. 6.8, we show the mean and RMS scatter of $\Delta\hat{C}_\ell$ for a full-sky survey (red) and a survey with the *Euclid* footprint (blue). The figure shows the difference for two different redshift bins: the left panel represents the auto-spectra in the redshift range $z \in (0.9, 1.1)$. while the lower panel includes the entire redshift range covered by the Euclid spectroscopic survey, namely $z \in (0.9, 1.8)$. We can immediately see that the observer velocity only affects $\ell = 1$ for the full-sky case, while it spreads across the odd multipoles ($\ell = 3, 5 \ldots$) for the *Euclid* type mask. This leakage to higher multipoles is due to



(a) ($\mathcal{O}$-$P_\ell^{\rm obs}$) SLR test.



(b) *($\mathcal{O}$-$P_\ell^{\rm obs}$) SLR test.

Figure 6.7: PDF of the likelihood-ratio test statistic for the LPP Kaiser model tests applied to the multipoles of the power spectrum.

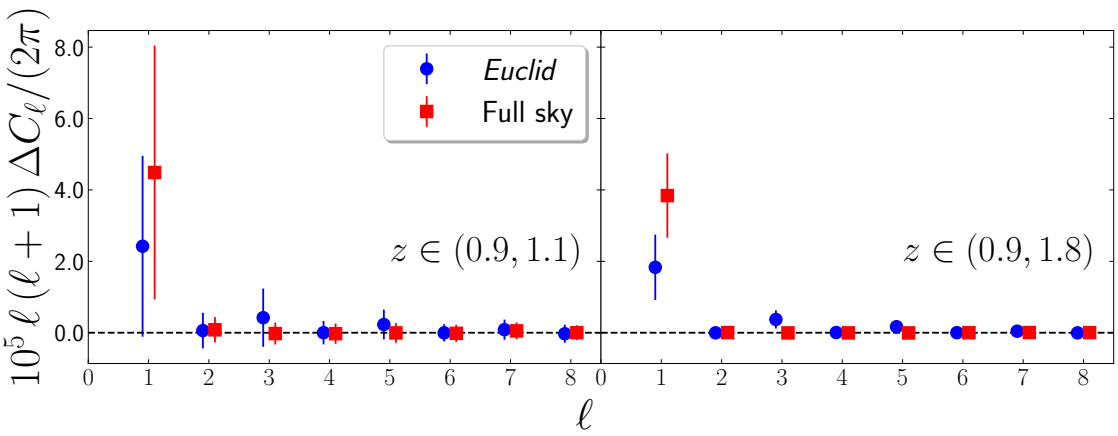


Figure 6.8: The difference between the angular power spectra of the $\mathcal{O}$ and $\mathcal{G}$ mocks is plotted for two angular masks: a full-sky mask (red) and the *Euclid* mask (blue). The symbols indicate the mean signal while the error bars show the RMS scatter. This plot is based on a similar figure from Paper III.

Table 6.3: The S/N results from the ($\mathcal{O}$-$\mathcal{G}$) SLR test on the angular power spectra. The signal-to-noise quantifies the significance of the velocity of the observer term in the $\ell \in [1, 10]$ modes.

$(z_{\rm min}, z_{\rm max})$	$(0.9, 1.1)$	$(1.1, 1.3)$	$(1.3, 1.5)$	$(1.5, 1.8)$	$(0.9, 1.8)$
$(0.9, 1.1)$	0.9				
$(1.1, 1.3)$	1.3	0.7			
$(1.3, 1.5)$	1.1	1.1	0.8		
$(1.5, 1.8)$	1.2	1.1	1.2	1.0	
$(0.9, 1.8)$	1.4	1.4	1.5	1.6	2.1

the partial coverage in the sky as the spherical harmonics are no longer orthogonal and consequently power from the full-sky dipole leaks to higher odd multipoles.

We perform the ($\mathcal{O}$-$\mathcal{G}$) SLR test (see Sec. 4.4.1) to quantify the significance of the kinematic dipole. We restrict the data vector to the first 10 multipoles of the $\mathcal{O}$ angular power spectra $\ell \in [1, 10]$. The resultant S/N values for all the considered redshift bins are shown in Tab. 6.3 for the *Euclid* footprint. We can see that the S/N is small for the tomographic bins, however, it is clear that the velocity of the observer signal becomes more prominent in the case of the broad redshift bin $z \in (0.9, 1.8)$ with a S/N = 2.1.

6.3.2 Weak Gravitational Lensing

We then turn our attention to the weak lensing effects. As seen in Fig. 6.8, the difference between the $\mathcal{O}$ and $\mathcal{G}$ angular spectra is minimal beyond $\ell > 5$. Consequently, we directly compare the $\mathcal{O}$ and $\mathcal{V}$ mocks instead of separately testing the

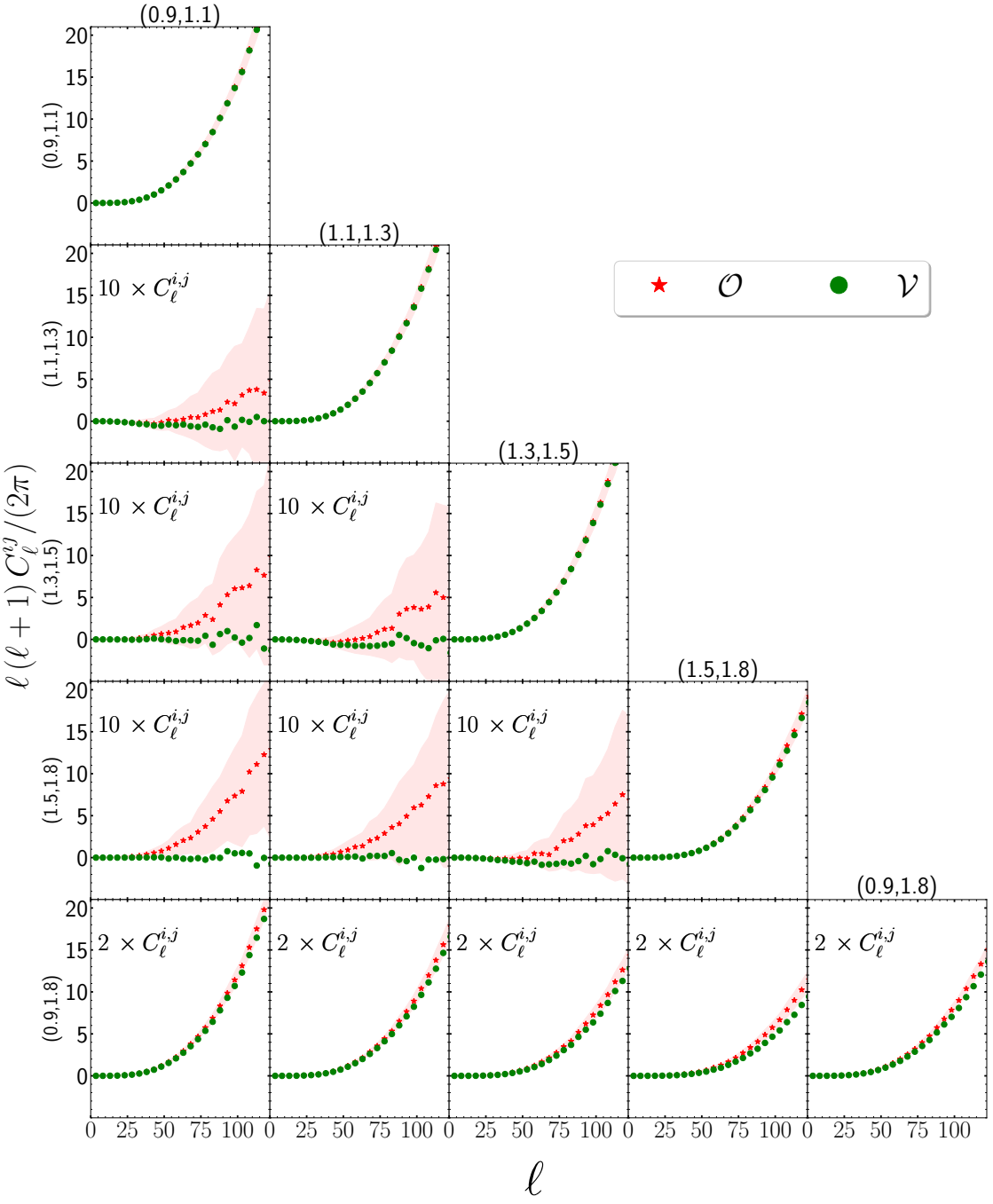


Figure 6.9: We plot the mean of the angular power spectra of the $\mathcal{O}$ mocks in blue dots while the mean of the spectra of the $\mathcal{V}$ mocks is shown in green. The pink shaded regions denote the standard deviation of the spectra of the $\mathcal{O}$ mocks. We show the auto-correlation spectra of the redshift bins on the diagonal while the cross-correlations between the different bins are shown on the off-diagonal cells. We note that the cross-correlation spectra (the last row) are multiplied by 10(2) to increase the clarity of the figure. This figure is taken from Paper III.

$\mathcal{O}$ and $\mathcal{G}$ against the $\mathcal{V}$ spectra.

We plot the mean of the spectra of the $\mathcal{O}$ mocks in red dots, while the $\mathcal{V}$ mocks are shown in green in Fig. 6.9. The salmon-shaded regions surrounding the red dots indicate the standard deviation of the $\mathcal{O}$ measurements. The auto-correlation results are shown on the diagonal, while the cross-correlation spectra are displayed in the off-diagonal cells. The bottom row shows the cross and auto-correlation results for the redshift bin containing the entire survey. As can be seen in Fig. 6.9, the $\mathcal{O}$ spectra are significantly enhanced with respect to the $\mathcal{V}$ spectra, both in the auto-correlation of the largest bin and in the cross-correlation of well-separated bins, e.g. the cross-correlation between $z \in (1.5, 1.8)$ and $z \in (0.9, 1.1)$. The non-vanishing cross-correlation in the $\mathcal{O}$ spectra is due to the integral terms in Eq. (2.181), in particular to the dominant weak lensing term. In the auto-correlation, it is clear that the weak-lensing effects are significant only in the highest redshift bin, as well as in the redshift bin that encompasses the entire survey.

Finally, we quantify the importance of the weak lensing term using the ($\mathcal{O}$-$\mathcal{V}$) SLR test. We resample the spectra into bins with $\Delta\ell = 5$ in the range $\ell \in (1, 120)$ as the Music simulations underestimate the angular power spectra beyond that (see Sec. 3.5). Consequently, we compute the model means and covariances with 24 bins. We show the S/N values obtained with the ($\mathcal{O}$-$\mathcal{G}$) SLR test in Fig. 6.10. The detection of the lensing term is highly significant in the cross-correlations of well-separated tomographic redshift bins and in all statistics involving the wide bin $z \in (0.9, 1.8)$. In fact, multiple panels demonstrate that all $\mathcal{O}$ mocks reject the $\mathcal{V}$ model at the 95% confidence level.

6.4 Two-point Correlation Function Multipoles

The last summary statistic we consider in this chapter is the 2PCF. We utilise the LS estimator (see Sec. 4.2) to measure the 2PCF multipoles for all 140 realizations of the different types of mocks $\{\mathcal{R}, \mathcal{V}, \mathcal{G}, \mathcal{O}\}$ in 4 tomographic redshift bins between the points $z \in \{0.9,\ 1.1,\ 1.3,\ 1.5,\ 1.8\}$. We estimate the 2PCF multipoles in 500 linearly-spaced bins in the range $r \in [20,\ 350]\, h^{-1}\,\mathrm{Mpc}$.

The 2PCF multipoles are shown in Fig. 6.11 . In the figure, we plot the mean of the estimated multipoles after rebinning to bins of size $\Delta r = 15\, h^{-1}\,\mathrm{Mpc}$. We show the mean of estimated $\hat{\xi}_\ell$ of the $\mathcal{R}$ (black), $\mathcal{V}$ (green), $\mathcal{G}$ (blue) and $\mathcal{O}$ (red) mocks in solid lines while the shaded region is the standard deviation computed from the multipoles of the $\mathcal{O}$ mocks. The real-space monopole moment is positive at small separations. At $r \simeq 100\, h^{-1}\,\mathrm{Mpc}$, we can see the baryonic-acoustic oscillation feature. We then note anticorrelation for $r \gtrsim 125\, h^{-1}\,\mathrm{Mpc}$ while $\hat{\xi}_2$ and $\hat{\xi}_4$ are consistent with zero as expected. RSD boost the monopole signal as well as generate a negative quadropole and a relativily small positive hexadecapole signal (see also Sec. 2.2.2).

We first concentrate on the observer velocity signal. However, the differences between the $\mathcal{G}$ and $\mathcal{O}$ 2PCF multipoles are difficult to detect in Fig. 6.11. To clearly

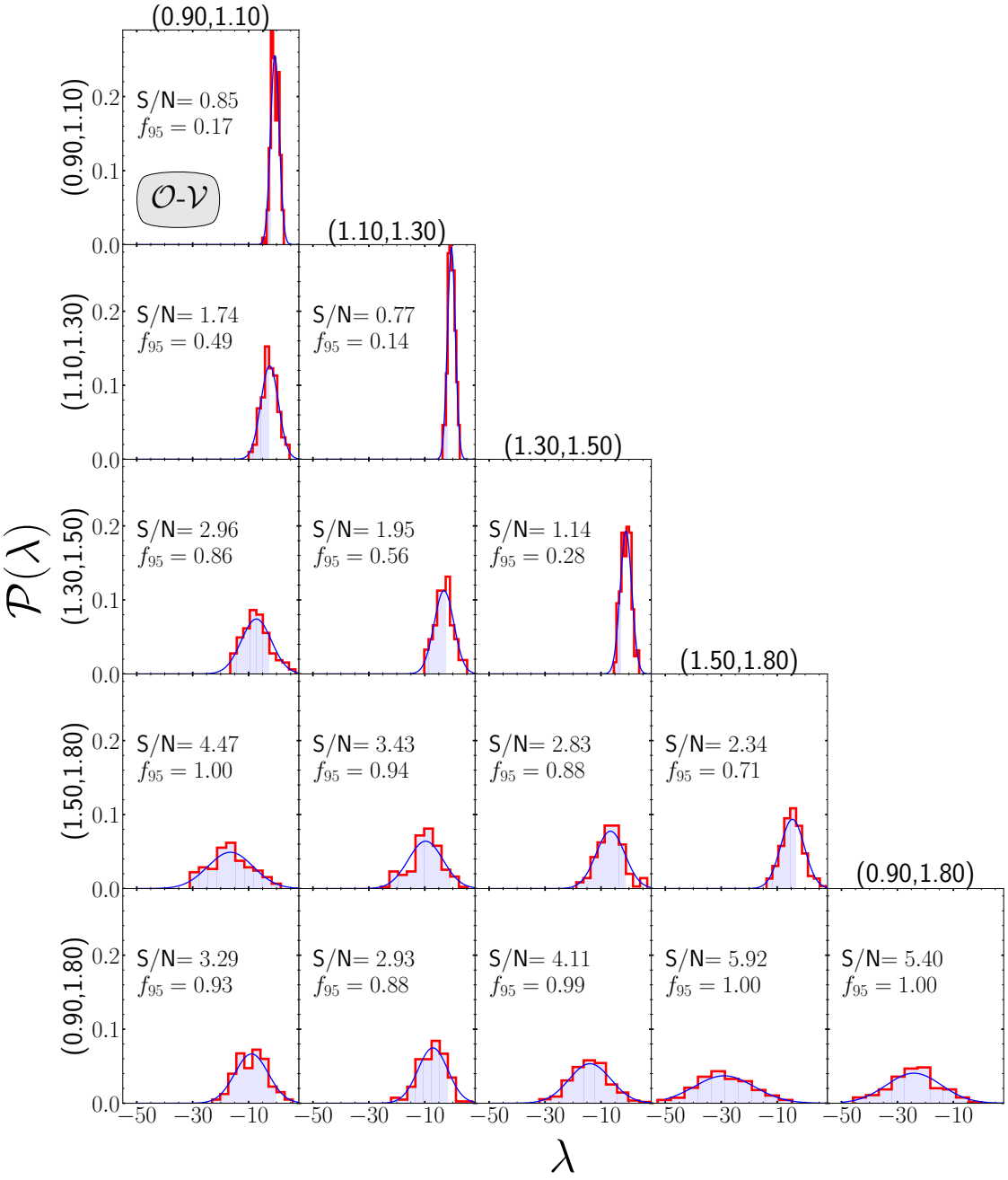


Figure 6.10: PDF of the likelihood-ratio test statistic for the ($\mathcal{O}$-$\mathcal{G}$) SLR test applied to the angular auto and cross spectra.

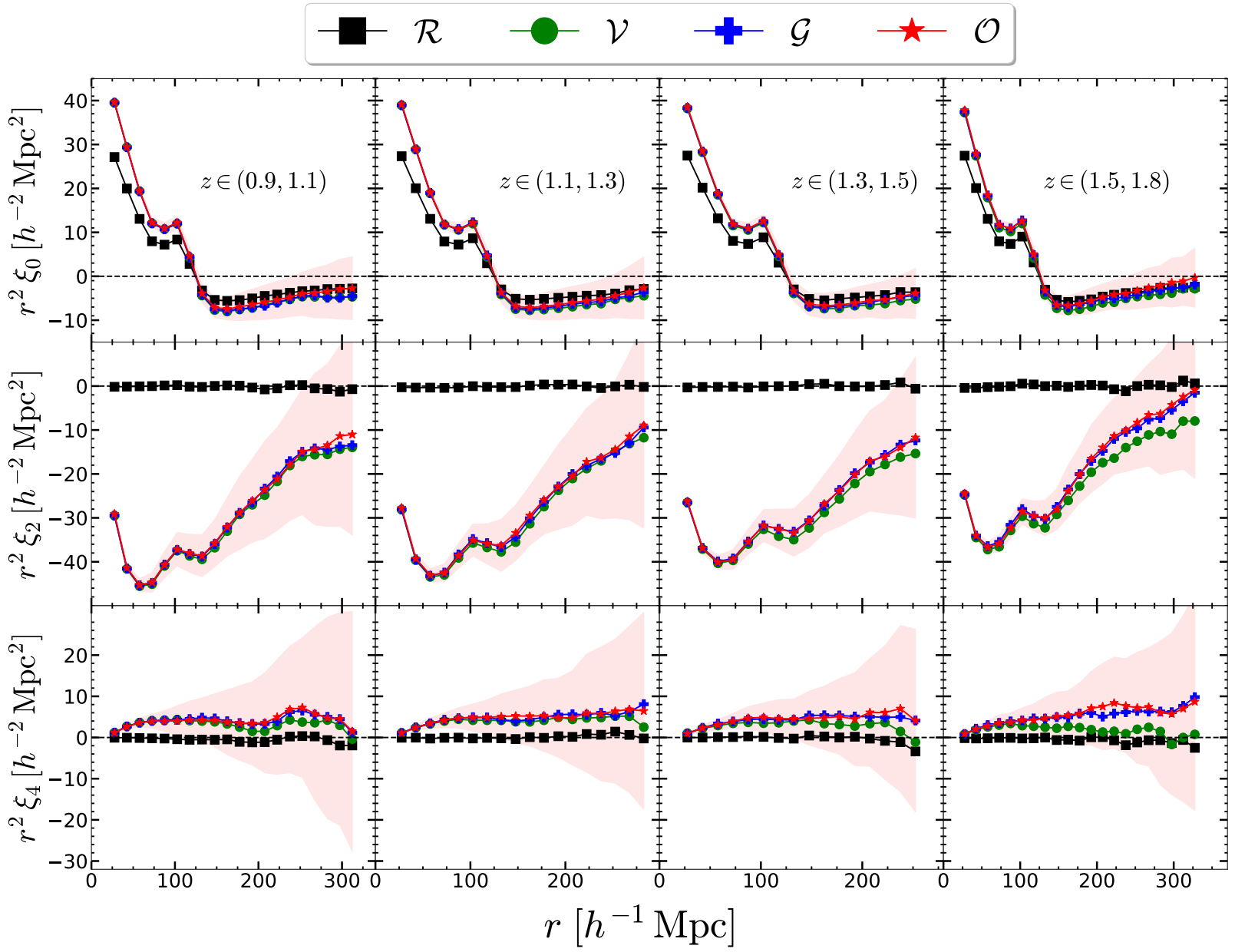


Figure 6.11: We show the mean of the 2PCF multipoles estimated from the $\mathcal{R}$ (black squares), $\mathcal{V}$ (green circles), $\mathcal{G}$ (blue crosses) and $\mathcal{O}$ (red stars) mocks. The salmon-shaded region denotes the RMS scatter estimated from the $\mathcal{O}$ multipoles. Each row shows one of the even multipole of the 2PCF while the different columns denote the different tomographic redshift bins considered.

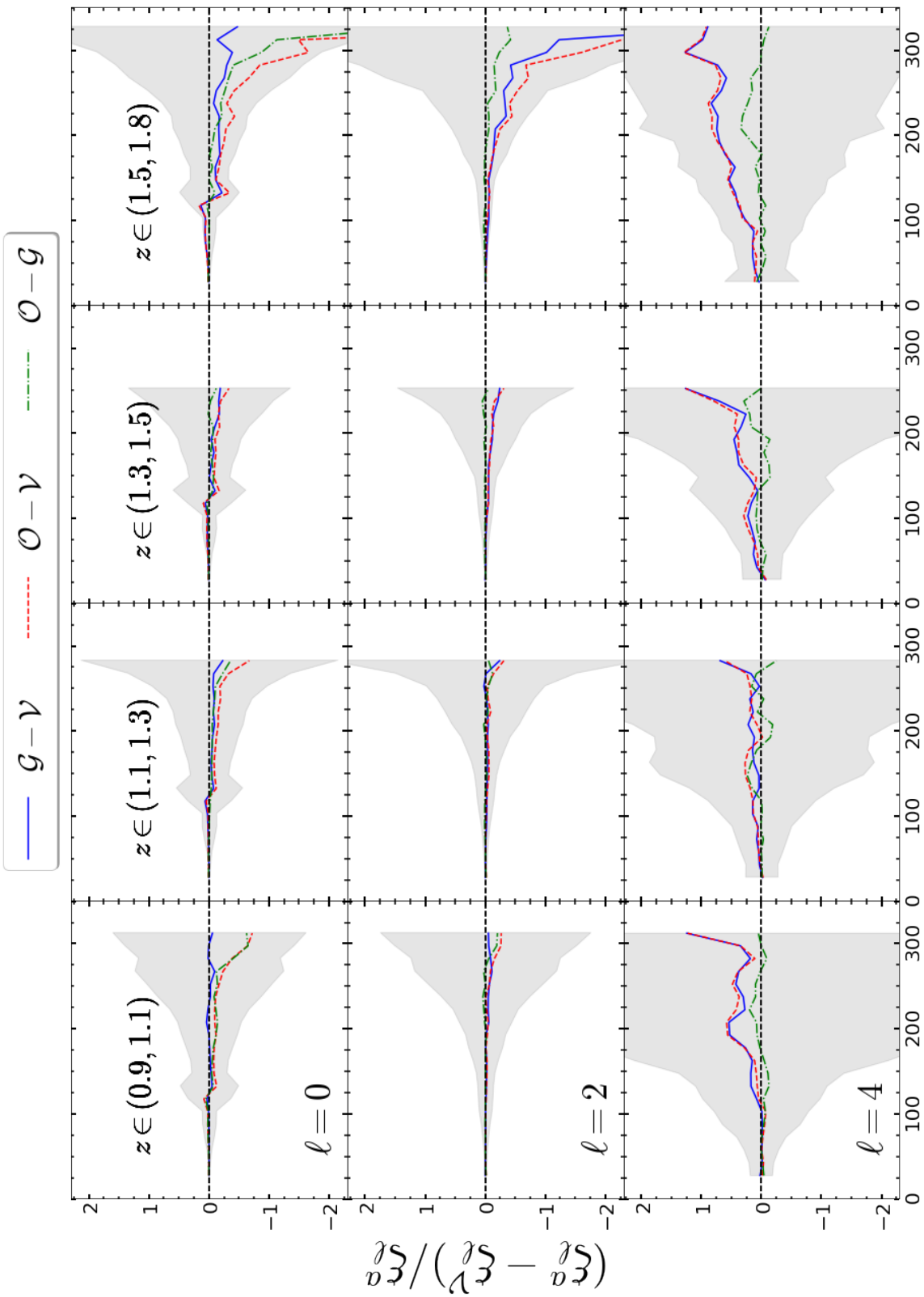


Figure 6.12: We present the fractional differences between the means of the $\mathcal{G}$ and $\mathcal{V}$ 2PCF multipoles using blue solid lines, the $\mathcal{O}$ and $\mathcal{V}$ 2PCF multipoles with red dashed lines, and the $\mathcal{O}$ and $\mathcal{G}$ 2PCF multipoles with green dash-dotted lines. Shaded regions represent the ratio between the standard deviation and the mean $(\sigma_{\xi_\ell}/\xi_\ell)$ of the $\mathcal{V}$ mocks. Each row corresponds to a different even multipole, while each column represents a different redshift bin.

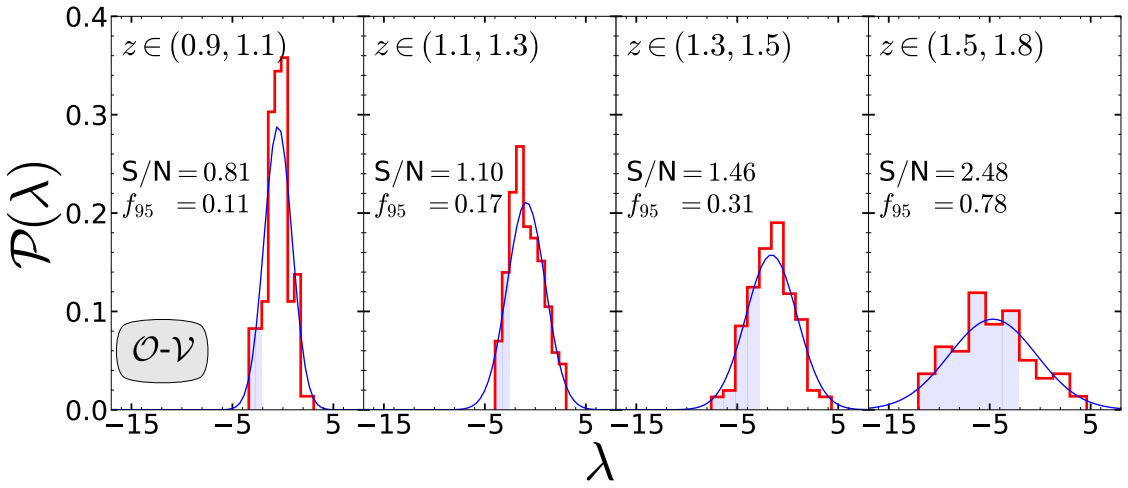


Figure 6.13: $\mathcal{P}(\lambda)$ for the ($\mathcal{O}$-$\mathcal{V}$) SLR test applied to 2PCF multipoles. This figure is taken from Paper III

observe the peculiar velocity signal, we show the difference between the means of the multipoles obtained from the $\mathcal{G}$ and $\mathcal{O}$ mocks in Fig. 6.12. The impact of the observer terms is represented by the green dash-dotted lines. It is evident that the monopole is the most affected, especially in the deepest redshift bin, as seen in the top row of Fig. 6.12. However, the higher multipoles seem to exhibit little to no signal due to the velocity of the observer compared to the measurement error (shown by the shaded regions).

We quantify the significance of the effect using the ($\mathcal{O}$-$\mathcal{G}$) SLR test. To that end, we bin the 2PCF multipoles $\xi_\ell(r)$ into 10 bins each, in the range $r \in [35\,h^{-1}\,\mathrm{Mpc}, \Delta x]$, where Δx is the width of the redshift bin. We then combine the multipoles to create a $n = 30$ data vector.

The results of the likelihood ratio test corroborate the visual inspection of Fig. 6.12 as it consistently gives values of S/N $\leq$ 1 (see Tab. 6.4) for all redshift bins. We note that Bertacca et al. (2020) finds a stronger impact due to the velocity of the observer term, however, this study employs larger separations and the BP convention for the 2PCF coordinate system (which changes the multipoles see Sec. 2.2.2, Raccanelli et al., 2014, Reimberg et al., 2016).

6.4.1 Weak Gravitational Lensing

Upon examining the difference between the multipoles of the $\mathcal{G}$ and $\mathcal{V}$ mocks depicted as blue solid lines in Fig. 6.12, we observe that the signal strength of all $\mathcal{G}$ multipoles increases with the separation distance r and the mean redshift. However, it is evident that the RMS scatter, represented by the shaded regions, also exhibits a significant increase with r. Furthermore, we notice that the discrepancy between the $\mathcal{G}$ and $\mathcal{V}$ mocks becomes more apparent for higher multipoles compared to the monopole.

This difference reaches its maximum in the deepest redshift bin, as expected

Table 6.4: S/N from the ($\mathcal{O}$-$\mathcal{V}$) ,($\mathcal{O}$-$\mathcal{G}$), and ($\mathcal{G}$-$\mathcal{V}$) tests for the 2PCF multipoles.

$(z_{\rm min}, z_{\rm max})$	$\mathcal{O}-\mathcal{G}$	$\mathcal{G}-\mathcal{V}$	$\mathcal{O}-\mathcal{V}$
$(0.9, 1.1)$	0.6	0.5	0.8
$(1.1, 1.3)$	0.7	0.6	1.1
$(1.3, 1.5)$	0.6	1.2	1.5
$(1.5, 1.8)$	1.0	2.1	2.5

due to the weak lensing sourced effect. From the behaviour of the monopole (solid blue line) and the quadrupole (dashed blue line) in that bin, we note that the weak-lensing terms affect the signal for scales greater than $r > 200\,h^{-1}\,$Mpc. The weak-lensing signal affects smaller separations in the hexadecapole. However, for all multipoles the increased signal is still within the scatter of the measurements.

Performing the ($\mathcal{G}$-$\mathcal{V}$) SLR test, we find that the S/N steadily grows from 0.5 to 2.1 from the lowest to the highest redshift bin as shown in Tab. 6.4.

6.4.2 All relativistic RSD

Finally, we consider the combined contribution of the observer velocity and the Weak-lensing signal using the ($\mathcal{O}$-$\mathcal{V}$) SLR test. We show the results in Fig. 6.13. We can see that $\mathcal{P}(\lambda)$ peaks at increasingly negative values as the mean redshift increases. In the deepest redshift bin, the likelihood-ratio test manages to reject the velocity-only model for RSD at the 95% confidence level in 78% of our mock catalogues. We quantify the S/N using Eq. (4.29) finding that the S/N = 2.5 for the highest redshift bin.

It is instructive to compare the S/N values obtained from the 2PCF multipoles to their power spectrum counterparts. The ($\mathcal{G}$-$\mathcal{V}$) SLR results reported in Tables 6.4 and 6.2 indicate that the lensing contribution similarly impacts both statistics. However, it is clear that the velocity of the observer is more prominent in the power spectrum case. The increase in S/N is caused by the observer-velocity-induced boost in the power spectrum multipoles (see Fig. 6.4) at very large scales as compared to the relatively minor contribution in the 2PCF shown in Fig. 6.12. It is important to note that the scales probed by the 2PCF multipoles are limited by the thickness of the redshift bins. By using a larger redshift bin, we access larger scales, which could result in a more pronounced impact of the observer's velocity (Bertacca, 2020).

6.5 Summary

In this chapter, we answered the first main objective of this thesis. Namely, we quantitatively measured the importance of different relativistic RSD in the *Euclid* survey. Our results can summarized as follows.

(i) Power spectrum monopole: In Sec. 6.1, we studied the impact of relativistic RSD on the power spectrum monopole. Firstly, we found that wide-angle correc-

tions arising from peculiar velocities become significant, leading to a statistically notable enhancement of power for $k < 0.02, h$ Mpc^{-1} compared to the GPP Kaiser model. Consequently, modelling these corrections becomes imperative when considering cosmological inferences at such large scales. Secondly, the influence of weak lensing, remains minor, given that existing models for the luminosity function of Hα emitters adequately capture the faint-end slope. However, for redshifts $z > 1.5$ and broad redshift bins, the weak-lensing effect systematically boosted the power-spectrum monopole by a factor comparable to the statistical uncertainties.

(ii) Multipole moments of the power spectrum: For the velocity of the observer, we have shown, in Sec. 6.2, that it induces a strong distortion in the amplitude of all multipoles for ($k < 0.015\,h\,\mathrm{Mpc}^{-1}$) which has similar amplitudes across the different redshift bins. The statistical test has shown that relativistic RSD are detectable with a S/N = 2.4 in the deepest redshift bin (see Table 6.2). We found that the weak lensing boosts the signal in the power spectrum monopole at large scales ($k < 0.02\,h\,\mathrm{Mpc}^{-1}$) as seen in Fig. 6.4. As for the higher multipoles, we found that the weak lensing affects the spectra up to relatively smaller scales ($k < 0.03\,h\,\mathrm{Mpc}^{-1}$), in particular, for the hexadecapole. Finally, we compared the spectra of the $\mathcal{O}$ mocks to the LPP Kaiser model, where we found, in the redshift bin $1.5 < z < 1.8$, the model is rejected with a S/N = 2.9 (3.7) by the $\mathcal{O}$ mocks depending on the adopted k-range. This result reinforces the conclusion that the weak lensing term cannot be disregarded at high redshift and on large scales.

(iii) Angular power spectrum: In Sec. 6.3, we found that the velocity of the observer induces a boost in the first few odd multipoles of the angular power spectrum as shown in Fig. 6.8. That boost is measured with S/N = 2.1 for the broadest bin in the *Euclid* survey. As for the weak lensing signal, it slightly increased the clustering signal for $\ell \gtrsim 50$ in the angular spectra. This increase was particularly evident in the broad redshift bin $0.9 < z < 1.8$ where we find S/N = 5.4 and in the cross-correlations of well-separated bins where the signal-to-noise reached S/N = 4.5. It is clear that the optimal strategy for measuring the lensing signal is using cross-correlations between well-separated bins. Consequently, one possible strategy to boost the overall lensing signal is to change the binning strategy, creating a larger number of thinner bins ($\Delta z = 0.01$) and then extracting the lensing signal from their cross-correlations. We leave that to future work.

(iv) Multipole moments of the 2PCF: The observer's peculiar velocity introduced slight distortions to the clustering signal, with the likelihood-ratio test consistently yielding a signal-to-noise ratio lower. At larger scales ($r > 150\,h^{-1}\,\mathrm{Mpc}$) and redshifts exceeding $z > 1.3$, the weak lensing contribution enhanced the quadrupole and hexadecapole moments of the 2PCF, as shown in Sec. 6.4. Moreover, the S/N of the combined contributions of all relativistic RSD increased with redshift, reaching 2.5 in the most distant tomographic bin considered ($1.5 < z < 1.8$). We note that similar values of S/N were found in a similar analysis in Euclid Collaboration: Jelic-Cizmek et al. (2023) for the *Euclid* survey, where it was shown that the weak lensing can lead to $0.4 - 0.7$ bias in the cosmological parameter constraints.

There are several intriguing extensions of the research presented in this chapter.

For example, estimating the impact of relativistic effects on the constraints of hyper-scale parameters, such as primordial non-Gaussianity and the matter-radiation equality scale, is of significant interest. Multiple studies have indicated that the boosted signal measured in the power spectrum multipoles can be erroneously interpreted as primordial non-Gaussianity (Raccanelli et al., 2018, Bahr-Kalus et al., 2021, Noorikuhani & Scoccimarro, 2023). Furthermore, exploring degeneracies with the constraints of the matter-radiation equality scale can prevent possible biases in the measurements of the neutrino mass sum from clustering statistics (Lesgourgues & Pastor, 2006).

Secondly, another approach is to estimate statistics dominated by relativistic RSD effects. For instance, we could analyze the odd multipoles of the 2PCF and the power spectrum (e.g., see Bertacca, 2020, Beutler & Di Dio, 2020, see also McDonald & Seljak, 2009, Bonvin et al., 2016, Hall & Bonvin, 2017, Bonvin & Fleury, 2018 for tests of the gravitional theory based on the odd multipoles of the 2PCF of different populations), the angular cross-correlation of different populations (Borzyszkowski et al., 2017), the cross-correlations between the large-scale structure and the cosmic microwave background (Renk et al., 2016), or the galaxy–gravitational wave cross-correlations (Bosi et al., 2023).

Additionally, a novel approach used in Abidi et al. (2023) constructs an estimator for the $E_{\rm G}$ statistic (first proposed in Zhang et al., 2007), which compares the evolution of the sum of the potentials with that of the velocity, using only clustering information obtained from two tracers of the large-scale structure.

Chapter 7

The Finger of the Observer Effect

This chapter is dedicated to exploring the phenomenon known as the FOTO effect. The FOTO effect refers to the enhancement observed in the power spectrum multipoles arising from the peculiar velocity of the observer. While we provided a brief discussion of the significance of this effect during the study of relativistic RSD in the *Euclid* survey in Section 6.2, this section delves deeper into its significance and possible uses for cosmological analysis.

We start by performing a comprehensive numerical investigation into the FOTO effect, focusing specifically on its impact on the monopole of the power spectrum. Through the $\mathcal{O}$ and $\mathcal{G}$ mocks, we aim to clarify the underlying mechanism driving this phenomenon and quantify its significance in the context of cosmological studies varying both the observer velocities and the fraction of the sky covered by the survey. Furthermore, we derive an analytical expression capturing the behaviour of the FOTO effect, providing valuable insights into its dependence on various parameters such as redshift, Hubble expansion, and observer velocity. By providing a theoretical framework for the FOTO effect, we aim to understand how it manifests in observational data.

Moreover, we explore the potential utility of the FOTO effect as a cosmological probe, extending beyond its role as a systematic bias effect. We demonstrate how measurements of the FOTO effect can yield valuable information regarding the magnitude and direction of the observer's peculiar velocity, as well as offer a unique avenue for probing fundamental cosmological parameters such as the matter density parameter. By leveraging the FOTO effect in conjunction with other observational constraints, we aim to fulfil the second objective of this thesis: utilising relativistic RSD in cosmological parameter inference.

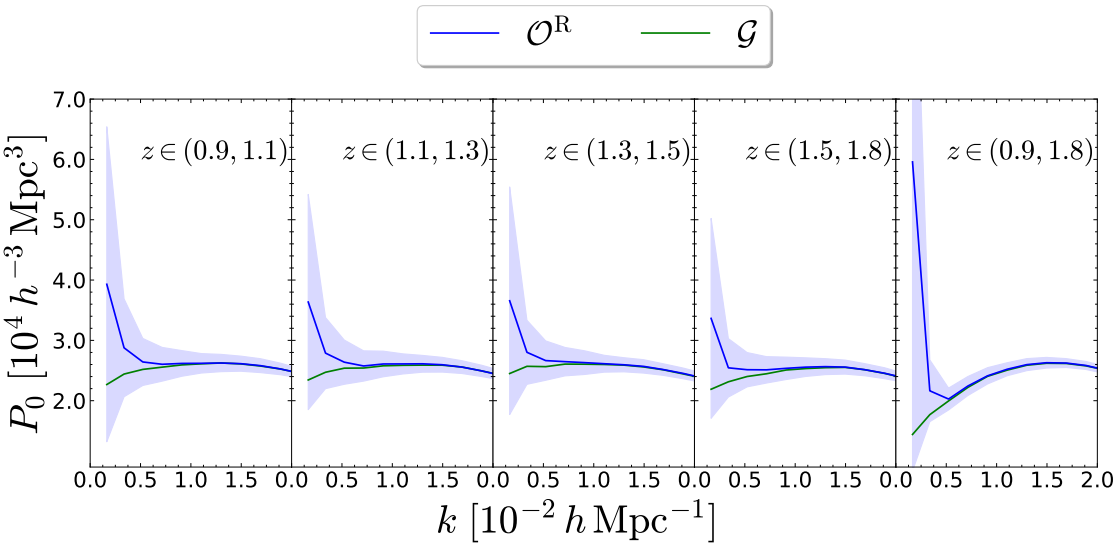


Figure 7.1: The solid curves show the mean monopole spectra estimated from the $\mathcal{O}^{\rm R}$ (blue) and $\mathcal{G}$ (green) mocks. The shaded regions indicate the RMS scatter of the $\mathcal{O}^{\rm R}$ spectra. This figure is based on a similar figure in Paper I.

7.1 Impact on the Monopole

We start with an example of the FOTO signal. To that end, we use the FKP estimator to compute the power spectrum monopole for all realizations of the $\mathcal{O}^{\rm R}$ and $\mathcal{G}$ mocks for the *Euclid* survey using the *Euclid* footprint (see Sec. 3.4.2). In Fig. 7.1, we plot the mean of the estimated spectra from the $\mathcal{O}^{\rm R}$ and $\mathcal{G}$ mocks in blue and green solid lines respectively. As seen in the Figure, the $\mathcal{O}^{\rm R}$ monopole spectra have substantially more power compared to their $\mathcal{G}$ counterpart at small k, i.e. $k < 0.5 \times 10^{-2}\,h\,{\rm Mpc}^{-1}$.

In order to derive the source of the signal shown in Fig. 7.1, let us start from Eq. (2.184)

$$\delta_{\rm obs}(\boldsymbol{x}) = \delta_{\rm com}(\boldsymbol{x}) + \frac{\alpha_{\rm o}(x)}{x}\,\frac{\boldsymbol{v}_{\rm o}\cdot\hat{\boldsymbol{x}}}{a\,H}\,, \tag{7.1}$$

where $\delta_{\rm obs} = \delta_{\rm g,s}$, the 'comoving redshift-space overdensity' $\delta_{\rm com}$ collects all terms that do not depend on the observer velocity and

$$\alpha_{\rm o}(x) = 2(1-\mathcal{Q}) - \frac{x\,H}{c\,(1+z)}\,\mathcal{E} + \frac{x\,H}{c\,(1+z)}\left[3 - \frac{1+z}{H}\,\frac{{\rm d}H}{{\rm d}z}\right]. \tag{7.2}$$

We can measure the impact of the dipole term in Eq. (7.2) on the monopole of the observed power spectrum as follows. Starting from Eq. (5.3), let us set $\ell = 0$ and compute the power spectrum without the angular average over ${\rm d}\Omega_k$ for the

moment. Ignoring the Integral constraint contribution, we compute

$$\hat{P}_0(\boldsymbol{k}) = \int W(\boldsymbol{x}_1)W(\boldsymbol{x}_2)\langle\delta_{\rm obs}(\boldsymbol{x}_1)\,\delta_{\rm obs}(\boldsymbol{x}_2)\rangle \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_2-\boldsymbol{x}_1)}\mathrm{d}^3x_1\;\mathrm{d}^3x_2\,. \tag{7.3}$$

We can assume that the cross-correlations between the velocity of the observer and $\delta_{\rm cos}$ are negligible if we exclude the local Universe ($z < 0.05$) from the analysis (Nadolny et al., 2021). We can thus directly see that the observed power spectrum $P_{0,\rm obs}$ can be related to the comoving power spectrum $P_{0,\rm com}$ through

$$P_0^{\rm obs}(\boldsymbol{k}) = P_0^{\rm com}(\boldsymbol{k}) + \frac{1}{A}\int B_1\,B_2\,(\boldsymbol{v}_{\rm o}\cdot\hat{\boldsymbol{x}}_1)\,(\boldsymbol{v}_{\rm o}\cdot\hat{\boldsymbol{x}}_2)\,e^{i\boldsymbol{k}\cdot(\boldsymbol{x}_2-\boldsymbol{x}_1)}\,\mathrm{d}^3x_1\,\mathrm{d}^3x_2\,, \tag{7.4}$$

$$\equiv P_0^{\rm com}(\boldsymbol{k}) + P_{\rm dip,0}^{\rm AN}(\boldsymbol{k})\,, \tag{7.5}$$

where $B_i = \bar{n}_i\,\alpha_{{\rm o},i}/(x_i\,a_i\,H_i)$ and $A = \int \bar{n}^2(x)\,\mathrm{d}^3x$.[1] There is no ensemble average in the definition $P_{\rm dip,0}^{\rm AN}(\boldsymbol{k})$ as there are no stochastic terms within the integral. We use the superscript AN in $P_{\rm dip,0}^{\rm AN}(\boldsymbol{k})$ to articulate that the result does not represent the angular averaged power spectrum monopole (it excludes the $\int \mathrm{d}\Omega_k/(4\pi)$).

In order to compute the dipole power spectrum $P_{\rm dip,0}^{\rm AN}(\boldsymbol{k})$, we separate the integrals in $P_{\rm dip,0}^{\rm AN}(\boldsymbol{k})$ and rewrite them as

$$P_{\rm dip,0}^{\rm AN}(\boldsymbol{k}) = \frac{1}{A}\left[\int B\,(\boldsymbol{v}_{\rm o}\cdot\hat{\boldsymbol{x}})e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\mathrm{d}^3x\right]\left[\int B(\boldsymbol{v}_{\rm o}\cdot\hat{\boldsymbol{x}})\,e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\;\mathrm{d}^3x\right]. \tag{7.6}$$

We solve the previous relation by first calculating the angular integrals,

$$\int_{4\pi}(\hat{\boldsymbol{v}}_{\rm o}\cdot\hat{\boldsymbol{x}})\,e^{\pm i\boldsymbol{k}\cdot\boldsymbol{x}}\,\mathrm{d}\Omega_x \equiv \int_{4\pi}\mathcal{L}_1\,(\hat{\boldsymbol{v}}_{\rm o}\cdot\hat{\boldsymbol{x}})\,e^{\pm i\boldsymbol{k}\cdot\boldsymbol{x}}\,\mathrm{d}\Omega_x\,, \tag{7.7}$$

over the solid angles spanned by $\boldsymbol{x}$. To that end, we first expand the exponential function into plane waves

$$e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = \sum_\ell (2\ell+1)\,i^\ell\,j_\ell(kx)\,\mathcal{L}_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}})\,. \tag{7.8}$$

Assuming full-sky coverage, we use the following identity of the Legendre polynomials,

$$\int \mathcal{L}_\ell(\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{x}})\mathcal{L}_m(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}})\,\mathrm{d}\Omega_x = \frac{4\pi}{2\ell+1}\,\delta^{\rm K}_{\ell m}\,\mathcal{L}_m(\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{k}})\,, \tag{7.9}$$

to obtain

$$\begin{aligned}\int_{4\pi}(\hat{\boldsymbol{v}}_{\rm o}\cdot\hat{\boldsymbol{x}})\,e^{\pm i\boldsymbol{k}\cdot\boldsymbol{x}}\,\mathrm{d}\Omega_x &= \sum_\ell (2\ell+1)\,(\pm i)^\ell\,j_\ell(kx)\int_{4\pi}\mathcal{L}_1\,(\hat{\boldsymbol{v}}_{\rm o}\cdot\hat{\boldsymbol{x}})\,\mathcal{L}_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}})\,\mathrm{d}\Omega_x \\ &= \pm 4\pi\,i\,j_1(kx)\,(\hat{\boldsymbol{v}}_{\rm o}\cdot\hat{\boldsymbol{k}})\,.\end{aligned} \tag{7.10}$$

[1] we set the FKP weights $w = 1$ in this derivation as they only affect the covariance not the signal by construction (Feldman et al., 1994).

Therefore, the anisotropic dipole power spectrum is given by

$$P^{\rm AN}_{\rm dip,0}(\boldsymbol{k}) = 16\pi^2\,(\boldsymbol{v}_{\rm o}\cdot\hat{\boldsymbol{k}})^2\,\frac{I^2(k)}{\int \bar{n}^2(x)\,{\rm d}^3x}\,, \tag{7.11}$$

where

$$I(k) = \int \frac{x\,\bar{n}\,\alpha}{a\,H}\,j_1(kx)\,{\rm d}x\,. \tag{7.12}$$

For a narrow shell of thickness Δx at distance x, we get

$$P^{\rm AN,T}_{\rm dip,0}(\hat{k}) \simeq 4\pi\,\frac{(\boldsymbol{v}_{\rm o}\cdot\hat{\boldsymbol{k}})^2}{a^2H^2}\,\alpha_{\rm o}^2\,j_1^2(kx)\Delta x\,, \tag{7.13}$$

where the T super script marks the thin shell approximation.

It is clear that $P^{\rm AN}_{\rm dip,0}$ is anisotropic as we have not calculated the angular average over all directions of $\hat{\boldsymbol{k}}$ (see Eq. 5.3). Doing so, we find

$$P_{\rm dip,0}(k) = \int P^{\rm AN}_{\rm dip,0}(\boldsymbol{k})\,\frac{{\rm d}\Omega_k}{4\pi} = \frac{16\pi^2}{3}\,v_{\rm o}^2\,\frac{I^2(k)}{\int \bar{n}^2(x)\,{\rm d}^3x}\,. \tag{7.14}$$

We could also average $\boldsymbol{v}_{\rm o}^2$ over an ensemble of observers which results in

$$\langle P_{\rm dip,0}(k)\rangle = 16\pi^2\,\sigma_v^2\,\frac{I^2(k)}{\int \bar{n}^2(x)\,{\rm d}^3x}\,, \tag{7.15}$$

where σ_v is the one-dimensional peculiar velocity dispersion of the observers. Similarly for the anisotropic signal,

$$\langle P^{\rm AN}_{\rm dip,0}(\boldsymbol{k})\rangle = 48\pi^2\,\sigma_v^2\,\frac{(\hat{\boldsymbol{v}}_{\rm o}\cdot\hat{\boldsymbol{k}})^2}{H_0^2}\,\frac{I^2(k)}{\int \bar{n}^2(x)\,{\rm d}^3x}\,, \tag{7.16}$$

where we assume that the three-dimensional velocity dispersion is a factor $\sqrt{3}$ larger than its one-dimensional counterpart (i.e. that $\hat{\boldsymbol{v}}_{\rm o}$ is isotropically distributed).

We are then able to predict the FOTO signal for both our example surveys; the *Euclid* and SKAO2 surveys in the full-sky footprint case by computing Eqs. (7.14) and (7.11). It is clear from the relations, that the functional dependence of $\alpha_{\rm o}$ is needed for both relations. We use Eq. (7.2) to compute the $\alpha_{\rm o}$ function for both surveys using their respective magnification and evolution biases shown in Sec. 3.3. We show the resultant functions in Fig. 7.2. We also show the $\alpha_{\rm c}$ which results from the Kaiser classical analysis (see Sec. 2.3.6 for more details) given by

$$\alpha_{\rm c}(x) = \frac{{\rm d}\,\ln x^2\,\bar{n}_{\rm g,s}(x)}{{\rm d}\ln x} + \frac{Hx}{c(1+z)}\left[1 - \frac{{\rm d}\ln H}{{\rm d}\ln(1+z)}\right]\,. \tag{7.17}$$

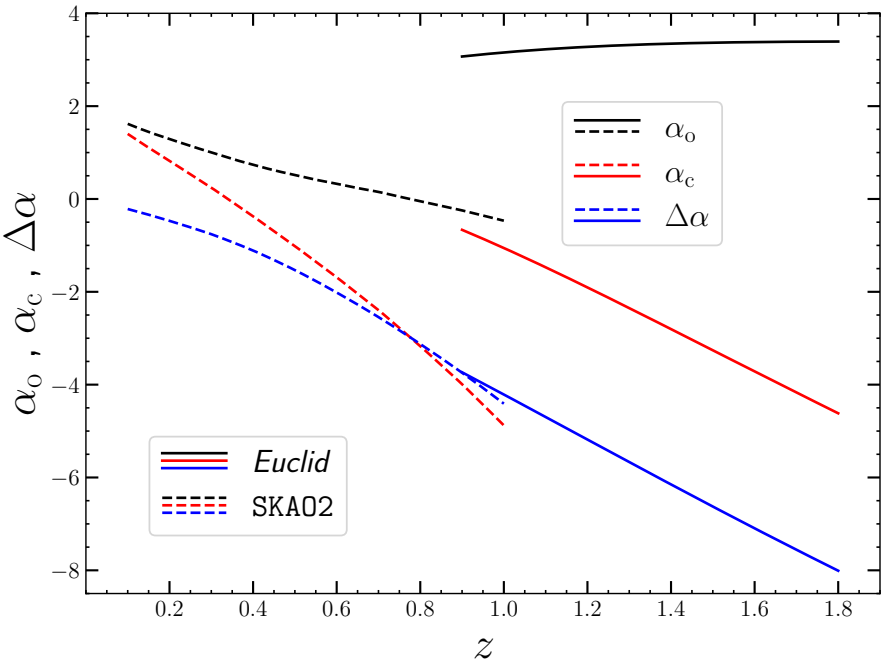


Figure 7.2: α_o, α_c and $\Delta\alpha$ are shown in black, red and blue respectively for the *Euclid* (solid) and SKA02 (dashed) selection functions.

As well as the difference between the two functions, which is given by

$$\Delta\alpha \equiv \alpha_c - \alpha_o = \frac{xH}{c(1+z)}\left(-2\mathcal{Q} - 2\right) \ , \tag{7.18}$$

which are pivotal for our subsequent analysis.

We can predict the behaviour of the FOTO signal through the α_o function. In Fig. 7.2, we plot that function as a solid (dashed) line for the *Euclid* (SKA02) survey. We can directly see that the *Euclid* survey has a relatively constant value of $\alpha_o \approx 3$ for its redshift range $0.9 < z < 1.8$. SKA02 on the other hand has a higher value for low redshifts that monotonically decreases over its redshift range $0.1 < z < 1.0$.

We show an example of the FOTO (Eq. 7.14) signal in Fig.7.3. The black solid line is computed using the *Euclid* α_o function for a deep redshift bin while the red solid line is computed using the SKA02 function in a shallower redshift bin. For both calculations, we set $v_o = 362\,\mathrm{km\ s^{-1}}$ which corresponds to the Planck measurement (Planck Collaboration: Aghanim et al., 2020b). We make three observations from the figure:

- The observer velocity imprints a strong oscillatory signal on the power spectrum monopole.
- That signal is dominant at very large scales (small k) but decays for intermediate scales.
- The frequency of the oscillation is determined by the mean redshift considered.

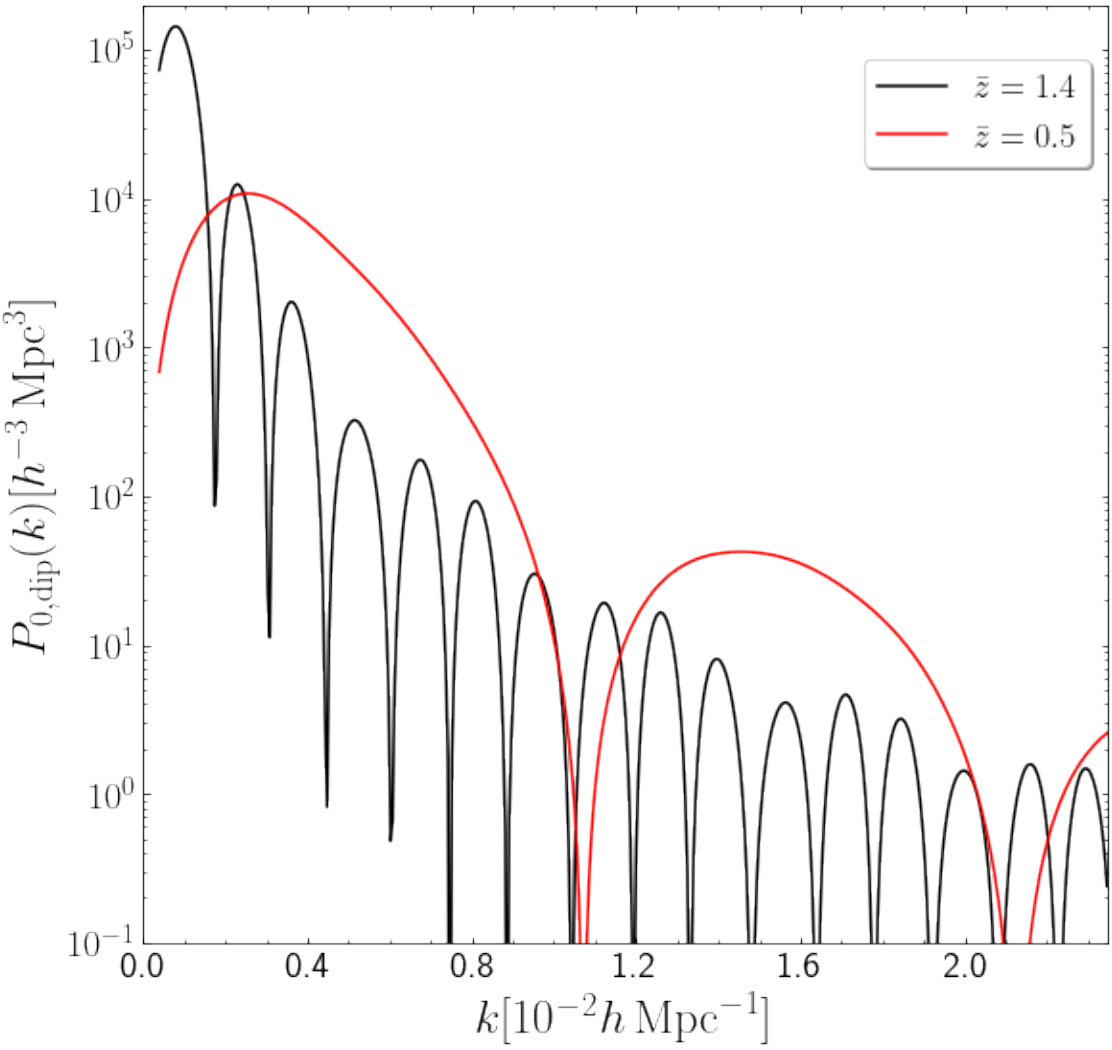


Figure 7.3: An example of the FOTO effect. We plot the results of Eq. (7.14) for two redshift bins; a shallow one in red and a deep redshift bin in black. The frequency dependence of the FOTO signal on the depth of the redshift bin is shown.

The conclusions drawn above relate to the frequency of the signal. Regarding its amplitude, Eq. (7.14) clearly shows that the FOTO signal is directly proportional to both the amplitude of the observer's velocity and the α_o function.

7.2 Comparison with Mocks

7.2.1 Ensemble of Observers

We first discuss the FOTO signal using the spectra inferred from the $\mathcal{O}^R$ mocks. Each observer in these mocks is assigned a random velocity vector. To validate the equations derived in the preceding section, we recalculate the power spectra of the $\mathcal{O}^R$ and $\mathcal{G}$ with the finest resolution possible in k. This entails performing the angular averaging for each individual value of k present in the FFT grid. In essence, instead of averaging the measurements in broader bins, we utilize the resolution of the grid itself. These spectra are computed using the FKP for the *Euclid* survey functions and considering both the *Euclid* and Full-sky footprints, as detailed in Sections 3.3 and 3.4.2.

We numerically evaluate the quantity $\Delta P_0 \equiv P_0^{\mathcal{O}^R} - P_0^{\mathcal{G}}$ for each realization of the mocks. We then show the comparison between the numerically derived measurement ΔP_0 and the theoretical model $\langle P_{0,\rm dip}\rangle$. In Fig. 7.4, we show the 68% percentile region of ΔP_0 for the full-sky footprint in salmon shading while its *Euclid* footprint counterpart is shown in black shading. We note that using the fine binning in k allows us to clearly detect the damped oscillatory features of the FOTO signal, which are not discernible in Fig. 7.1 due to the coarser k-binning employed in that figure. In Fig. 7.4, we also plot the theoretical model given in Eq. (7.15) in dashed blue lines. We compute the theoretical model using the *Euclid* α_o function and the velocity dispersion σ_v which is calculated by using the actual velocity dispersion of the observers in the mock catalogues of the $\mathcal{O}^R$ suite.

As seen in Fig. 7.4, the analytical model agrees quite well with the results of the full-sky *Euclid* mocks. As expected due to the monotonic behaviour observed in the α_o function (black solid line) of the *Euclid* survey shown in Fig. 7.2, the amplitude of the oscillations does not vary strongly with the mean redshift. However, for the mocks with the *Euclid* footprint, the FOTO signal is noticeably affected by the window function. This effect manifests as a suppression in the amplitude of the initial peak and a distortion in the shape surrounding the first minimum. Such discrepancies illustrate the direct influence of the window function(see Chap.5), which mixes the power across different k modes.

We then turn to the anisotropic relation shown in Eq. (7.16). We compute the monopole moment of the power spectrum binning in both the magnitude of the wave vector k as well as the cosine of the angle between the wavevector and the peculiar velocity of the observer,i.e, $\hat{\boldsymbol{v}}_o \cdot \hat{\boldsymbol{k}}$. We label this quantity as the 'two-dimensional' monopole moment of the power spectrum and denote it by $\Delta P_{0,2d}$. In

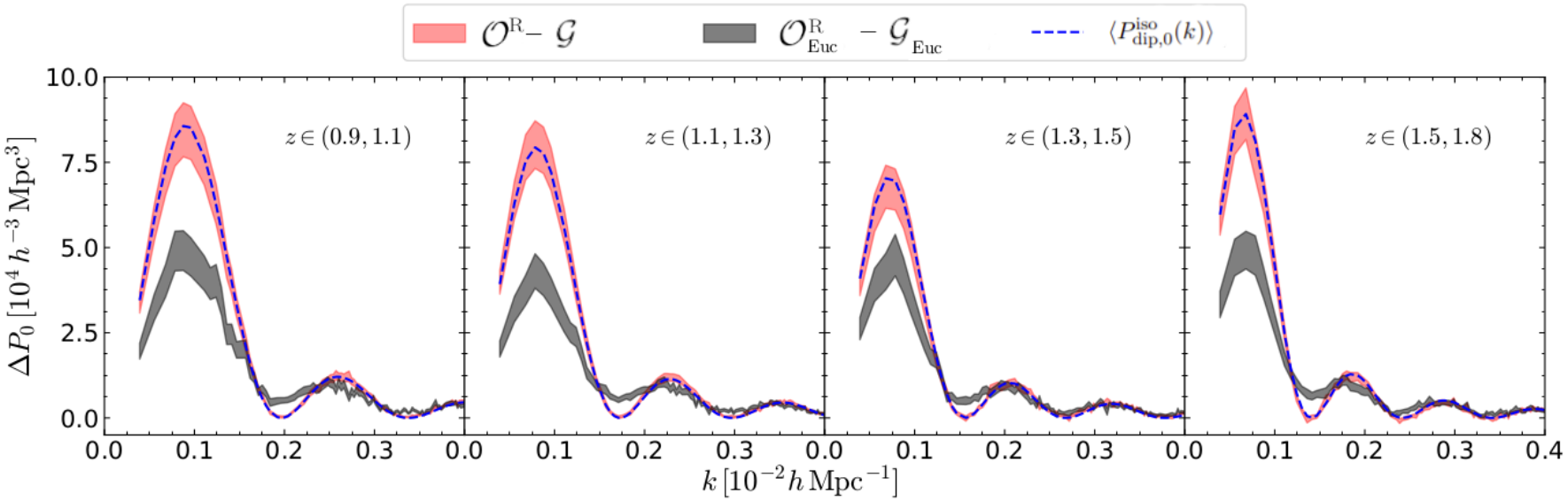


Figure 7.4: The figure illustrates the difference between the monopole moments of the power spectra measured from the $\mathcal{O}^{\rm R}$ and $\mathcal{G}$ mocks across various redshift bins. The shaded regions represent the RMS scatter of the differences, with black colour-shading indicating the *Euclid* footprint and salmon shading for the full-sky footprint. The blue dashed line corresponds to the theoretical prediction from Eq. (7.15), utilizing the $\alpha_{\rm o}$ function defined in Eq. (7.2). In all cases, the *Euclid* survey functions are employed. This figure is adapted from Paper I.

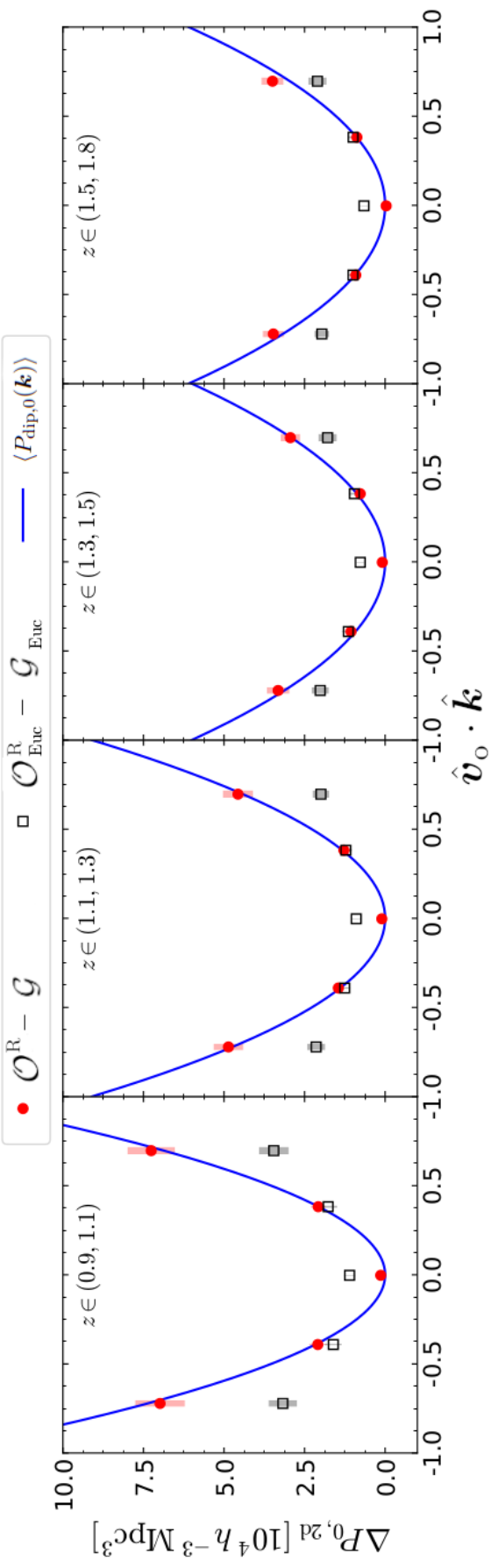


Figure 7.5: The difference between the two-dimensional monopole moments of the power spectra measured from the $\mathcal{O}^{\mathrm{R}}$ and $\mathcal{G}$ mocks as a function of $\hat{\boldsymbol{v}}_{\mathrm{o}} \cdot \hat{\boldsymbol{k}}$ in the bin $k_{\mathrm{FFT}} \leq k < 6\,k_{\mathrm{FFT}}$. The colour scheme is as in Fig. 7.4 but this time the catalogues' measurements are shown as symbols with error bars and the solid lines are computed using Eq. (7.16). This figure is taken from Paper I.

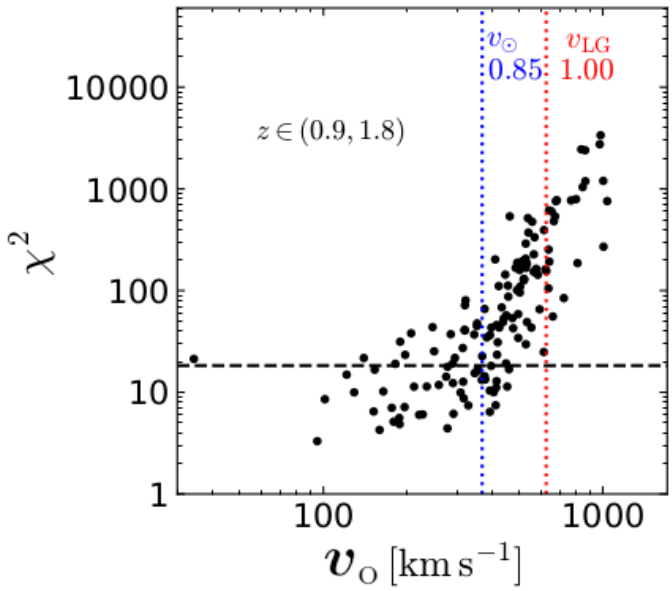


Figure 7.6: The χ^2 goodness-of-fit statistic measured by fitting the mean of the $\mathcal{G}$ spectra for the *Euclid* footprint to the individual measurements of the $\mathcal{O}^{\rm R}$ suite of catalogues (with each realization corresponding to a different $v_{\rm o}$). The horizontal dashed line indicates the 95% confidence limit for 10 degrees of freedom. For the realisations that lie above this line, the difference between $\hat{P}_0^{\mathcal{O}^{\rm R}}$ and the expected monopole of the $\mathcal{G}$ mocks is statistically significant. The blue and red dotted lines represent the peculiar velocity of the solar system and the Local Group, respectively. Next to them, we report the fraction of realisations with $v > v_\odot$ (or $v > v_{\rm LG}$) in which the χ^2 is above the 95% confidence limit. This Figure is taken from Paper I.

order to validate Eq. (7.16), we numerically estimate

$$\Delta P_{0,2{\rm d}}(k, \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{v}}_{\rm obs}) = P^{\mathcal{O}^{\rm R}}_{0,2{\rm d}}(k, \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{v}}_{\rm obs}) - P^{\mathcal{G}}_{0,2{\rm d}}(k, \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{v}}_{\rm obs})\,. \tag{7.19}$$

We show the results in Fig. 7.5. In the figure, we plot $\Delta P_{0,2{\rm d}}$ against $\hat{\boldsymbol{v}}_{\rm o} \cdot \hat{\boldsymbol{k}}$ after averaging over all the Fourier modes with $k \in [1,\, 6]\, k_{\rm FFT}$. The markers with error bars represent the mean and RMS scatter of the spectra of the mocks, while the solid lines are computed using Eq.(7.16). The colour scheme is similar to Fig. 7.4. We again find spectacular agreement between the model and the measurements for the full-sky mock catalogues. As expected, the window function of the *Euclid* footprint mixes the modes, and consequently, the angular dependence is less prominent in the black squares.

At What Speeds Do We Detect The FOTO Signal?

We examine the detectability of the FOTO signal, specifically focusing on determining the required magnitude of the observer velocity for detection for a survey with the *Euclid* survey functions and footprint.

To assess the detectability of the FOTO signal from a single survey marked by

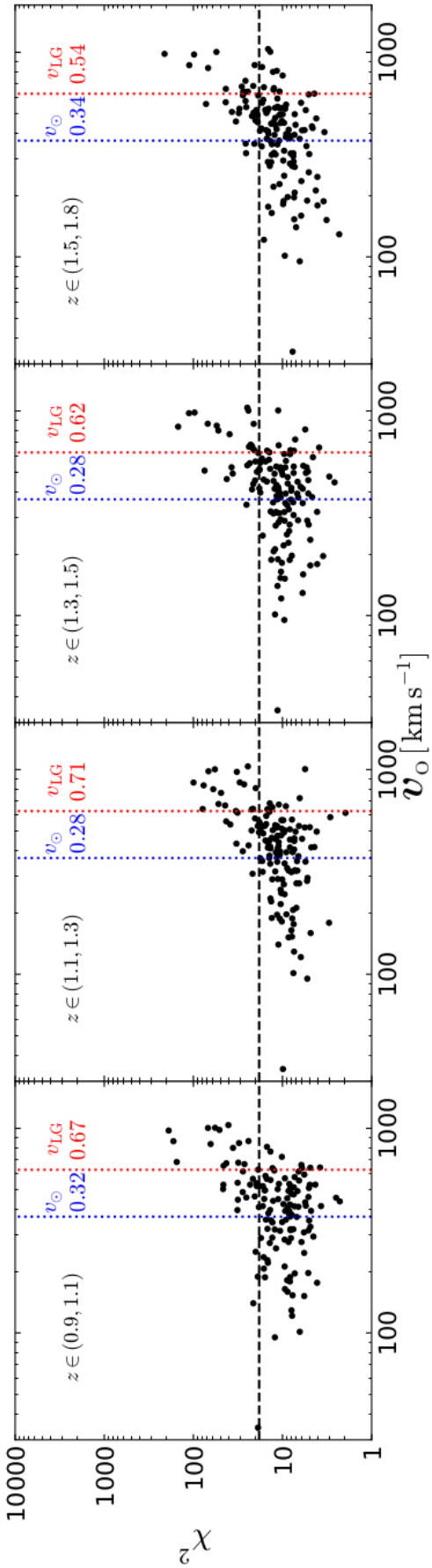


Figure 7.7: The χ^2 goodness-of-fit statistic measured by fitting the mean of the $\mathcal{G}$ spectra for the *Euclid* footprint to the individual measurements of the $\mathcal{O}^{\rm R}$ suite of catalogues. The horizontal dashed line indicates the 95% confidence limit for 10 degrees of freedom. The blue and red dotted lines represent the peculiar velocity of the solar system and the Local Group, respectively. Next to them, we report the fraction of realisations with $v > v_\odot$ (or $v > v_{\rm LG}$) in which the χ^2 is above the 95% confidence limit. This Figure is taken from Paper I.

i, we compare $\hat{P}^{\mathcal{O}^{\rm R}}_{i,0}$ extracted from one of the $\mathcal{O}^{\rm R}$ mock catalogues to $\boldsymbol{\mu}_{\mathcal{G}} = E[\hat{P}^{\mathcal{G}}_0]$, which encompasses all relativistic effects except those originating from $\boldsymbol{v}_{\rm o}$. The objective is to determine if the model accurately describing the monopole moment in the CMB frame is consistent with the data in the observer's frame using the χ^2 goodness-of-fit statistic. Mathematically, we compute (see Sec. 4.4.1 for covariance definition) after binning the spectra in 10 k bins

$$\chi^2_i = \frac{1}{2}\left(\hat{P}^{\mathcal{O}^{\rm R}}_{i,0} - \boldsymbol{\mu}_{\mathcal{G}}\right)^{\rm T} \mathsf{C}^{-1}_{\mathcal{G}} \left(\hat{P}^{\mathcal{O}^{\rm R}}_{i,0} - \boldsymbol{\mu}_{\mathcal{G}}\right), \tag{7.20}$$

for each realization. We then compare the resultant χ^2 values to the 95% percentile of $\mathcal{P}(\chi^2)$ distribution with 10 degrees of freedom.

In Figs. 7.6 and 7.7 , we show the resultant χ^2 values plotted against $v_{\rm o}$ for the catalogues with the *Euclid* footprints. The horizontal dashed line represents the 95% confidence level for a χ^2 distribution with 10 degrees of freedom. When a point exceeds this threshold, it signifies the rejection of the model without observer effects by the data at the chosen confidence level. In essence, the impact of $\boldsymbol{v}_{\rm o}$ should be detectable in these realizations. As expected, the detectability of the signal varies depending on $v_{\rm o}$ and the considered redshift interval. For the broader redshift range depicted in Fig. 7.6, the model in the CMB frame is rejected in all mocks with $v_{\rm o} > 470$ km s^{-1} and in the majority with $v_{\rm o} \simeq 300$ km s^{-1}.

7.2.2 Heliocentric Velocities

In the last section, we treated the velocity of the observer term stochastically, however, the observer velocity is a deterministic term as we as observers have only one peculiar velocity. In this section, we explore a more realistic scenario using the $\mathcal{O}$ mocks, ensuring that the observer velocity of all realizations matches the heliocentric velocity measured by the Planck survey (Planck Collaboration: Aghanim et al., 2020b). Using these mocks enables us to forecast the significance of the signal for realistic surveys and explore potential avenues for parameter inference.

For the *Euclid* survey, we divide the redshift range $z \in (0.9, 1.8)$ into 4 tomographic redshift bins, along with a bin covering the entire survey. Similarly, for the `SKAO2` survey, we utilize 4 tomographic bins covering the range $z \in (0.1, 1.0)$, as well as the widest redshift bin available.

Focusing on the monopole of the spectra, estimated using the FKP estimator with the finest grid possible, we present the results for both surveys in Fig. 7.8. In the Figure, the solid lines represent the mean $P^{\mathcal{O}}_0$ obtained by averaging over the 140 $\mathcal{O}$ catalogues, while the dashed lines represent their $\mathcal{G}$ counterparts. The shaded regions highlight the RMS scatter extracted from the $\mathcal{O}$ set. As expected, in the $\mathcal{O}$ mocks, P_0 exhibits evident oscillations on large scales, which are absent in the $\mathcal{G}$ mocks. The wavenumbers at which the peaks and valleys appear change with the characteristic redshift of the galaxy samples.

In Fig. 7.9, we examine the FOTO signal itself; we show the mean difference

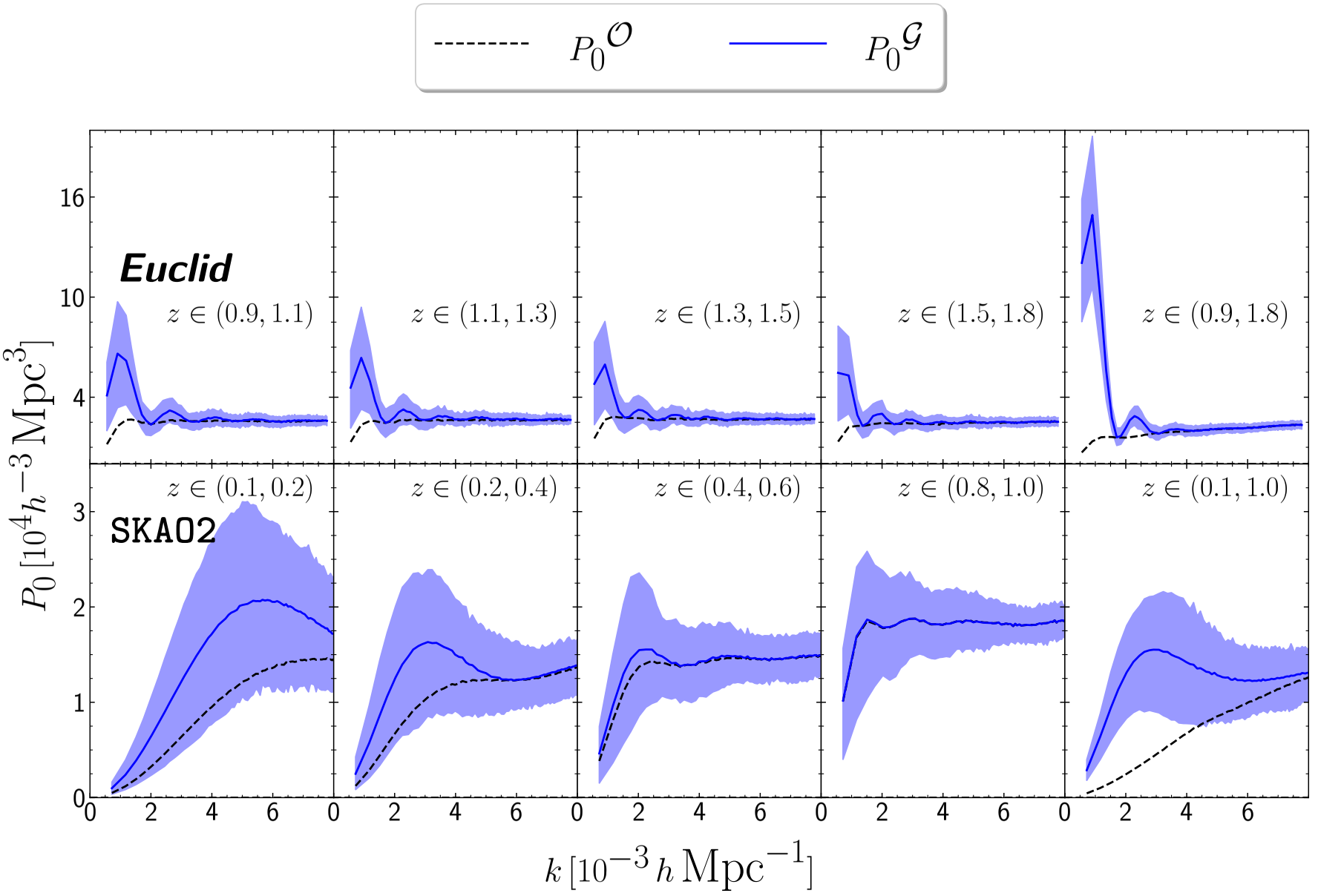


Figure 7.8: We plot the mean of the $\mathcal{O}$ power spectra in blue and the mean of the $\mathcal{G}$ spectra in black. The blue-shaded regions are the RMS scatter of the $\mathcal{O}$ spectra. The top (bottom) panels are for the *Euclid* (SKA02) survey.

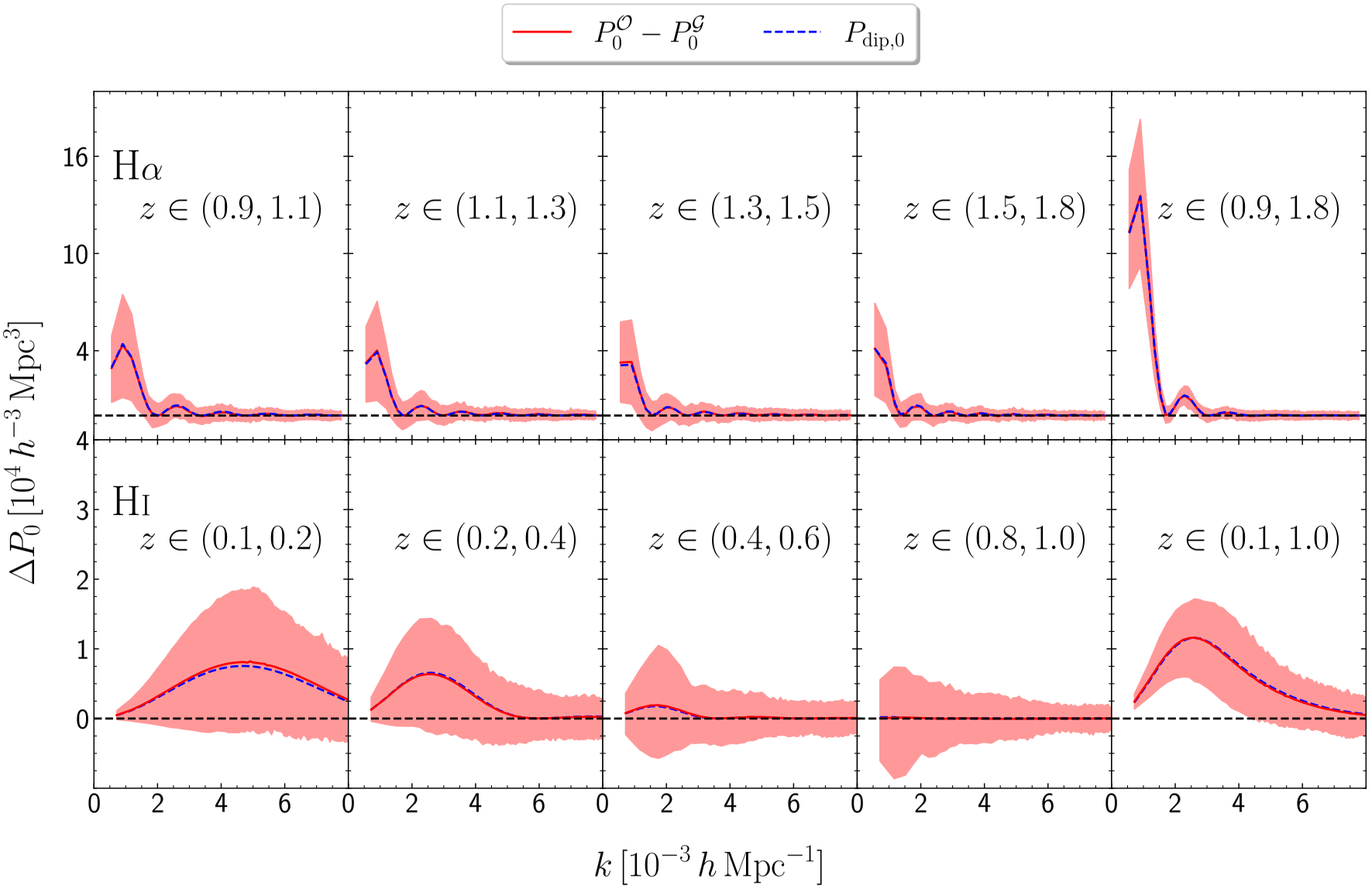


Figure 7.9: We plot the mean of the difference between the $\mathcal{O}$ power spectra and the mean of the $\mathcal{G}$ power spectra monopole in red. The salmon-shaded regions are the RMS scatter for the difference computed for all 140 mocks. The dashed blue line is computed using Eqs. (7.14) for the *Euclid* (Top) and `SKAO2` (bottom) survey functions.

between the monopole spectra obtained from the $\mathcal{O}$ set w.r.t. the $\mathcal{G}$ set. The shaded regions similarly represent the RMS scatter for the difference of the power spectra. The blue dashed line denotes the analytical results computed using Eq. 7.14.

As clear from the figure, the analytical formula agrees spectacularly well with the mock estimates. We note that the FOTO signal differs in both amplitude and frequency between the SKA02 and *Euclid* example surveys.

The amplitude difference can be attributed to the α_o term in Eq. (7.12). We can see the α_o for both surveys in Fig. 7.2. In Fig. 7.2, we notice that the amplitude of the α_o function of the *Euclid* survey is larger than its SKA02 counterpart for all of their respective redshift ranges. We can also note that the *Euclid* amplitude varies slowly with redshift while the α_o function of SKA02 decreases strongly with redshift. This behaviour is reflected in the amplitude of the FOTO effect. For the narrow bins in Fig. 7.9, we see that the FOTO signal of the *Euclid* remains fairly constant while the SKA02 FOTO signal has a significantly smaller amplitude that decreases with redshift. Regarding the difference in frequency, it varies depending on the redshift bin under consideration. In summary, the peak frequency of the signal shifts to smaller wavevectors as the survey depth increases (see also Fig. 7.3).

Is The FOTO Signal Detectable?

In Sec.7.2.1, we employed a χ^2 goodness-of-fit test to assess the detectability of the FOTO signal for a group of observers with random velocities. In this section, we utilize the likelihood ratio test (as discussed in Sec.4.4.1) to quantify the signal-to-noise ratio for detecting the FOTO signal using the *Euclid* survey functions for one singular velocity that is set to match the CMB dipole. Specifically, our objective is to investigate how the signal-to-noise ratio is influenced by the survey mask, considering both its geometry and the fraction of the sky covered, denoted as $f_{\rm sky}$.

To that end, we test if the effect displayed in Fig. 7.9 is measurable for a survey with a realistic survey mask. We use the ($\mathcal{O}$-$\mathcal{G}$) DLR considering three distinct types of angular footprints (see Fig. 3.5):

- a full-sky footprint (see Fig. 3.5d),
- a footprint which excludes all galaxies that are within θ° from the Galactic plane (see Fig. 3.5c). We use three different values for $\theta \in \{10, 15, 20\}$,
- the *Euclid* footprint which avoids the Galactic and Ecliptic planes (see Fig. 3.5b).

We consider the *Euclid* example survey as the FOTO signal of the SKA02 is much lower. We compute the power spectra for these three window function types using the *Euclid* example survey in the redshift bin $z \in (0.9, 1.8)$ for all realizations of $\mathcal{O}$ and $\mathcal{G}$ spectra. We rebin the spectra into 5 bins in the range $k \in (0.39, 5.34) \times 10^{-3}\, h\,\mathrm{Mpc}^{-1}$.

We present the results of the hypothesis test for the three survey mask types considered in Fig. 7.10. The full-sky mocks have the largest S/N while the black points to its left denote the footrprints with the galactic cuts. The red x denotes

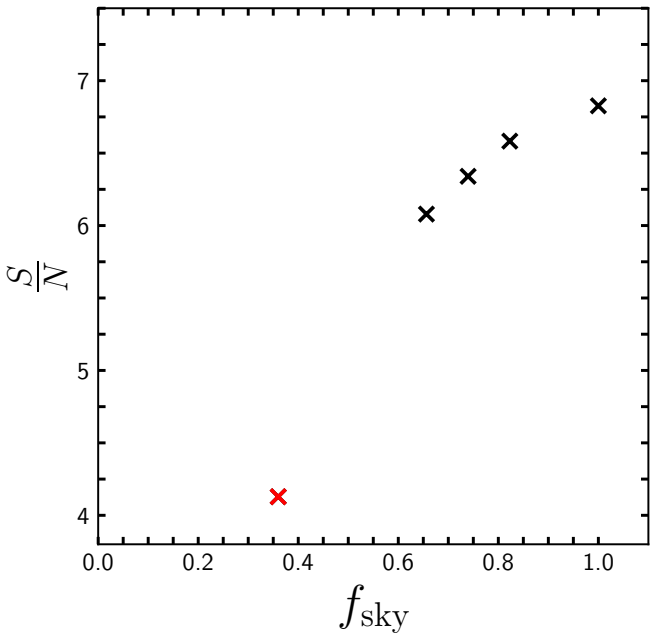


Figure 7.10: We plot the Signal-to-Noise ratio of the FOTO signal measured using the $0.9 < z < 1.8$ redshift bin of the *Euclid* survey using different angular masks against the fraction of the sky covered by each angular mask. The *Euclid* footprint (see Fig. 3.5b) is coloured in red.

the *Euclid* footprint. As anticipated, the S/N increases with the fraction of the sky covered. However, it's noteworthy that the effect remains measurable even for the smallest sky fraction.

An interesting observation from Fig. 7.10 is the pronounced drop in S/N for the *Euclid* mask. This highlights the significant impact of the mask shape on the resulting power spectrum signal.

Alcock-Paczyński Distortions

In measuring the power spectra (and in turn FOTO signal shown in Fig.7.9), we rely on a fiducial cosmological model to convert observed redshifts into spatial Cartesian coordinates. Discrepancies between the actual cosmological parameters, denoted by $\Omega_i^{\rm T}$, and the assumed fiducial cosmological parameters, denoted by $\Omega_i^{\rm F}$, introduce anisotropic distortions in the measured power spectra which are known as Alcock-Paczyński (AP) distortions (Alcock & Paczynski, 1979). In Figs.7.8 & 7.9, we employ the same cosmological parameters for both the mock construction and the power spectrum estimation, thereby ensuring that they are unaffected by the AP effect. However, as measuring the cosmological parameters is a significant aim of redshift surveys, their values are not known *apriori* and thus, the observed FOTO signal is impacted by AP distortions. In this Section, we measure the power spectra using fiducial parameters that differ from the true parameters to explore the impact of the AP effect on the measured FOTO signal.

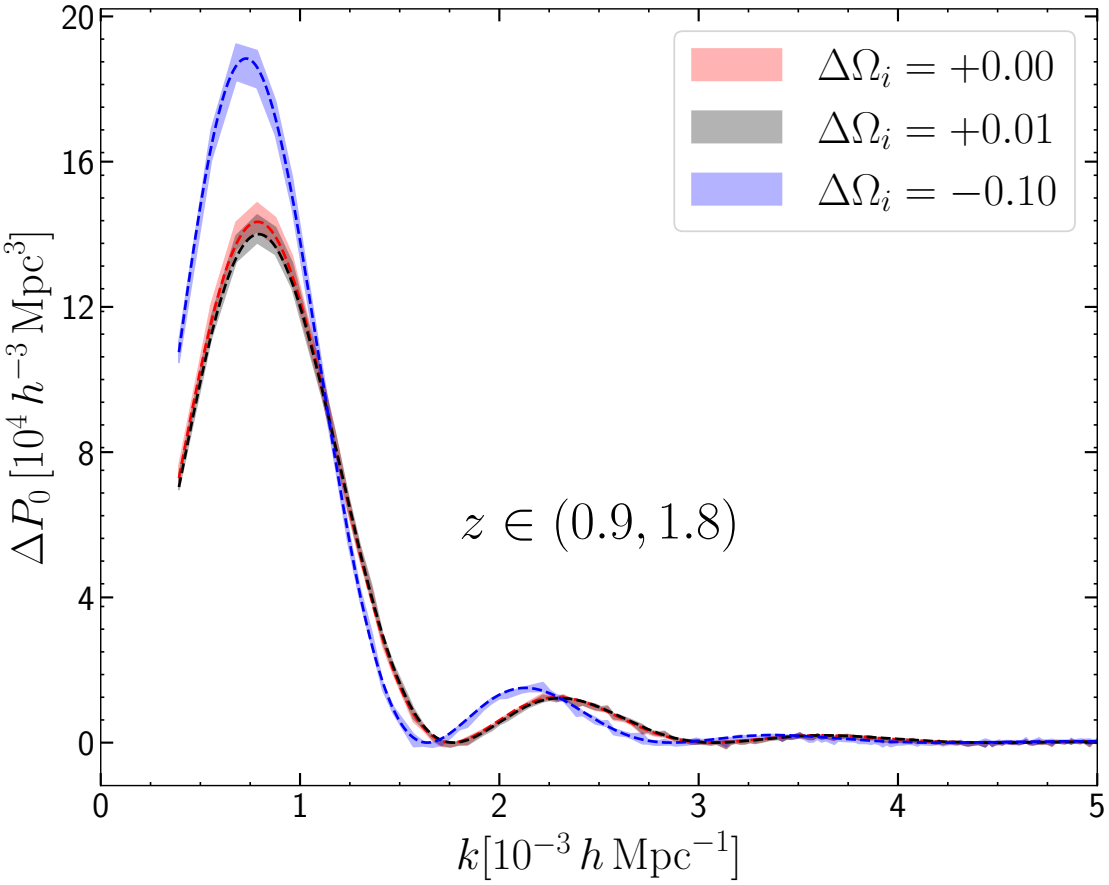


Figure 7.11: We showcase the AP distortion on the FOTO effect. The shaded regions are computed from the mocks. They denote the standard error of the measurement. The dashed lines represent the model predictions computed using Eq. (7.25). The different colours in both the mocks and the model prediction correspond to the different shifts in the cosmological parameters used for the spectra and model estimation.

The AP distortion affects the orthogonal and parallel modes of the power spectrum differently. Denoting the true and assumed cosmology with T and F respectively, we can relate the elements $\{k_\parallel\,, k_\perp\}$ of the true wavevector $\boldsymbol{k}$ to the elements $\{q_\parallel\,, q_\perp\}$ of measured wavevector $\boldsymbol{q}$ through

$$k_\parallel = \frac{H^{\rm T}}{H^{\rm F}}\, q_\parallel \equiv \frac{q_\parallel}{a_\parallel}\,, \qquad k_\perp = \frac{D_{\rm A}^{\rm F}}{D_{\rm A}^{\rm T}}\, q_\perp \equiv \frac{q_\perp}{a_\perp}\,. \tag{7.21}$$

Consequently the distortion to the wavevector amplitude k is given by (noting that $q_\parallel = (\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{x}})\, q$)

$$k(q, \hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{x}}) = \frac{q}{a_\perp}\sqrt{1 - \frac{(a_\parallel^2 - a_\perp^2)}{a_\parallel^2}(\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{x}})^2}\,. \tag{7.22}$$

The impact of the AP effect on the measured power spectrum multipoles $\widehat{P}_\ell(q)$ can then be written as (Ballinger et al., 1996, Beutler et al., 2014)

$$\widehat{P}_\ell(q) = \frac{(2\ell+1)}{a_\perp^2\, a_\parallel}\int P_{\rm g}(\boldsymbol{k})\mathcal{L}_\ell(\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{x}})\frac{{\rm d}\Omega_x}{4\pi}\,, \tag{7.23}$$

where $P_{\rm g}(\boldsymbol{k})$ is the power spectrum computed at the true wavevector $\boldsymbol{k}$.

To find how the AP effect alters the FOTO signal in the power spectrum monopole, we set $\ell = 0$ and replace $P_{\rm obs}$ with the model prediction in Eq. (7.14)

$$\widehat{P}_{\rm dip,0}(q) = \frac{1}{a_\perp^2\, a_\parallel}\int P_{\rm dip,0}[k(q, \hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{x}})]\frac{{\rm d}\Omega_x}{4\pi}\,. \tag{7.24}$$

Assuming that $(a_\parallel^2 - a_\perp^2)/a_\parallel^2 \ll 1$, we can write the measured signal as

$$\widehat{P}_{\rm dip,0}(q) \approx \frac{1}{a_\perp^2\, a_\parallel} P_{\rm dip,0}\left(\frac{q}{a_\perp}\right)\,. \tag{7.25}$$

In order to test how well Eq.(7.25) recovers the estimated power spectrum, we measure the power spectra assuming three different sets of cosmological parameters. The first set is the true cosmological parameters described in Sec. 3.2, that measurement is unaffected by the AP distortions. The other two sets of cosmological parameters are built by shifting $(\Omega_{\rm m}, \Omega_\Lambda) \to (\Omega_{\rm m} + \Delta\Omega, \Omega_\Lambda - \Delta\Omega)$, where we use two values for $\Delta\Omega \in \{0.01, -0.1\}$. The first value is the same order of magnitude as the the errors on $\Omega_{\rm m}$ given by the CMB constraints (Planck Collaboration: Aghanim et al., 2020a).

The results can be seen in Fig. 7.11. The shaded regions are centred around the mean of the estimated spectra and bounded by the error on the mean. The black-shaded region represents the true parameters while the grey and blue-shaded regions represent the small and large shifts respectively. The dashed lines are the model predictions using Eq. (7.25). The colour scheme is the same for the model

predictions. As seen in the figure, the small parameter shift $\Delta\Omega = 0.01$ leads to a slight change in the amplitude while the frequency dependence is not affected. On the other hand, the $\Delta\Omega = 0.1$ shift strongly affects both the amplitude and frequency of the FOTO signal. In both cases, we are able to recover both the amplitude and frequency shift using Eq. (7.25). We note that our model is equally successful for all the redshift bins considered in this work.

7.3 Higher Multipoles

In the previous sections, we mainly concentrated on the FOTO signal in the monopole of the power spectra. In this section, we derive the FOTO effect for higher multipoles. To that end, we compute the autocorrelation of the observer dipole term for any multipole using Eq. (5.3)

$$P_{\mathrm{dip},\ell}(k) = \frac{2\ell+1}{A}\int\int\int B_1 B_2 (\boldsymbol{v}_\mathrm{o}\cdot\hat{\boldsymbol{x}}_1)(\boldsymbol{v}_\mathrm{o}\cdot\hat{\boldsymbol{x}}_2)\mathrm{e}^{i\boldsymbol{k}\cdot(\boldsymbol{x}_1-\boldsymbol{x}_2)}\mathcal{L}_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}}_2)\mathrm{d}^3x_1\mathrm{d}^3x_2\frac{\mathrm{d}\Omega_k}{4\pi}\,.$$

We split the integrals into two parts

$$P_{\mathrm{dip},\ell}(k) = \frac{2\ell+1}{A}\int\left[\int B_1(\boldsymbol{v}_\mathrm{o}\cdot\hat{\boldsymbol{x}}_1)\mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{x}_1}\mathrm{d}^3x_1\right] \left[\int B_2(\boldsymbol{v}_\mathrm{o}\cdot\hat{\boldsymbol{x}}_2)\mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{x}_2}\mathcal{L}_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}}_2)\mathrm{d}^3x_2\right]\frac{\mathrm{d}\Omega_k}{4\pi}\,. \tag{7.26}$$

In evaluating the previous integrals, we use the following identities

$$\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} = \sum_\ell (2\ell+1)\,\mathrm{i}^\ell\, j_\ell(kx)\mathcal{L}_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}})\,, \tag{7.27}$$

$$\int \mathcal{L}_\ell(\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{x}})\mathcal{L}_{\ell_1}(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}})\,\mathrm{d}\Omega_x = \frac{4\pi}{2\ell_1+1}\mathcal{L}_{\ell_1}(\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{k}})\;\delta^\mathrm{K}_{\ell\ell_1}\,, \tag{7.28}$$

$$\mathcal{L}_\ell(\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{x}})\mathcal{L}_{\ell_1}(\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{x}}) = \sum_{n=\ell-\ell_1}^{\ell+\ell_1}\begin{pmatrix}\ell & \ell_1 & n\\ 0 & 0 & 0\end{pmatrix}^2(2n+1)\mathcal{L}_n(\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{x}})\,, \tag{7.29}$$

$$I_n = \int x^2\, B_1\, j_n(kx)\,\mathrm{d}x\,. \tag{7.30}$$

For the first square bracket in Eq. (7.26), we expand the exponential and split the integral into radial and angular parts

$$\begin{aligned}\int B_1(\boldsymbol{v}_\mathrm{o}\cdot\hat{\boldsymbol{x}})\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\,\mathrm{d}^3x &= \sum_{\ell_1}(2\ell_1+1)\,\mathrm{i}^{\ell_1}\int A(\boldsymbol{v}_\mathrm{o}\cdot\hat{\boldsymbol{x}})\,j_{\ell_1}(kx)\,\mathcal{L}_{\ell_1}(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}})\,\mathrm{d}\Omega_x\,\mathrm{d}x\\ &= v_\mathrm{o}\sum_{\ell_1}(2\ell_1+1)\mathrm{i}^{\ell_1}I_{\ell_1}\frac{4\pi}{2\ell_1+1}\mathcal{L}_{\ell_1}(\hat{\boldsymbol{v}}_o\cdot\hat{\boldsymbol{k}})\delta^\mathrm{D}_{\ell_1 1}\\ &= 4\pi\mathrm{i}\,v_\mathrm{o}\,I_1\mathcal{L}_1(\hat{\boldsymbol{v}}_o\cdot\hat{\boldsymbol{k}})\,.\end{aligned} \tag{7.31}$$

For the second square bracket in Eq. (7.26),

$$
\begin{aligned}
&\int B_2(\boldsymbol{v}_\mathrm{o}\cdot\hat{\boldsymbol{x}})\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\mathcal{L}_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}})\mathrm{d}^3x = \\
&\quad\sum_{\ell_1}(2\ell_1+1)(-\mathrm{i})^{\ell_1}I_{\ell_1}\int(\hat{\boldsymbol{v}}_o\cdot\hat{\boldsymbol{x}})\mathcal{L}_{\ell_1}(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}})\mathcal{L}_l(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}})\mathrm{d}\Omega_x \\
&=v_\mathrm{o}\sum_{\ell_1}(2\ell_1+1)(-\mathrm{i})^{\ell_1}I_{\ell_1}\int(\hat{\boldsymbol{v}}_o\cdot\hat{\boldsymbol{x}})\sum_{n=\ell-\ell_1}^{\ell+\ell_1}\begin{pmatrix}\ell & \ell_1 & n\\ 0 & 0 & 0\end{pmatrix}^2(2n+1)\mathcal{L}_n(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}})\mathrm{d}\Omega_x \\
&=v_\mathrm{o}\sum_{\ell_1}\sum_{n=\ell-\ell_1}^{\ell+\ell_1}\begin{pmatrix}\ell & \ell_1 & n\\ 0 & 0 & 0\end{pmatrix}^2(2\ell_1+1)(2n+1)(-\mathrm{i})^{\ell_1}I_{\ell_1}\int\mathcal{L}_1(\hat{\boldsymbol{v}}_o\cdot\hat{\boldsymbol{x}})\mathcal{L}_n(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{x}})\mathrm{d}\Omega_x \\
&=v_\mathrm{o}\sum_{\ell_1}\sum_{n=\ell-\ell_1}^{\ell+\ell_1}\begin{pmatrix}\ell & \ell_1 & n\\ 0 & 0 & 0\end{pmatrix}^2(2\ell_1+1)(2n+1)(-\mathrm{i})^{\ell_1}I_{\ell_1}\left[\frac{4\pi}{2n+1}\mathcal{L}_n(\hat{\boldsymbol{v}}_o\cdot\hat{\boldsymbol{k}})\delta^\mathrm{D}_{n1}\right] \\
&=4\pi\, v_\mathrm{o}\sum_{\ell_1}\begin{pmatrix}\ell & \ell_1 & 1\\ 0 & 0 & 0\end{pmatrix}^2(2\ell_1+1)(-\mathrm{i})^{\ell_1}I_{\ell_1}\mathcal{L}_1(\hat{\boldsymbol{v}}_o\cdot\hat{\boldsymbol{k}})\,. \qquad (7.32)
\end{aligned}
$$

Plugging in both results in Eq. (7.26)

$$
\begin{aligned}
P_{\mathrm{dip},\ell}(k) &= \frac{2\ell+1}{A}\,v_\mathrm{o}^2\int\left[4\pi\mathrm{i}I_1\mathcal{L}_1(\hat{\boldsymbol{v}}_o\cdot\hat{\boldsymbol{k}})\right] \\
&\quad\times\left[4\pi\sum_{\ell_1}\begin{pmatrix}\ell & \ell_1 & 1\\ 0 & 0 & 0\end{pmatrix}^2(2\ell_1+1)(-\mathrm{i})^{\ell_1}I_{\ell_1}\mathcal{L}_1(\hat{\boldsymbol{v}}_o\cdot\hat{\boldsymbol{k}})\right]\frac{\mathrm{d}\Omega_k}{4\pi} \\
&= \frac{16\pi^2}{3}\frac{2\ell+1}{A}\,v_\mathrm{o}^2\,I_1\sum_{\ell_1}\mathrm{i}\,(-\mathrm{i})^{\ell_1}\begin{pmatrix}\ell & \ell_1 & 1\\ 0 & 0 & 0\end{pmatrix}^2(2\ell_1+1\,)I_{\ell_1}\,. \qquad (7.33)
\end{aligned}
$$

The contribution to each multipole will be given by computing the Wigner 3j symbol. We explicitly provide the FOTO contribution to the first few power spectrum multipoles

$$P_{\mathrm{dip},0}(k) = \frac{16\pi^2}{3A}v_\mathrm{o}^2\,I_1^2(k)\,, \qquad (7.34)$$

$$P_{\mathrm{dip},1}(k) = \frac{16\pi^2\mathrm{i}}{3A}v_\mathrm{o}^2\,\left[I_0(k)-2I_2(k)\right]\,, \qquad (7.35)$$

$$P_{\mathrm{dip},2}(k) = -\frac{16\pi^2}{5A}v_\mathrm{o}^2\,\left[2I_1(k)+3I_3(k)\right]\,, \qquad (7.36)$$

$$P_{\mathrm{dip},3}(k) = \frac{16\pi^2\mathrm{i}}{7A}v_\mathrm{o}^2\,\left[-3I_2(k)+4I_4(k)\right]\,, \qquad (7.37)$$

$$P_{\mathrm{dip},4}(k) = -\frac{16\pi^2}{9A}v_\mathrm{o}^2\,\left[4I_3(k)+5I_5(k)\right]\,. \qquad (7.38)$$

Examining the previous result, we observe that the monopole exhibits a positive-

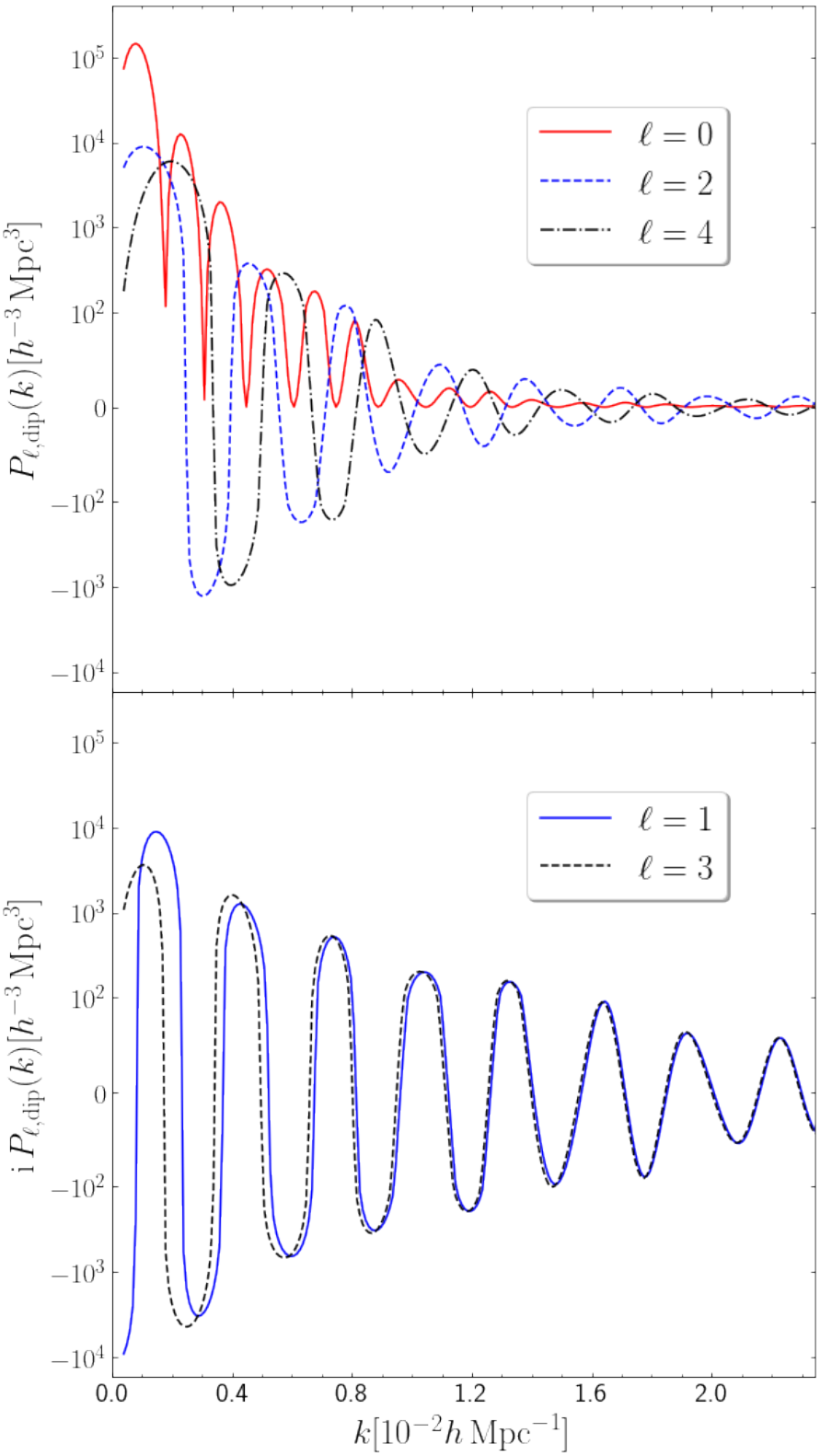


Figure 7.12: We showcase the FOTO signal for the different multipoles. The even multipoles are shown in the top panel while the odd multipoles are shown in the bottom panel. We note that the odd multipoles are purely imaginary.

definite contribution, whereas the sign of the FOTO signal for the higher multipoles varies depending on the cumulative effects of the I_n functions. Additionally, the odd multipoles are purely imaginary, while the even multipoles are purely real. This phenomenon can be attributed to the inherent asymmetry of the estimator employed. In Eq. (7.3), the first spatial integral demonstrates symmetry under reflection in the wavevector ($\boldsymbol{k} \to -\boldsymbol{k}$), while the second spatial integral only exhibits symmetry for even multipoles and antisymmetry for odd ones. Consequently, the real part of the estimator averages to zero for odd multipoles, and the imaginary part averages to zero for even multipoles. This indicates the limitation of the YB estimator in accurately capturing clustering at extremely large scales.

The first 5 multipoles of the FOTO signal are depicted in Fig. 7.12. As expected, the monopole is the only positive-definite signal. Moreover, we observe that the power induced in the odd multipoles significantly exceeds that of the quadrupole and hexadecapole. We also note that the signal for the odd multipoles does not decline as quickly as the even multipoles.

7.4 Can we cancel the FOTO Signal?

In most models of galaxy clustering, the adopted reference frame typically assumes that we are in motion with the cosmological fluid, i.e., the $\mathcal{G}$ reference frame. Therefore, for cosmological purposes, one might seek to invert Eq. (2.62) and deduce individual galaxy redshifts in the $\mathcal{G}$ frame by employing (e.g. Hamilton, 1998, Glanville et al., 2021, Peterson et al., 2022)

$$1 + z_{\rm cos} = \frac{1 + z_{\rm obs}}{1 + z_{\rm pec,o}} \simeq \frac{1 + z_{\rm obs}}{1 - v_{\parallel,\rm o}/c} \simeq (1 + z_{\rm obs}) \left(1 + \frac{v_{\parallel,\rm o}}{c}\right) . \tag{7.39}$$

That transformation leads to the inverse of the observer velocity radial shift shown in Eq. (2.67), i.e.,

$$\delta x = \frac{v_{\parallel,\rm o}}{a\, H(z_{\rm obs})} . \tag{7.40}$$

At linear order, the product $a\,H$ is unchanged by the redshift transformation, in other words, it is equivalent to compute $a\,H$ at $z_{\rm obs}$ or $z_{\rm cos}$.

However, Eq. (7.40) does not cancel the FOTO signal as it only accounts for the radial shift but fails to incorporate the transversal shift due to relativistic aberration or the change in the observed flux (see Sec. 2.3.6).

We can trace the impact of the redshift correction shown in Eq. (7.39) on the density contrast using a similar derivation to the one presented in Sec. (2.2.1). Denoting the galaxy overdensity after the correction by $\delta_{\rm corr}$, we find

$$\delta_{\rm corr}(\boldsymbol{x}) = \delta_{\rm obs}(\boldsymbol{x}) - \alpha_{\rm c}\, \frac{\boldsymbol{v}_{\rm o} \cdot \hat{\boldsymbol{x}}}{a\, H\, x} , \tag{7.41}$$

where (see Sec. 2.3.6)

$$\alpha_{\rm c}(x) = 2\,(1-Q) - \frac{H\,x}{c\,(1+z)}\left[2Q + \mathcal{E} + 1 - \frac{{\rm d}\ln H}{{\rm d}\ln(1+z)}\right] \,. \tag{7.42}$$

We can then use Eq. (5.1) to relate the corrected density contrast to the one defined in the comoving frame

$$\delta_{\rm corr} = \delta_{\rm com} + (\alpha_{\rm o} - \alpha_{\rm c})\,\frac{\boldsymbol{v}_{\rm o}\cdot\hat{\boldsymbol{x}}}{a\,H\,x} \,. \tag{7.43}$$

It is clear from Eq. (7.43) that the corrected overdensity still depends on the observer velocity and does not coincide with $\delta_{\rm com}$ ($\mathcal{G}$ frame).

We can find the impact of the redshift correction on the power spectrum by simply replacing $\alpha_{\rm o}$ by $\Delta\alpha \equiv \alpha_c - \alpha_{\rm o}$ in Eq. (7.12) as the remainder of the derivation is unaffected. The result is given by [2]

$$P^{\rm corr}_{\rm dip,0} \equiv P^{\rm corr}_0(k) - P^{\rm com}_0(k) = D_{\Delta\alpha\Delta\alpha}(k) = \frac{16\pi^2}{3}\,\frac{v_{\rm o}^2\,I^2_{\Delta\alpha}(k)}{\int \bar{n}^2(x)\,{\rm d}^3x} \,, \tag{7.44}$$

where

$$I_{\Delta\alpha}(k) = \int_{x(z_i)}^{x(z_f)} \frac{x\,\bar{n}\,\Delta\alpha}{a\,H}\,j_1(kx)\,{\rm d}x \,. \tag{7.45}$$

From the previous equation, it is clear that the correction succeeds in cancelling the FOTO signal only when $\Delta\alpha$ vanishes. That function is given by

$$\Delta\alpha = \frac{xH}{c(1+z)}\,(-2\mathcal{Q} - 2) \,. \tag{7.46}$$

To predict where the signal vanishes, we show $\Delta\alpha$ in Fig. 7.2 in solid blue for the *Euclid* survey and dashed blue for the SKAO2 survey. From the figure, we can deduce that the redshift correction only cancels the signal for the low redshift bins ($z \lesssim 0.4$) of the SKAO2 survey. This result is consistent with the $\Delta\alpha \propto x$ scaling in Eq. (7.46). It is also clear that the redshift transformation boosts the signal in the remainder of the redshift range as $(\Delta\alpha)^2 > \alpha_{\rm o}^2$ for the *Euclid* survey in its entire redshift range and for the SKAO2 survey at $z \gtrsim 0.4$.

In order to validate Eq. (7.44) and the impact of the redshift correction, we numerically correct the redshift of each galaxy in a given $\mathcal{O}$ mock realization according to Eq. (7.39) and then compute the power spectrum of corrected mocks. We show the difference of the spectra after the redshift correction, $P_0^{\mathcal{O},\rm corr}$, w.r.t the $\mathcal{G}$ power spectra in Fig. 7.13. In particular, we plot the mean and 68% percentile range of

[2] The superscript in the power spectrum symbol denotes the subscript of the overdensity it is computed from, e.g., $P_0^{\rm corr}$ is the monopole power spectrum of the $\delta_{\rm corr}$.

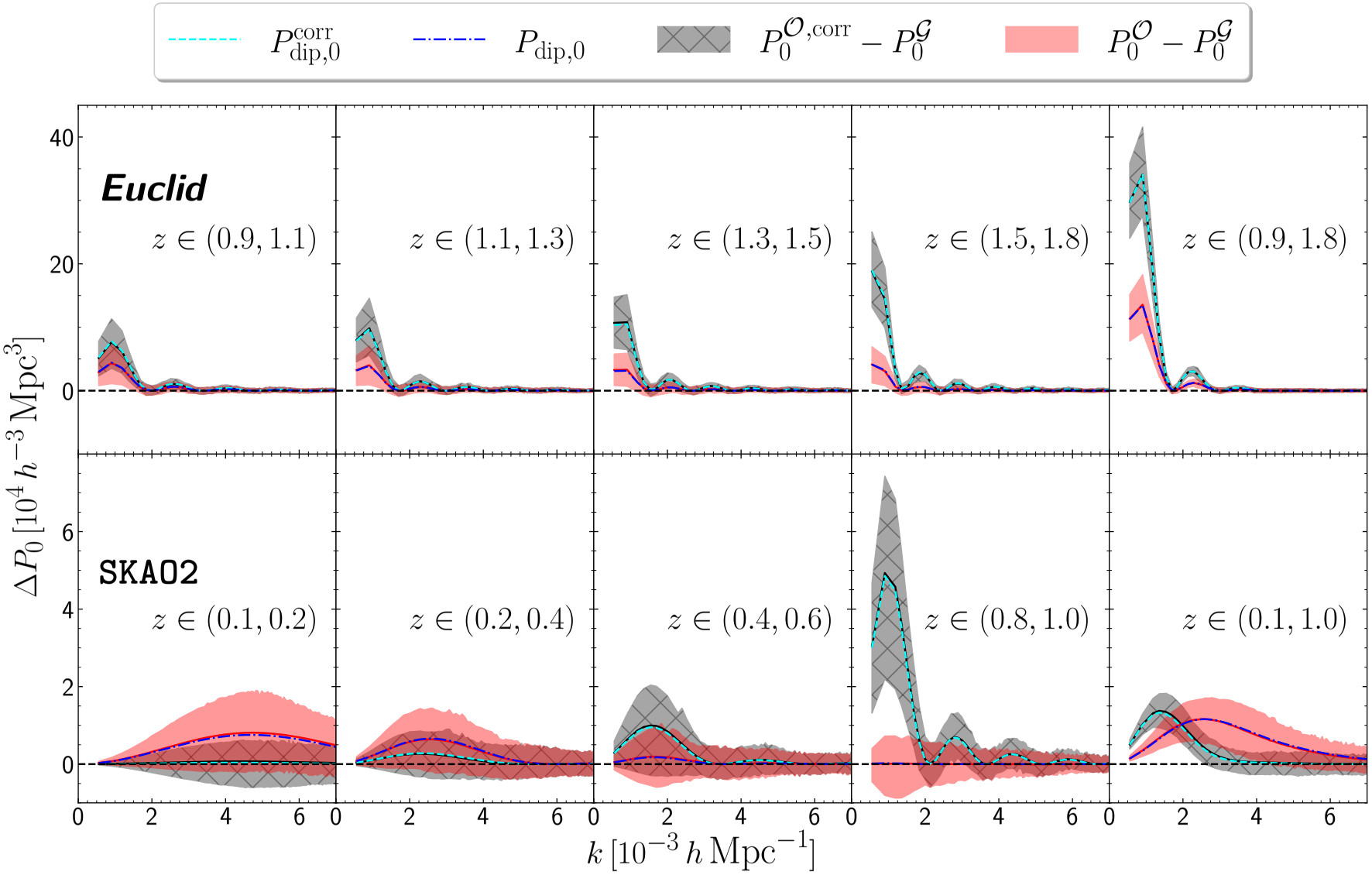


Figure 7.13: The black solid line represents the mean monopole of the spectra after applying the redshift correction, while the grey-shaded regions indicate the RMS scatter. Similarly, the red solid lines and pink shaded regions denote the means and RMS scatter of the original signals imported from Fig. 7.9. The cyan dashed lines correspond to the theoretical model incorporating the redshift correction, as described by Eq. (7.44). Finally, the blue dash-dotted lines represent models of the original signal computed using Eq. (7.14). This Figure is based on a similar one from Paper II

the corrected signal in black and grey shading respectively. The mean and 68% percentile range of the original uncorrected signal are also shown as the red solid line and salmon shading respectively. Finally, we plot the analytical prediction after the redshift correction (see Eq. 7.44) in dashed cyan while the theoretical prediction of the original signal (see Eq. 7.14) is plotted in dashed blue. As expected, the manual correction only cancels the signal at low redshifts for the SKAO2 example survey. Moreover, the signal is boosted for the high redshift bins of the SKAO2 survey ($z \gtrsim 0.4$) and for all bins of the *Euclid* survey. Secondly, our model predictions for redshift corrected spectra are able to recover the results from the mocks extremely well as seen by comparing the cyan and black lines. Moreover, we observe that the peak frequency of the corrected FOTO signal in the largest redshift bin of the SKAO2 survey is lower. We attribute this frequency shift to the presence of $\Delta\alpha$ within the integral of Eq. (7.44). Since $\Delta\alpha$ approaches zero for $z \lesssim 0.4$, the integral is more influenced by the higher redshift range of the redshift bin. Consequently, the resulting peak frequency shifts towards smaller k values. In conclusion, it is clear that in high redshift spectroscopic surveys ($z > 0.4$), the redshift transformation correction of each redshift by the observer velocity can introduce spurious signals at large scales.

7.4.1 Can we boost the FOTO Signal?

The previous section provides a method to amplify the FOTO signal. The approach is straightforward: we begin by selecting an arbitrary peculiar velocity, denoted as $\boldsymbol{v}_{\rm art}$. We then shift the redshifts of each galaxy in the observed galaxy catalogue by the following relation

$$1 + z_{\rm art} = \frac{1 + z_{\rm obs}}{1 + z_{\rm sh}}\,, \tag{7.47}$$

where

$$z_{\rm sh} = \frac{\sqrt{1 - v_{\rm art}^2/c^2}}{1 - \boldsymbol{v}_{\rm art}/c \cdot \hat{\boldsymbol{x}}}\,. \tag{7.48}$$

We then measure the power spectrum monopole of the artificially shifted catalogue of galaxies, $P_0^{obs,{\rm art}}$, with modified redshifts. The overdensity $\delta_{\rm art}$ from the aforementioned redshift-only boost at linear order altered by the shift is then given by

$$\delta_{\rm art} = \delta_{\rm com} + \frac{\alpha_{\rm o}}{a\,H\,x}\,(\boldsymbol{v}_{\rm o} \cdot \hat{\boldsymbol{x}}) + \frac{\alpha_{\rm c}}{a\,H\,x}(\boldsymbol{v}_{\rm art} \cdot \hat{\boldsymbol{x}})\,. \tag{7.49}$$

Writing the previous equation in terms of $\Delta\alpha$

$$\delta_{\rm art} - \delta_{\rm com} = \frac{\alpha_{\rm o}}{a\,H\,x}[(\boldsymbol{v}_{\rm art} + \boldsymbol{v}_{\rm o}) \cdot \hat{\boldsymbol{x}}] + \frac{\Delta\alpha}{a\,H\,x}(\boldsymbol{v}_{\rm art} \cdot \hat{\boldsymbol{x}})\,. \tag{7.50}$$

We can compute the autocorrelation of the resulting dipole term following the procedure presented in Sec. 7.3. The final result for the power spectrum monopole

	Velocity boosts				
$v_{\rm art}(v_{\rm o})$	3	1	2	2	3
$\cos^{-1}(\boldsymbol{v}_{\rm art}\cdot\boldsymbol{v}_{\rm o})$	$\pi/6$	$\pi/3$	π	0	$\pi/2$

Table 7.1: The velocity vectors used for the ARBs for both the velocity and cosmological parameter measurements. $v_{\rm art}$ is the norm of the velocity in units of the true observer velocity and $\cos^{-1}(\boldsymbol{v}_{\rm art}\cdot\boldsymbol{v}_{\rm o})$ is the angle between the artificial velocity vector and the true velocity.

is

$$\begin{aligned} P^{\rm art}_{\rm dip,0} &= P^{\rm art}_0(k) - P^{\rm com}_0(k) \\ &= \frac{16\pi^2}{3\int \bar{n}^2(x)\,{\rm d}^3x}\left[D_{\alpha_{\rm o}\alpha_{\rm o}}(k) + 2D_{\alpha_{\rm o}\Delta\alpha}(k) + D_{\Delta\alpha\Delta\alpha}(k)\right] , \end{aligned} \tag{7.51}$$

where

$$D_{\alpha_{\rm o}\alpha_{\rm o}}(k) = |\boldsymbol{v}_{\rm o} + \boldsymbol{v}_{\rm art}|^2\ I^2_{\alpha_{\rm o}}(k)\ , \tag{7.52}$$

$$D_{\Delta\alpha\Delta\alpha}(k) = v^2_{\rm art}\, I^2_{\Delta\alpha}(k)\ , \tag{7.53}$$

$$D_{\alpha_{\rm o}\Delta\alpha}(k) = [(\boldsymbol{v}_{\rm o} + \boldsymbol{v}_{art})\cdot\boldsymbol{v}_{\rm art}]\ I_{\alpha_{\rm o}}(k) I_{\Delta\alpha}(k)\ . \tag{7.54}$$

It is clear from Eq. (7.51) that a careful selection of the magnitude and orientation of $\boldsymbol{v}_{art}$ can substantially enhance the FOTO signal. That artificially amplified FOTO signal offers two primary advantages: firstly, the aforementioned boost in the signal; secondly, and more significantly, the signal's behaviour is dependant upon both the magnitude and the **direction** of $\boldsymbol{v}_{\rm o}$ through $D_{\alpha_{\rm o}\alpha_{\rm o}}$ and $D_{\alpha_{\rm o}\Delta\alpha}$. In essence, the artificially enhanced signal relies on the complete observer velocity vector, cosmological parameters and survey functions that make up the $\alpha_{\rm o}$ and $\Delta\alpha$ functions.

In the subsequent section, we illustrate how the application of the artificial shift enables us to measure the full velocity vector of the observer. We also show the potential of constraining cosmological parameters based on the signal. We denote the procedure of applying the artificial radial redshift boost to each galaxy in the catalogue by the Artificial Redshift Boost (ARB).

7.5 Measurements from the FOTO Signal.

As shown in the previous section, applying the ARB procedure can significantly amplify the FOTO signal. In this section, we utilise this technique to extract both the true velocity of the observer and essential cosmological parameters.

We present a case study of this methodology applied to both example surveys; the *Euclid* and `SKA02` using the full-sky footprint. We demonstrate the efficacy of this approach in two distinct scenarios

- Estimating the velocity of the observer assuming prior knowledge of the cosmological parameters (Hereafter, $\boldsymbol{v}_{\rm o}$ measurements).
- Measuring cosmological parameters after adopting the kinematic interpretation of the CMB (Hereafter, $\Omega_{\rm i}$ measurements), in other words assuming the peculiar velocity of the observer is given by the Planck measurement (Planck Collaboration: Aghanim et al., 2020b).

The procedure for each of these measurements using each $\mathcal{O}$ catalogue is outlined below:

i) We apply the ARB to the observed $\mathcal{O}$ catalogues of both surveys using five distinct velocities, shown in Table 7.1.

ii) Utilizing the FKP estimator, we compute the power spectrum monopole, denoted as $\hat{P}_0^{\rm art,\mathcal{O}}(k;\boldsymbol{v}_{\rm art})$ after each application, followed by rebinning the spectra.

iii) The rebinned spectra resulting from the ARBs with different velocities are concatenated to form the combined power spectrum monopole $\hat{P}_0^{\rm art,\mathcal{O}}(k;\boldsymbol{v}_1..\boldsymbol{v}_5)$, where the indices each refer to one the shifts shown in Tab. 7.1.

iv) Next, we construct a data vector $\boldsymbol{d} = \hat{P}_0^{\rm art,\mathcal{O}}(k;\boldsymbol{v}_1..\boldsymbol{v}_5) - \mathrm{E}[P^{\mathcal{G}}(k)],.$

v) We employ a MCMC algorithm to extract the parameter of interest using the data vector and the model presented in Eq. (7.51).

7.5.1 Details of the MCMC Analysis.

As outlined in Sec. 4.4.2, the MCMC algorithm requires three inputs to constrain the parameters of interest. Namely, the data vector d, the likelihood function $\mathcal{L}$ which depends on both the data vector and its corresponding model parameters (the model $\boldsymbol{m}$ and covariance C), and the parameters of interest Σ.

The data vector

The data vector is constructed from the concatenated spectra of the $\mathcal{O}$ catalogues after the ARBs

$$\boldsymbol{d} = \hat{P}_0^{\rm art,\mathcal{O}}(k;\boldsymbol{v}_1..\boldsymbol{v}_5) - \mathrm{E}[P^{\mathcal{G}}(k)],. \tag{7.55}$$

The FOTO signal's wavenumber range varies depending on the survey and the considered redshift, as evident from Fig. 7.13. Consequently, when constructing d, we adopt different binning strategies. For the *Euclid* survey, we consistently utilize the largest accessible redshift range, $z \in (0.9, 1.8)$, and bin the spectra into five bins within the wavenumber range $k \in (0.39, 5.34) \times 10^{-3}\,h^{-1}\,\mathrm{Mpc}$.

In contrast, for the SKAO2 survey, we employ 6 bins within the range $k \in (0.39, 8.9) \times 10^{-3}\,h^{-1}\,\mathrm{Mpc}$ for the redshift bin $z \in (0.1, 1.0)$ in the $\boldsymbol{v}_{\rm o}$ measurement, while we opt for 5 bins within the range $k \in (0.39, 5.1) \times 10^{-3}\,h^{-1}\,\mathrm{Mpc}$ for the redshift bin $z \in (0.8, 1.0)$ in the $\Omega_{\rm i}$ measurements.

The likelihood function

In all cases, we assume a Gaussian likelihood

$$\mathcal{L}(\Sigma|\boldsymbol{d}) = \frac{1}{(2\pi)^{n/2}\det \mathsf{C}} \exp\left\{\frac{-\left[\boldsymbol{d}-\boldsymbol{m}(\Sigma)\right]^{T}\mathsf{C}^{-1}\left[\boldsymbol{d}-\boldsymbol{m}(\Sigma)\right]}{2}\right\}, \tag{7.56}$$

as we are working at extremely large scales where the overdensity is expected to be Gaussian. We compute the covariance of the data numerically using

$$\begin{aligned}\boldsymbol{C} &= \mathrm{E}\left\{[\boldsymbol{d}-\mathrm{E}(\boldsymbol{d})][\boldsymbol{d}-\mathrm{E}(\boldsymbol{d})]^{\mathrm{T}}\right\} \\ &= \mathrm{E}\left\{[\hat{P}_0^{\mathrm{art},\mathcal{O}}-\mathrm{E}(\hat{P}_0^{\mathrm{art},\mathcal{O}})][\hat{P}_0^{\mathrm{art},\mathcal{O}}-\mathrm{E}(\hat{P}_0^{\mathrm{art},\mathcal{O}})]^{\mathrm{T}}\right\}.\end{aligned} \tag{7.57}$$

Subsequently, We correct the precision matrix C^{-1} by dividing by the factor $(N_{\mathrm{real}}-1)/(N_{\mathrm{bins}}-n-2)$ (Kaufman, 1967, see also Hartlap et al. (2007) for a first application in cosmology). As for the model, we use Eq. (7.51) for all measurements.

The model parameter set

We employ different assumptions on the terms of the model $\boldsymbol{m}$ that are known *apriori* thus changing Σ depending on the application. However, in all applications, we assume the functional forms of the magnification and Evolution biases are known. For all the MCMC chains discussed below, we employ the public code emcee (Foreman-Mackey et al., 2013).

A note regarding the ARBs.

The selection of peculiar shifts listed in Table 7.1 follows a predominantly random approach, ensuring primarily that they point in diverse directions across the sky. Notably, we report solely the angle formed by the boost velocity and the true velocity, as this angle is the key factor influencing the signal (refer to Eq. (7.51)). We limit the selection to five shifts to minimize the bias and variance of numerically estimated covariance (Hartlap et al., 2007). However, using fewer shifts degrades the FOTO signal's ability to constrain the direction of the observer velocity.

Additionally, we impose a cap on the magnitude of the boost, restricting it to $3v_{\mathrm{o}}$. This restriction serves dual purposes: Firstly, for the velocity vector measurements, it ensures that the artificial signal remains comparable in magnitude to the original signal, which is essential to constrain the velocity magnitude. Secondly, for the cosmological parameters' measurements, a higher amplitude shift may marginally enhance the precision of the extracted parameters, however, it introduces a bias in the parameter constraints when accounting for redshift errors.

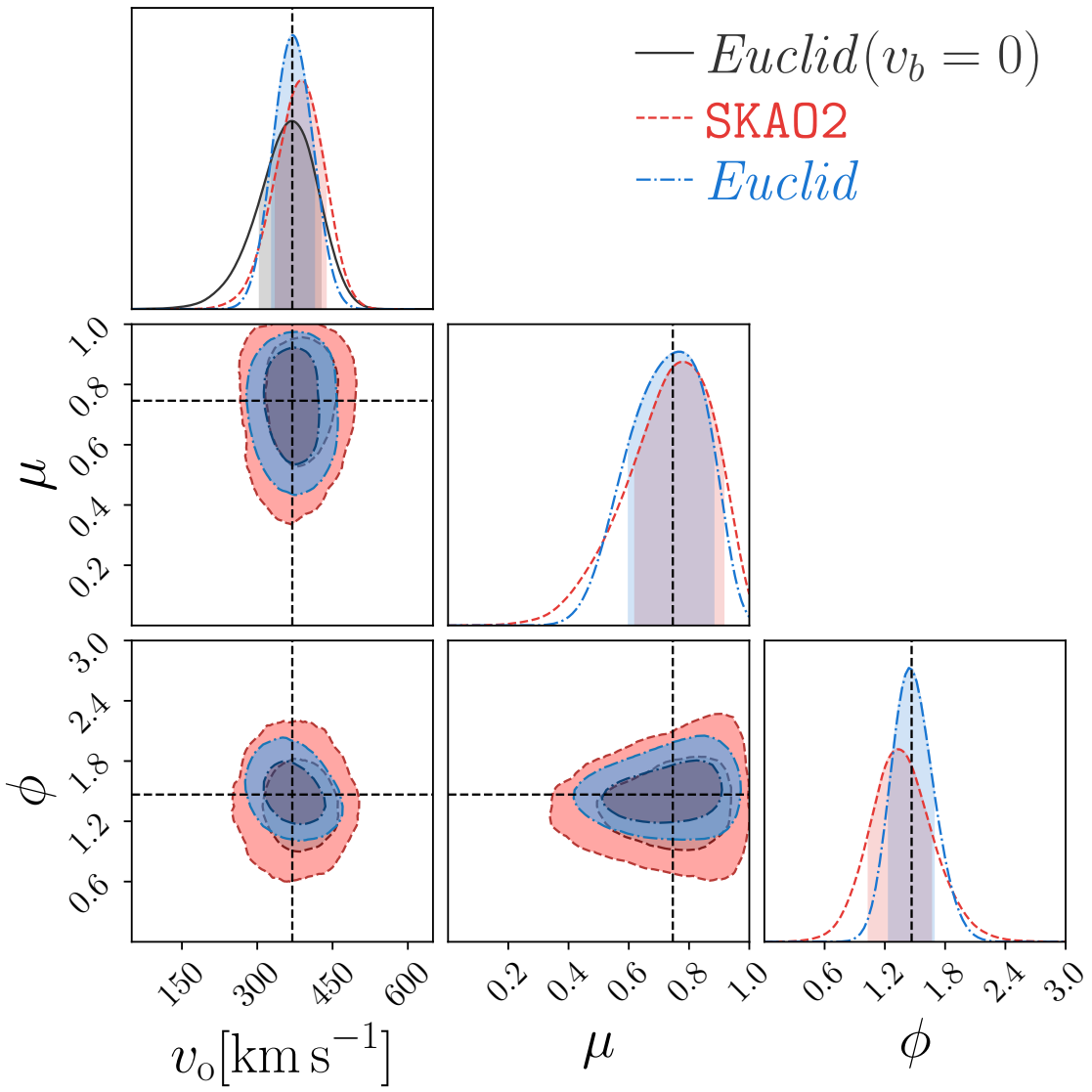


Figure 7.14: The 68-95% credibility contours for the observer velocity vector components measurements are plotted for the two example surveys: *Euclid* (blue) and SKA02 (red). We also show the constraints on the observer velocity magnitude without ARBs in grey in the top left panel.

7.5.2 The velocity vector.

We start with the measurements of the observer's peculiar velocity. In order to measure the velocity vector, we assume a flat cosmology with a known $\Omega_{\mathrm{m},0}$. Therefore, the only parameters left to determine in Eq. (7.51) are the observer velocity components. We represent the observer velocity by its magnitude v_o and two directional parameters: μ, which represents the cosine of the azimuthal angle, and ϕ, which represents the polar angle. In summary, the parameters of the MCMC chain are $\Sigma = \{v_\mathrm{o}, \mu, \phi\}$.

For the priors, the angular parameters are allowed to vary across the entire sky while for the velocity magnitude, we use a Maxwell distribution (Sheth & Diaferio, 2001, Hamana et al., 2003) given by

$$\mathcal{P}(v_\mathrm{o}) = \frac{2}{\pi}\frac{\exp\left(-v_\mathrm{o}^2/2\sigma^2\right)}{\sigma^3}\,, \tag{7.58}$$

with $\sigma = 289\mathrm{km\ s^{-1}}$.

We begin by plotting the posterior distributions for one realization from each example survey in Fig. 7.14. It is clear that the outlined procedure successfully constrains both the magnitude and direction of the observer's velocity. A notable observation is that the *Euclid* survey exhibits greater constraining power compared to the SKAO2 survey. This difference is traced to the higher overall signal present in the *Euclid* survey, as demonstrated in Fig. 7.9.

We repeat this analysis for all $\mathcal{O}$ realizations, presenting the constraints on the velocity magnitude for one redshift bin from each survey in Fig. 7.15. In the figure, the 68% highest probability density (HPD) regions are plotted to illustrate the constraints, with the fiducial value denoted by the black vertical line. The higher significance of the *Euclid* survey's constraints is clear, due to its inherently higher signal. This indicates that the accuracy of the measured signal strongly depends on the survey's intrinsic characteristics, represented by the $\mathcal{Q}$ and $\mathcal{E}$ functions, which consequently influence the redshift behaviour depicted in Fig. 7.2.

Averaging overall realizations, we report that the observer velocity magnitude can be determined with an average precision of 21% for the *Euclid* survey, while the SKAO2 survey yields a slightly lower average precision of 26%.

For the directional estimates, we display the combined posterior distributions in Fig. 7.16. Each pixel's colour indicates the number of realizations where the pixel falls within its 68% highest posterior density (HPD) region. The true direction of $\hat{\boldsymbol{v}}_o$ is denoted by $\circ$, while the direction measured by National Radio Astronomy Observatory (NRAO) Sky Survey (NVSS, Blake & Wall, 2002) is marked by X.

Both the *Euclid* (top plot) and SKAO2 (bottom plot) surveys successfully recover the direction of the velocity observer vector in most realizations. However, the *Euclid* survey demonstrates greater constraining power compared to SKAO2. Similar to the magnitude constraints, this enhancement in constraint quality can be attributed to the higher signal-to-noise ratio in the *Euclid* survey data.

We then examine the posterior distributions associated with the directional angle

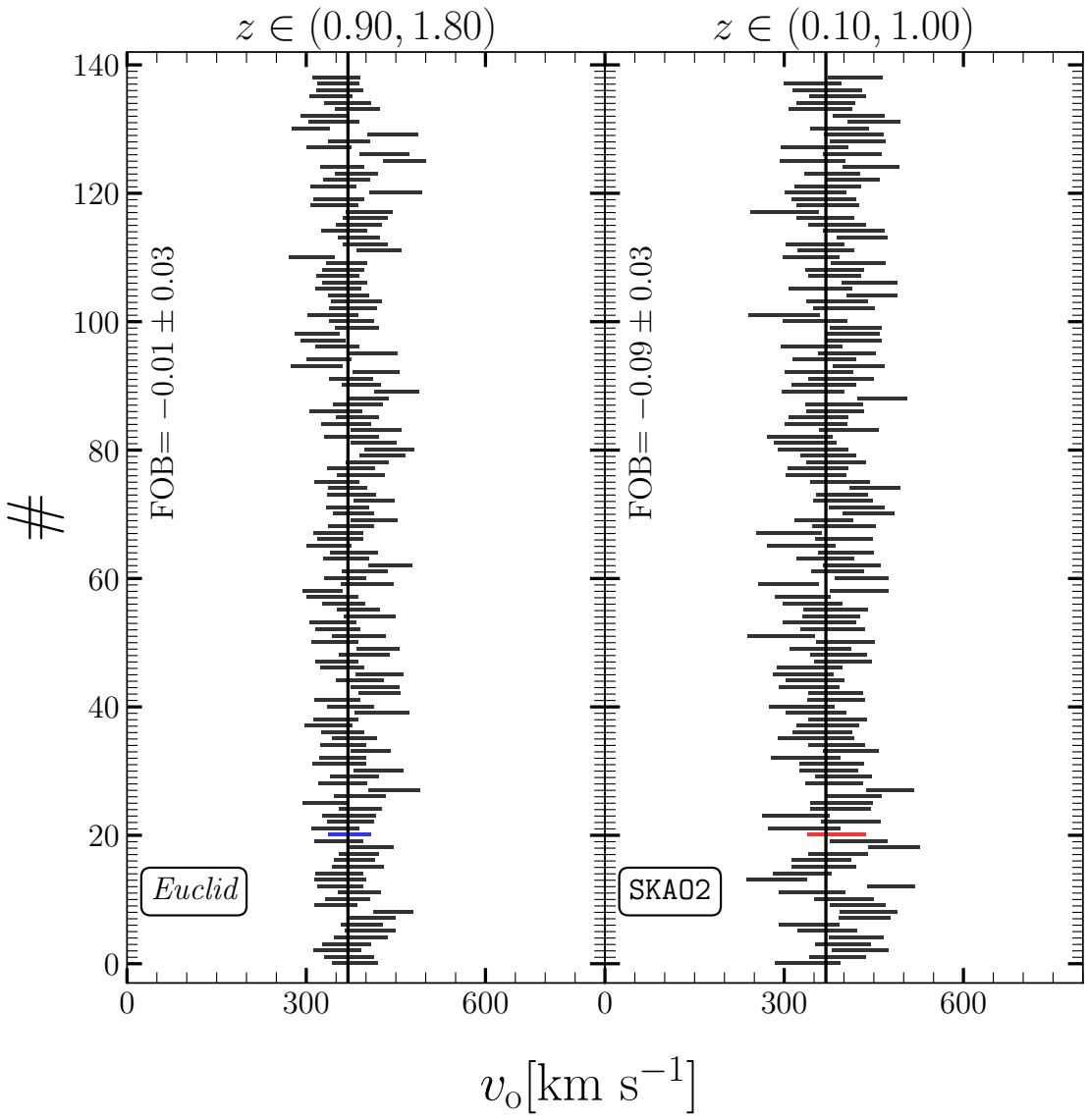


Figure 7.15: The 68% highest probability density (HPD) intervals obtained from the MCMC results for the observer velocity magnitude are depicted for each realization. In the figure, the results corresponding to the *Euclid* selection function are displayed on the left, while those for the SKA02 selection function are shown on the right. The true value is indicated by a vertical black line. Additionally, the average figure of bias (FOB) value is computed across all realizations. We mark the distributions plotted in Fig. 7.14 for the *Euclid* (SKA02)in blue (red) for clarity.

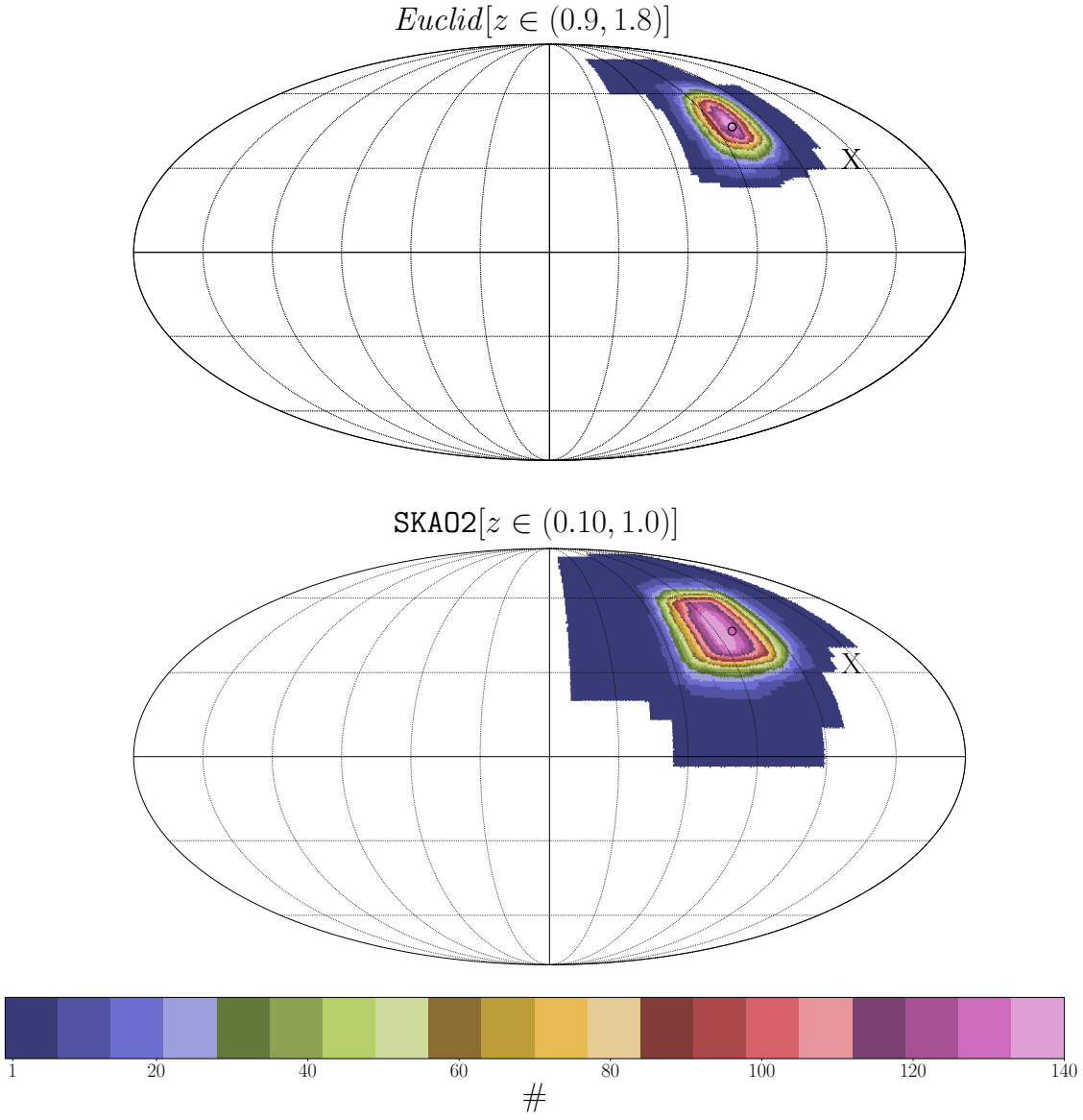


Figure 7.16: The combined posterior distributions of the directions for all realizations are presented for both the *Euclid* (top) and SKA02 (bottom) surveys. The colour scheme reflects the frequency of realizations. In other words, it counts how many times each pixel falls within the 68% HPD interval of the different realizations. The true direction is denoted by a circle (○), while the directions inferred by the NVSS survey (Blake & Wall, 2002) are marked by an X and a black triangle (▲).

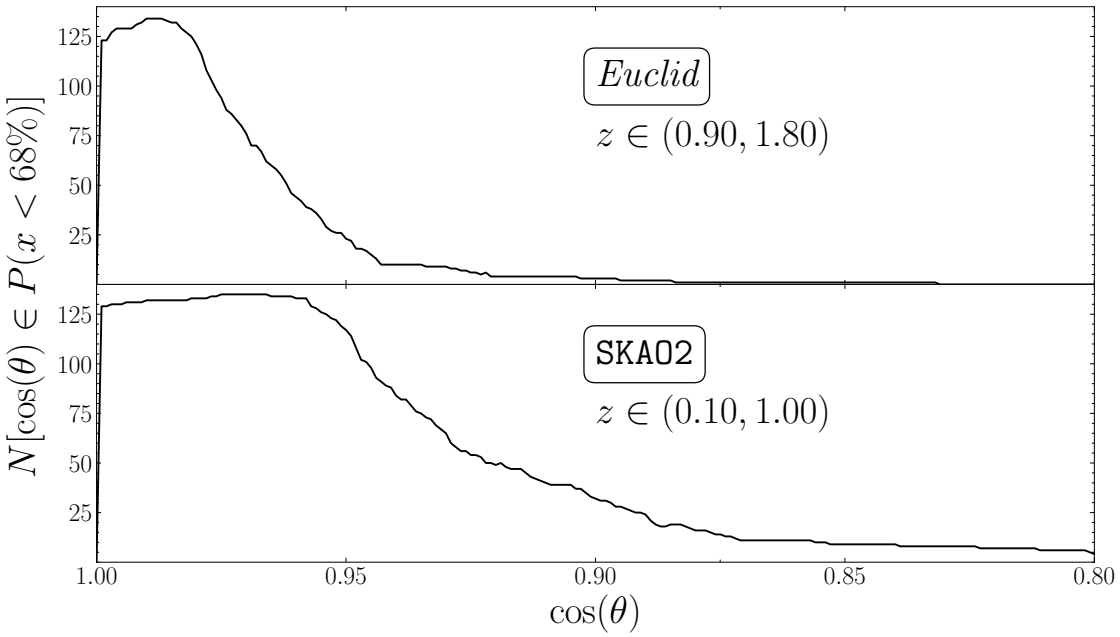


Figure 7.17: The number of realizations that include $\cos(\theta)$ in their 68% HPD regions are plotted on the y-axis for *Euclid* (left) and SKA02 (right).

offset. Here, we define the angle θ as the arccosine of the dot product between the true velocity observer unit vector $\hat{\boldsymbol{v}}_{\rm o,true}$ and the estimated velocity observer unit vector $\hat{\boldsymbol{v}}_{\rm o,est}$, given by:

$$\theta = \cos^{-1}(\hat{\boldsymbol{v}}_{\rm o,true} \cdot \hat{\boldsymbol{v}}_{\rm o,est}), . \tag{7.59}$$

To visualize the distribution of θ, we create Fig. 7.17. Where the y-axis represents the number of realizations that include each value of $\cos(\theta)$ in the 68% HPD. We compute that quantity for both the *Euclid* and SKA02 selection functions, positioned on the left and right sides of the plot, respectively. We deduce from the Figure that we are able to accurately recover the direction of the velocity of the observer.

In summary, our analysis demonstrates that across the majority of realizations, we are able to constrain both the magnitude and the direction of the observer velocity vector using the FOTO signal.

7.5.3 Constraining cosmological parameters.

In this study, we also explore the potential of the FOTO effect in constraining cosmological parameters after adopting the Kinematic interpretation of the CMB. In this section, we set the observer's velocity to the precision determined by Planck measurements and use the FOTO effect to constrain the underlying cosmological parameters.

We adopt a flat ΛCDM model, where the Hubble parameter is described by the relation:

$$H^2(z) = H_0^2 \left[\Omega_{\rm m,0}(1+z)^3 + (1-\Omega_{\rm m,0})(1+z)^{3(1+w)}\right] , \tag{7.60}$$

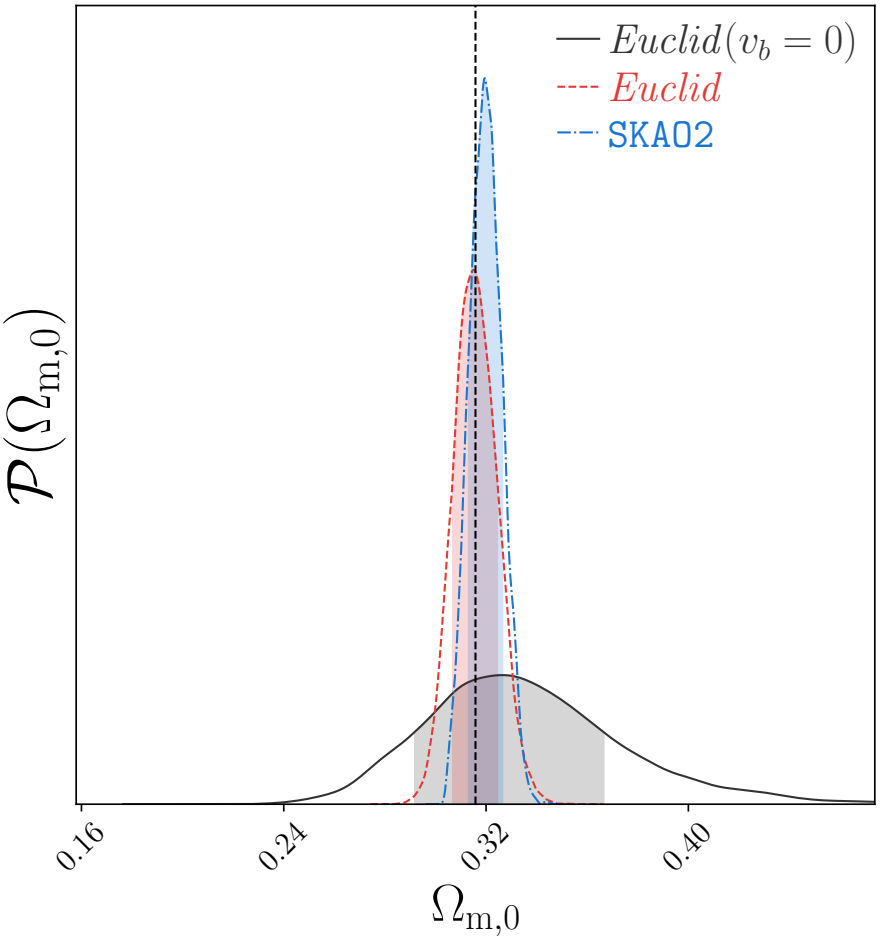


Figure 7.18: The posterior distribution for $\Omega_{\mathrm{m},0}$ for one realization for the *Euclid* survey (blue) and the SKA02 (red). We also show the constraint of the *Euclid* without applying the ARBs in grey.

where H_0 is the present-day Hubble constant, $\Omega_{\mathrm{m},0}$ represents the matter density parameter, and w denotes the equation-of-state parameter for dark energy.

We aim to fit two distinct models: initially, we assume a cosmological constant with $w = -1$, focusing solely on extracting the matter density parameter from the simulated data. Subsequently, we allow both $\Omega_{\mathrm{m},0}$ and w to vary, intending to assess whether this method can effectively constrain the evolution of dark energy.

Measuring $\Omega_{\mathrm{m},0}$.

To estimate the present-day matter density parameter $\Omega_{\mathrm{m},0}$, we employ a four-parameter model for the boosted FOTO signal. This model encompasses three parameters for the velocity vector components and one for the matter density parameter. Hence, our parameter set is defined as $\Sigma = \{\Omega_{\mathrm{m},0}, v_{\mathrm{o}}, \mu, \phi\}$.

For the velocity vector parameters, we apply Gaussian priors with means and standard deviations obtained from measurements reported in Planck Collaboration: Aghanim et al. (2020b). Regarding the matter density parameter $\Omega_{\mathrm{m},0}$, we adopt a

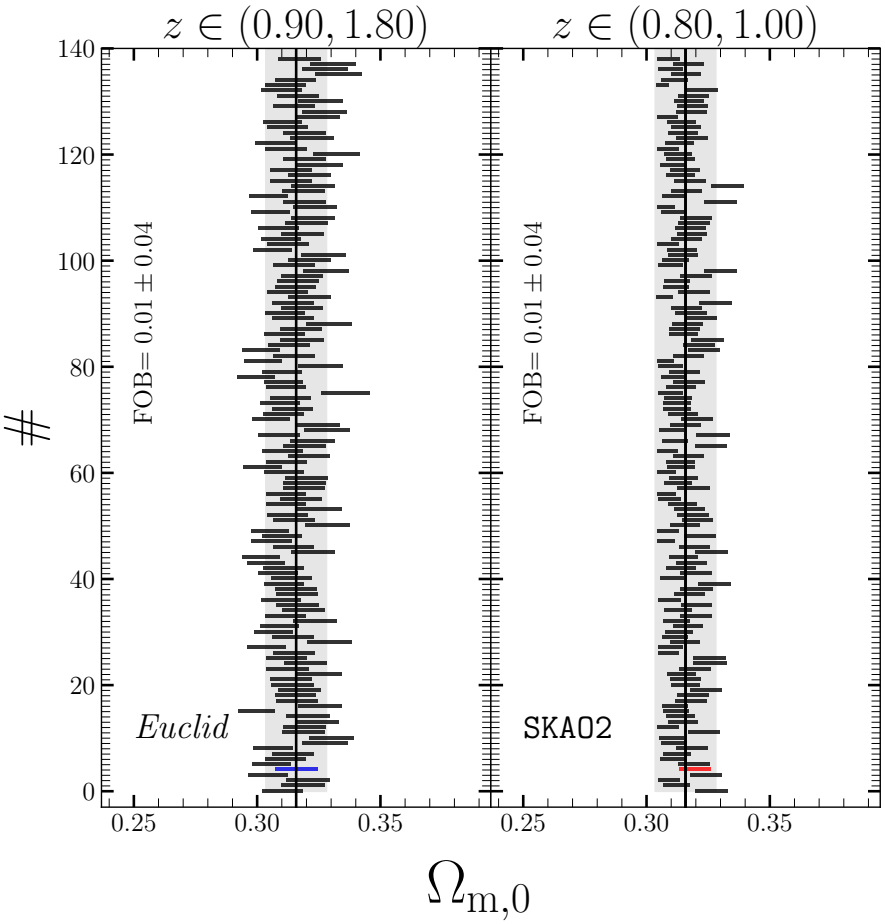


Figure 7.19: Similar to Fig.7.15, this plot displays the constraints on $\Omega_{\mathrm{m},0}$. Additionally, we include the 68% confidence interval for $\Omega_{\mathrm{m},0}$ from the CMB (Planck Collaboration: Aghanim et al., 2020a), represented by a grey band centred around the fiducial value. The blue and red bars correspond to the *Euclid* and SKAO2 distributions, respectively, as depicted in Fig.7.18.

flat prior within the interval $\Omega_{\mathrm{m},0} \in [0, 1]$. This choice of prior indicates that any value within this range is equally likely *apriori*, allowing us to explore the entire parameter space.

We present the resulting constraints from one realization for both surveys in Fig. 7.18. The plot illustrates the *Euclid* constraint without any ARBs in grey, while the constraints for *Euclid* (SKAO2) after incorporating the ARBs are shown in red (blue). The substantial improvement in constraints achieved by boosting the signal using ARBs are clear. The cosmological model's constraint is not solely influenced by the logarithmic derivative $\mathrm{d}\ln H/\mathrm{d}\ln(1+z)$ in α_{o}, but also by the Hubble function and comoving distance terms present in Eq.(7.51).

Additionally, we observe that the constraint from SKAO2 is slightly better than its *Euclid* counterpart. This enhancement can be attributed to the wider frequency range of the SKAO2 FOTO signal, which provides more information for constraining cosmological parameters.

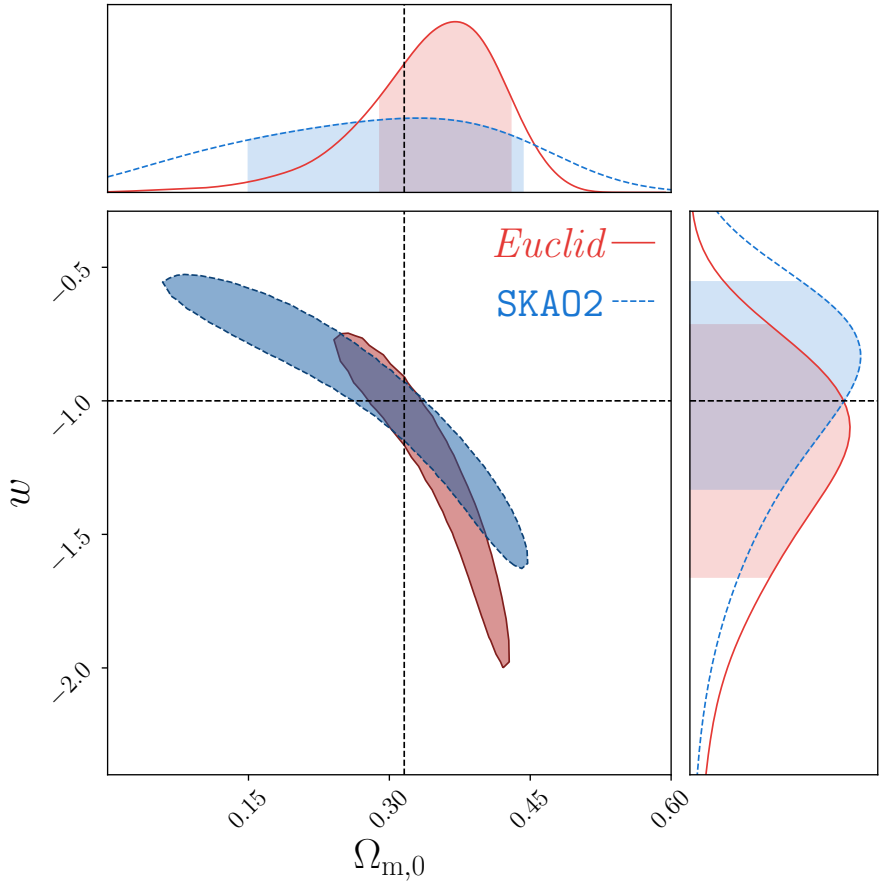
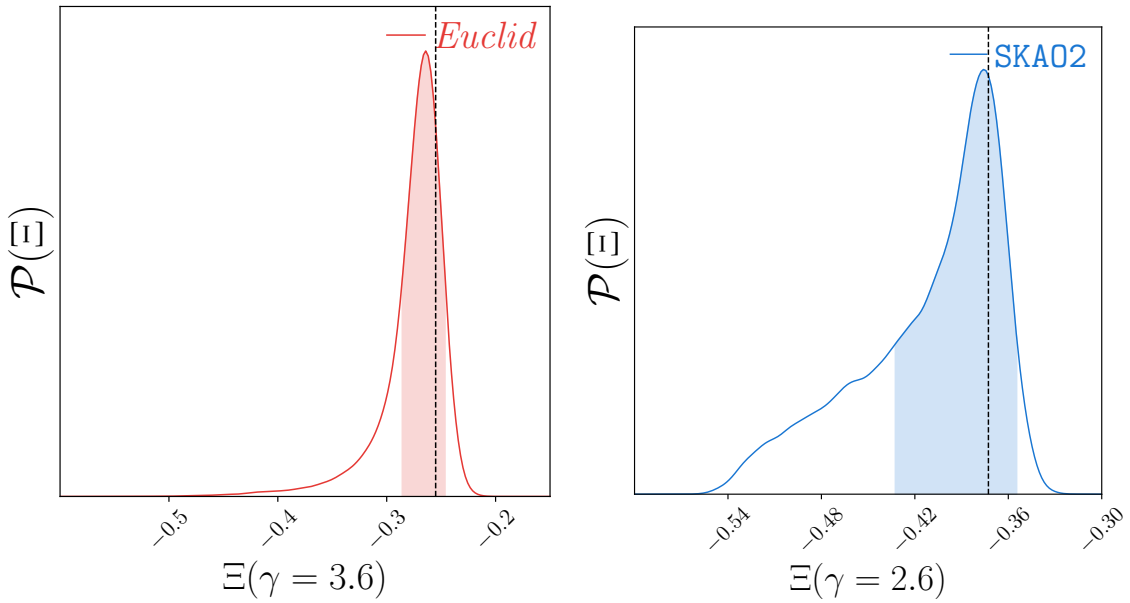


Figure 7.20: The posterior distribution for $\Omega_{m,0}$ and w for one realization for the *Euclid* survey (blue) and the SKA02 (red) surveys.

We derive constraints on the matter density parameter, $\Omega_{m,0}$, from all available realizations. Fig. 7.19 presents the 68% HPD regions of these constraints, with the *Euclid* survey results displayed in the left panel and those from the SKA02 survey spectra shown in the right panel. In addition to the constraints derived from our method, we include the $\Omega_{m,0}$ constraint provided by Planck Collaboration: Aghanim et al. (2020a) as a grey band centred around the fiducial value. The figure illustrates that the constraints obtained through our approach are competitive with the credibility interval derived from the CMB.

It's important to note that the constraints depicted in Fig. 7.19 are optimistic as they rely on measuring clustering at very large scales, which can be affected by various observational systematic effects. Furthermore, these constraints are anticipated to degrade for smaller survey volumes. However, as demonstrated by the SKA02 case, successful extraction of $\Omega_{m,0}$ is possible even when the original signal is not particularly strong. In future surveys, leveraging higher multipoles could further enhance the constraining power of this method using the model derived in Sec. 7.3.



(a) The *Euclid* survey constraints on Ξ. (b) The SKAO2 survey constraints on Ξ.

Figure 7.21: The posterior distribution for the derived parameter $\Xi = (1 - \Omega_m)^\gamma w$ for one realization for both example surveys.

Measuring $\Omega_{\mathrm{m},0}$ and w.

In the last part of this work, we test the possibility of constraining the dark energy equation of state using the FOTO signal. In that case, we allow both $\Omega_{\mathrm{m},0}$ and w in the Hubble equation (Eq. 7.60) to vary. We employ a five-parameter model; three parameters for the velocity vector and two cosmological parameters. Hence, our parameter set is defined as $\Sigma = \{\Omega_{\mathrm{m},0}, \omega, v_{\mathrm{o}}, \mu, \phi\}$. Similar to the $\Omega_{\mathrm{m},0}$ constraints, we use Planck priors for the velocity parameters, we let $\Omega_{\mathrm{m},0}$ and w vary between $\Omega_{\mathrm{m},0} \in [0, 1]$ and $w \in [-6, 1]$ respectively.

An example of the obtained posterior distributions is presented in Fig. 7.20. The constraints from the *Euclid* survey are depicted in red, while those from the SKAO2 survey are shown in blue. The figure reveals a high degree of degeneracy between the two parameters under investigation. It's clear that the FOTO signal can only constrain a combination of these parameters rather than each one independently. Interestingly, the posterior distribution from the SKAO2 survey appears to constrain the parameter w more effectively compared to its counterpart from the *Euclid* survey. Conversely, the roles are reversed in the constraints on $\Omega_{\mathrm{m},0}$.

This discrepancy can be attributed to the different redshift bins utilized in each survey. The *Euclid* survey employs deeper binning where the Universe is matter-dominated, resulting in a stronger constraint on $\Omega_{\mathrm{m},0}$. On the other hand, the redshift bin of the SKAO2 survey is shallower (in DE domination), causing the value of w to exert a greater influence on the FOTO signal. Consequently, we predict that the posterior slope inferred from this method aligns more closely with $\Omega_{\mathrm{m},0}$ as

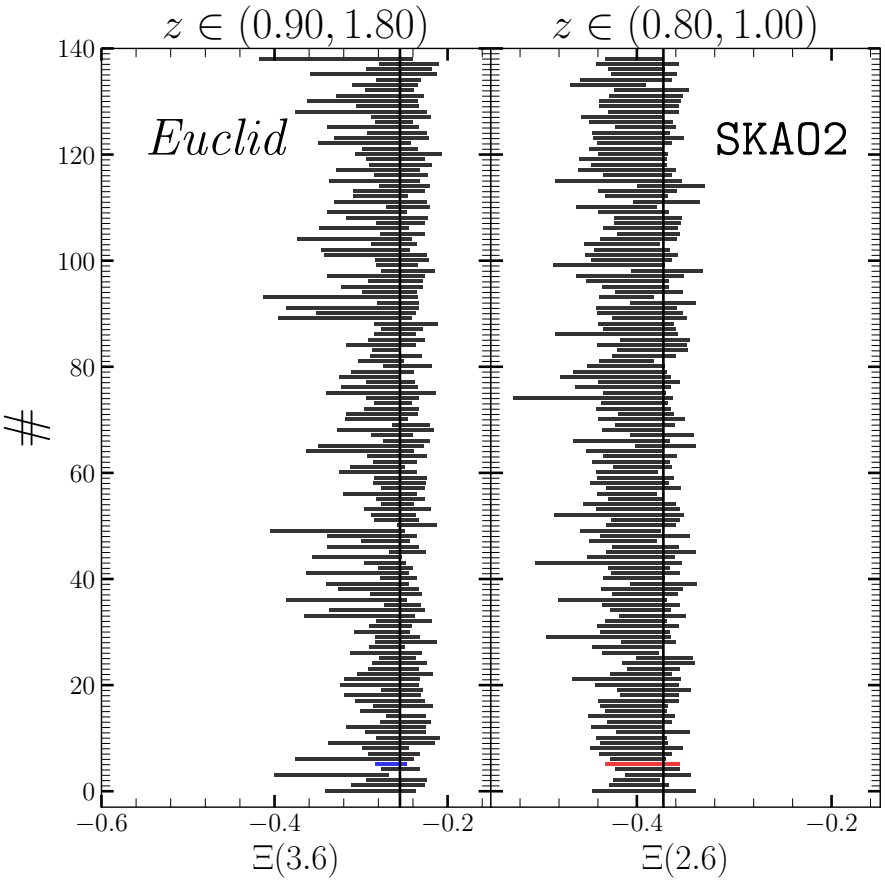


Figure 7.22: We plot the 68% HPD regions for the constraints on the derived parameter $\Xi = (1 - \Omega_m)^\gamma w$. The blue and red marked bars represent the *Euclid* and SKA02 distributions shown in Fig. 7.21, respectively.

the redshift increases (assuming, of course, the parametric relation in Eq. 7.60).

To further elucidate this point, we investigate the constraining power on the derived parameter $\Xi = (1 - \Omega_\mathrm{m})^\gamma w$, where γ is a constant rational number. Based on Fig. 7.20, we predict that γ should be proportional to the mean redshift of the bin utilized in the parameter constraints.

Utilizing the computed MCMC chains, we determine that setting $\gamma = 3.6$ yields the highest constraining power on Ξ for the *Euclid* survey, while $\gamma = 2.6$ produces the strongest constraints for the SKA02 survey. The posterior distribution of Ξ is illustrated in Fig. 7.21. As seen in the Figure, the constraints on Ξ are tighter compared to its individual components. Additionally, $\mathcal{P}(\Xi)$ exhibits a long tail towards more negative values. We attribute this tail to the curved nature of the combined $\mathcal{P}(\Omega_\mathrm{m}, w)$.

We show the constraints over all realizations for both surveys in Fig. 7.22. We plot the 68% HPD density in black for the *Euclid* survey on the left and the SKA02 on the right. The blue and red bars mark the posteriors shown in Figs. 7.21 and 7.22. We can see from the figure that the tail in the $\mathcal{P}(\Xi)$(shown in Fig. 7.21) leads to asymmetric constraints for almost all realizations.

7.6 Summary

The velocity of the observer imprints an oscillatory signature signal on the observed power spectrum multipoles as seen in Sec. 7.2 that we name the Finger Of The Observer effect. We derived the analytical relation that governs this effect for the monopole of the power spectrum as well as all higher multipoles of the observed power spectrum (see Sec. 7.3). We found that the FOTO signal is detectable in most realizations (see Fig. 7.10) assuming the observer velocity $v_o \geq v_\odot$, where $v_\odot$ is the Planck measurement of the heliocentric velocity. In Sec. 7.2.2, we also measured the dependence of the signal-to-noise of the FOTO signal (assuming $v_o = v_\odot$) on the angular footprint of the survey. We found that the S/N is directly proportional to the fraction of the sky. However, we note that the signal is detectable with S/N ≈ 4 in the smallest fraction considered, corresponding to the *Euclid* footprint.

We then tested the possibility of removing the FOTO signal by a radial redshift transformation in Sec. 7.4. We found that even if you shift the redshift by the exact velocity of the observer, the FOTO signal is only cancelled for $z < 0.4$. On the other hand, the FOTO signal is actually boosted for all higher redshifts. This can be understood by the fact that this redshift transformation only removes the line of sight contribution of the observer-induced RSD, while the relativistic aberration and magnification effects remain.

In Sec. 7.5, we made use of that boost in the FOTO signal to tackle the second main objective of this work; extracting cosmological information from relativistic RSD. We wrote down and verified numerically the impact of these redshift boosts on the power spectrum monopole. We then shifted the particles in the catalogues of the two example surveys with five different velocities. After measuring the power spectrum monopole for the shifted catalogues, we combined the power spectrum monopoles of the different boosted catalogues and used it to:

- extract the magnitude and direction of the velocity of the observer,
- constrain cosmological parameters.

We found that for both of our example surveys, we were able to constrain both the magnitude and direction of the velocity of the observer (see Figs. 7.14, 7.15, and 7.16), with the *Euclid* survey providing stronger constraints due to its inherently higher FOTO signal (see Fig. 7.9). Moreover, adopting the kinematic interpretation of the CMB, we tested the potential of the FOTO signal in measuring the matter density parameter. In optimistic scenarios, we found that the derived constraints on $\Omega_{m,0}$ are competitive to the CMB as shown in Fig. 7.19. Finally, we have shown that the FOTO signal can be used to extract valuable information using more complicated cosmological models as seen in Fig. 7.22.

In the work presented in this chapter, we focused on the monopole. However, adopting a tomographic approach and combining different multipoles can significantly enhance the signal, especially when the assumptions that the magnification and evolution biases are perfectly known are relaxed. Combining different surveys

that span a larger redshift range can be used to constrain the redshift evolution of DE. As for the higher multipoles, the odd multipoles offer a distinct advantage: the $P_\ell^{\rm com}(k)$ signal is substantially lower than that of the power spectrum monopole. This is primarily because most of the signal is traced to the wide-angle contributions, which diminish at higher redshifts, and to the leakage of even multipoles due to the asymmetric choice of the estimator (Castorina & White, 2018). Consequently, it is possible to isolate the FOTO signal with a higher signal-to-noise ratio in these multipoles.

Chapter 8

Summary and Outlook

Redshift Space Distortions (RSD) shape the galaxy maps that we construct of the Universe. Traditionally, in earlier surveys, the modelling of RSD was confined to accounting for the peculiar velocities of sources under the GPP (or more recently the LPP) approximation. However, that approach neglects the relativistic and lightcone effects induced by the gravitational potential encountered by photons from distant galaxies as they traverse from the source to the observer. While these effects are relatively minor at smaller scales, they become increasingly significant when exploring larger cosmic scales. With the advent of Stage IV surveys, which aim to probe unprecedentedly vast expanses of the Universe, understanding the influence of these relativistic RSD becomes paramount.

This thesis addresses two pivotal inquiries. Firstly, we aim to understand the extent to which relativistic RSD impact the two-point statistics measured from Stage-IV surveys. Secondly, we investigate the feasibility of leveraging these effects as tools for measuring cosmological parameters and advancing our understanding of the Universe.

The `LIGER` method

In addressing both inquiries, we utilise and improve the `LIGER` method, as outlined in Chapter 3, which is a computational framework for constructing galaxy catalogues tailored to specific surveys after accounting for relativistic RSD at linear order in cosmological perturbations. The `LIGER` code implements a coordinate transformation between the real-space positions of galaxies and their observed counterparts. We only utilize the large-box mode of the `LIGER` method, which employs this coordinate transformation to generate lightcones of DM particles. These dark matter particles' original real-space positions are determined through low-resolution N-body simulations.

As part of this thesis, we improve the coordinate shift mechanism within the `LIGER` code and introduce a supplementary toolkit known as `BUILDCONE`. This toolkit complements the `LIGER` method by ingesting the DM light cones generated through the coordinate transformation and producing galaxy catalogues tailored to

the specifications of the target survey.

We use the LIGER method to build 140 mock galaxy catalogues for two example surveys:
(i) The Euclid survey:
The Euclid Wide Spectroscopic Sample aims to map nearly 30 million emission-line galaxies across approximately one-third of the celestial sphere, spanning a redshift range of $0.9 \lesssim z \lesssim 1.8$. With careful management of observational systematic effects, *Euclid* presents a unique opportunity to investigate galaxy clustering on unprecedentedly large cosmic scales.
(i) The SKA02 survey:
Anticipated as the "Phase 2" of the SKAO mission, the envisioned SKA02 survey aspires to cover an expansive angular footprint spanning 30,000 deg^2. This ambitious follow-up campaign is poised to extend its reach across a broader range of cosmic history, targeting galaxies within a redshift interval of $0.1 < z < 2.0$. The expected output of this project includes the detection of 1 billion galaxies.

We investigate the influence of various relativistic RSD by turning specific RSD effects on and off during the construction of the galaxy catalogues. Specifically, for each of the 140 simulations, we generate the galaxy distribution in real space ($\mathcal{R}$) along with three distinct versions of its redshift space counterpart. In the $\mathcal{V}$ realizations, we solely incorporate the spatial derivative of the peculiar velocity of galaxies along the line of sight. Conversely, the $\mathcal{G}$ catalogues include all integral terms, primarily influenced by the weak lensing phenomenon (including both the magnification bias and convergence). Lastly, the $\mathcal{O}$ simulations account for distortions resulting from a non-zero peculiar velocity of the observer, calibrated to match the observed dipole temperature anisotropy of the cosmic microwave background or assigned random observer velocities in the $\mathcal{O}^{\mathrm{R}}$ catalogues. The $\mathcal{O}$ catalogues closely mimic the observed catalogues expected in real-world surveys.

We validate the generated mock catalogues by cross-referencing the angular power spectra estimated from the $\mathcal{R}$ and $\mathcal{G}$ mocks, against the theoretical predictions computed by two distinct Boltzmann Solvers: CLASS and CAMB. As demonstrated in Section 3.5, we observe a remarkable sub-percent agreement in both assessments within the large-scale regimes pertinent to this study.

Wide Angle effects.

In Chapter 5, We concentrate on evaluating the significance of wide-angle effects in contrast to various approximations commonly used in modelling the power spectrum. We compare these models to power spectrum multipoles of the $\mathcal{V}$ *Euclid* mocks which are computed using the estimators presented in Chapter 4. We conduct two distinct analyses. Initially, we compare the power spectrum monopole of the *Euclid* mocks using the *Euclid* footprint (see Fig. 3.5) against predictions derived from the GPP Kaiser model which assumes that all galaxies have a single constant line of sight. We find that wide-angle corrections – sourced by the Kaiser velocity gradient term – cause a statistically significant enhancement of power for

modes with $k < 0.02, h$ Mpc^{-1} compared to the GPP model.

Subsequently, we revisit the analysis, this time comparing the power spectrum multipoles of the *Euclid* mocks to predictions derived from the LPP Kaiser model, which partially accounts for the variable lines-of-sight of different galaxies. To construct the LPP prediction, we rely on the mixing-matrix framework outlined in Eqs. (5.8) and (5.9). We note that we apply this framework to all disconnected patches of the *Euclid* survey in contrast to treating them separately. Consequently, the elements of the mixing matrix exhibit high-frequency oscillations (see Fig. 5.5) due to the disjoint nature of the *Euclid* footprint. However, these oscillations have little impact on the power spectrum predictions as our results show a remarkable agreement between the LPP Kaiser model and the power spectrum multipoles of the $\mathcal{V}$ *Euclid* mocks, as depicted in Fig. 5.6. This success of the LPP model indicates that the discrepancy found in the GPP Kaiser model can be substantially alleviated by using the LPP Kaiser model. It also simultaneously displays the accuracy of the `LIGER` mock catalogues,

Relativistic RSD in *Euclid*.

In Chapter 6, we answer the first of the two main research questions of this work. We adopt the *Euclid* survey alongside its corresponding footprint as a case study. Our focus is centred on quantifying the significance of the weak lensing and observer velocity signals in three distinct two-point statistics. To that end, we use the estimators presented in Chapter 4 to compute the two-point statistics for all different versions of the galaxy catalogues.

Subsequently, we employ the likelihood-ratio test, as detailed in Section 4.4.1, to measure the statistical significance that the *Euclid* measurements which include particular RSD effects reject the models that neglect these effects. For instance, we quantify the impact of observer velocity effects by comparing the likelihoods that the $\mathcal{O}$ mocks are realizations of the models derived from averaging the clustering signal across the $\mathcal{O}$ and $\mathcal{G}$ mocks. Our results for each two-point statistic are summarized as follows.

(i) Multipole moments of the power spectrum: The peculiar velocity of the observer leaves distinct imprints on very large scales ($k < 0.015, h$,Mpc^{-1}). However, these imprints are challenging to detect with the *Euclid* survey within narrow redshift bins due to the substantial statistical errors at these scales. Across all tomographic redshift bins we explore, the signal-to-noise ratio (S/N) consistently remains around unity (Tab. 6.2). On the other hand, the weak gravitational lensing signal has a stronger signal that spans a large k-range. The weak lensing signal is particularly important within the tomographic slice $1.5 < z < 1.8$, where it achieves a detection significance of S/N = 2.1. Combining the two effects results in a maximum S/N = 2.4. We additionally compare the $\mathcal{O}$ power spectrum multipoles with the LPP Kaiser model, revealing that within the redshift bin $1.5 < z < 1.8$, the model is rejected with a S/N = 2.9. This result articulates the importance of the weak lensing term at high redshift and on large scales.

(i) Angular power spectrum:

The observer's peculiar velocity notably enhances the amplitudes of the first three odd multipole signals when measured in the heliocentric frame (ee Fig. 6.8). This augmentation can be quantified with a signal-to-noise ratio of 2.1 for the redshift bin covering the entire survey (Tab. 6.3). Additionally, the weak lensing signal boosts the clustering signal for $\ell \gtrsim 50$. This effect is particularly pronounced in the wide redshift bin $0.9 < z < 1.8$, where it is discernible with a signal-to-noise ratio of 5.4 (see Fig. 6.10). Cross-correlating narrower bins at greater distances yields a slightly lower statistical significance (up to S/N = 4.5).

(ii) Multipole moments of the 2PCF:

The observer's peculiar velocity has little impact on the clustering signal (as seen in Fig. 6.12), with the likelihood-ratio test consistently yielding a signal-to-noise ratio lower than unity. Additionally, the weak lensing effect enhances the quadrupole and hexadecapole moments of the 2PCF at large scales ($r > 150\,h^{-1}\,\mathrm{Mpc}$) for redshifts exceeding $z > 1.3$ (refer to Fig. 6.11). The combined contribution of lensing and observer's velocity leads to a S/N that increases with redshift, peaking at 2.5 within the furthest tomographic bin considered ($1.5 < z < 1.8$).

The FOTO Signal and Cosmological Measurements.

In Chapter 7, we delve into the second major inquiry of this study. Here, we utilise the FOTO effect, which is a characteristic oscillatory signal present across all power spectrum multipoles, due to the peculiar velocity of the observer. Our investigation spans both the *Euclid* and `SKA02` surveys assuming a full-sky footprint.

We first analytically derive the FOTO signal for all power spectrum multipoles. Subsequently, we validate our theoretical model, focusing on the power spectrum monopole. This validation is conducted in two distinct scenarios: one that uses an ensemble of observers with varying observer velocities (the $\mathcal{O}^{\mathrm{R}}$ mocks), and the other featuring mocks with the observer velocities set to the Planck value (the $\mathcal{O}$ mocks). We find excellent agreement between our analytical model and the spectra estimated from the mocks across all scales and redshifts considered in both cases.

We then quantify the signal-to-noise ratio of the FOTO effect for the survey exhibiting the strongest signal (the *Euclid* survey), under various angular masks. As anticipated, we observe an increase in the S/N ratio with the fraction of the sky covered in the footprint. However, we find that the signal remains detectable with an approximate S/N of 4 even for a survey covering only a third of the sky.

We then investigate whether the FOTO signal can be cancelled through a radial redshift transformation. We find that even if the redshift is shifted precisely by the observer's velocity, the FOTO signal is only removed for redshifts below 0.4. Conversely, for higher redshifts, the FOTO signal is actually amplified. We explain that observation by analytically showing that such redshift transformations solely eliminate the line-of-sight contribution of the observer-induced RSD, while relativistic aberration and magnification effects persist.

We then leverage this boost in the signal due to radial redshift transformations to

extract cosmological parameters. We first derive and numerically verify the impact of these redshift boosts on the power spectrum monopole. Subsequently, we outline the following procedure to use radial redshift transformations for cosmological measurements: shifting galaxies within our example survey catalogues with varying velocities, measuring the resultant power spectrum monopole, and then employing standard parameter inference techniques (MCMC chains) to extract parameters from the amplified spectra. Through this process, we are able to measure **both** the magnitude and direction of the observer's velocity from the shifted mocks, as well as constrain cosmological parameters such as the dark matter density parameter in optimistic scenarios.

Remarkably, for both of our example surveys, we achieve constraints on the observer's velocity with a precision of approximately 25%. Moreover, adopting a kinematic interpretation of the CMB, we find constraints on $\Omega_{\mathrm{m},0}$ that are comparable to those obtained from Planck measurements. Lastly, in optimistic scenarios, we discover that this signal holds promise for extracting information about the equation of state of Dark Energy.

Outlook

There are numerous avenues for extending the scope of this work. Firstly, while our primary focus has been on assessing the significance and utility of relativistic RSD, our mock catalogues have overlooked observational systematics and redshift errors. To bridge this gap between theory and practical application, it's imperative to consider the impact of observational angular systematics such as Galactic extinction, particularly as they influence galaxy clustering at the same large scales where relativistic effects are prominent.

Secondly, a natural progression would involve exploring higher-order summary statistics, such as the three-point correlation function (3PCF) or the bispectrum. However, advancing to these higher orders necessitates extending the coordinate transformation of the `LIGER` method into second-order cosmological perturbations to accurately model effects at such scales. One promising avenue, which does not require substantial modifications to existing methodologies, is to investigate the impact of relativistic RSD cross-correlations with other observables such as the CMB temperature maps and weak lensing shear maps. The latter approach can be particularly successful in isolating effects that are sub-dominant in the auto-correlation of each observable, such as the Sachs–Wolfe effect.

Lastly, regarding parameter inference for the FOTO signal, to apply this method to real data, we must extend our approach to catalogues with partial sky coverage. Moreover, leveraging higher multipoles of the power spectrum can significantly enhance the constraints obtained through this technique. These extensions not only broaden the applicability of our findings but also deepen our understanding of the underlying cosmological phenomena.

List of Figures

List of Tables

List Of Acronyms

RSD Redshift Space Distortions x

GPP Global Plane Parallel 3

LPP Local Plane Parallel 4

FOTO Finger Of The Observer 1

ARB Artificial Redshift Boost 152

MCMC Markov chain Monte Carlo 81

HPD highest probability density 157

NRAO National Radio Astronomy Observatory 156

CMB Cosmic Microwave Background 5

GPP Global Plane-Parallel 3

DM cold dark matter 7

DE dark energy 7

Bibliography

Abidi, M. M., Bonvin, C., Jalilvand, M., & Kunz, M. 2023, Phys. Rev. D, 107, 063514

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

Amanatides, J. & Woo, A. 1987, in EG 1987-Technical Papers (Eurographics Association)

Amendola, L., Appleby, S., Avgoustidis, A., et al. 2018, Living Reviews in Relativity, 21

Bacon, D. J., Battye, R. A., Bull, P., et al. 2020, Publications of the Astronomical Society of Australia, 37

Bagley, M. B., Scarlata, C., Mehta, V., et al. 2020, ApJ, 897, 98

Bahr-Kalus, B., Bertacca, D., Verde, L., & Heavens, A. 2021, J. Cosmology Astropart. Phys., 11, 027

Ballinger, W. E., Peacock, J. A., & Heavens, A. F. 1996, MNRAS, 282, 877–888

Bardeen, J. M. 1980, Phys. Rev. D, 22, 1882

Bartolo, N., Matarrese, S., Pantano, O., & Riotto, A. 2010, Classical and Quantum Gravity, 27, 124009

Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, ApJSupplement Series, 208, 20

Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, Phys. Rep., 367, 1

Bernardeau, F. & van de Weygaert, R. 1996, MNRAS, 279, 693

Bertacca, D. 2015, Classical and Quantum Gravity, 32, 195011

Bertacca, D. 2015, Classical and Quantum Gravity, 32, 195011

Bertacca, D. 2020, International Journal of Modern Physics D, 29, 2050085

Bertacca, D., Maartens, R., & Clarkson, C. 2014a, J. Cosmology Astropart. Phys., 2014, 037, ADS Bibcode: 2014JCAP...09..037B

Bertacca, D., Maartens, R., & Clarkson, C. 2014b, J. Cosmology Astropart. Phys., 2014, 013

Bertacca, D., Maartens, R., Raccanelli, A., & Clarkson, C. 2012, J. Cosmology Astropart. Phys., 10, 025

Bertacca, D., Raccanelli, A., Bartolo, N., et al. 2018, Phys. Rev. D, 97, 023531

Bertacca, D., Ricciardone, A., Bellomo, N., et al. 2020, Phys. Rev. D, 101, 103513

Beutler, F., Castorina, E., & Zhang, P. 2019, J. Cosmology Astropart. Phys., 3, 040

Beutler, F. & Di Dio, E. 2020, J. Cosmology Astropart. Phys., 2020, 048

Beutler, F. & McDonald, P. 2021, J. Cosmology Astropart. Phys., 11, 031

Beutler, F., Saito, S., Seo, H.-J., et al. 2014, MNRAS, 443, 1065

Bianchi, D., Gil-Marín, H., Ruggeri, R., & Percival, W. J. 2015, MNRAS: Letters, 453, L11

Blake, C. & Wall, J. 2002, Nature, 416, 150

Blas, D., Lesgourgues, J., & Tram, T. 2011, J. Cosmology Astropart. Phys., 2011, 034

Bonvin, C. & Durrer, R. 2011, Phys. Rev. D, 84, 063505

Bonvin, C. & Fleury, P. 2018, J. Cosmology Astropart. Phys., 2018, 061–061

Bonvin, C., Hui, L., & Gaztanaga, E. 2016, J. Cosmology Astropart. Phys., 2016, 021–021

Borzyszkowski, M., Bertacca, D., & Porciani, C. 2017, MNRAS, 471, 3899

Bosi, M., Bellomo, N., & Raccanelli, A. 2023, J. Cosmology Astropart. Phys., 2023, 086

Breton, M.-A., de la Torre, S., & Piat, J. 2022, A&A, 661, A154

Broadhurst, T. J., Taylor, A. N., & Peacock, J. A. 1995, ApJ, 438, 49

Bruni, M., Hidalgo, J. C., Meures, N., & Wands, D. 2014, The Astrophysical Journal, 785, 2

Bull, P. 2016, ApJ, 817, 26

Canuto, C., Hussaini, M., Quarteroni, A., & Zang, T. 2007, Spectral Methods: Evolution to Complex Geometries and Applications to Fluid Dynamics, Scientific Computation (Springer Berlin Heidelberg)

Castorina, E. & Di Dio, E. 2022, J. Cosmology Astropart. Phys., 01, 061

Castorina, E., Hand, N., Seljak, U., et al. 2019, J. Cosmology Astropart. Phys., 09, 010

Castorina, E. & White, M. 2018, MNRAS, 476, 4403

Challinor, A. & Lewis, A. 2011, Phys. Rev. D, 84, 043516

Chan, K. C. & Blot, L. 2017, Phys. Rev. D, 96, 023528

Chisari, N. E. & Zaldarriaga, M. 2011, Phys. Rev. D, 83, 123505

Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, MNRAS, 362, 505

Dalang, C. & Bonvin, C. 2022, MNRAS, 512, 3895

de Mattia, A. & Ruhlmann-Kleider, V. 2019, J. Cosmology Astropart. Phys., 2019, 036

Delouis, J. M., Puget, J. L., & Vibert, L. 2021, A&A, 650, A82

Di Dio, E., Montanari, F., Lesgourgues, J., & Durrer, R. 2013, J. Cosmology Astropart. Phys., 11, 044

Dodelson, S. 2003, Modern Cosmology (Elsevier Science)

Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, ApJ, 633, 560

Elkhashab, M. Y., Porciani, C., & Bertacca, D. 2021, MNRAS, 509, 1626

Elkhashab, M. Y., Porciani, C., & Bertacca, D. TBDa, To be submitted to JCAP

Elkhashab, M. Y., Porciani, C., Bertacca, D., Salvalaggio, J., & et al. TBDb, Submitted to the ECEB

Ellis, G. F. R. & Baldwin, J. E. 1984, MNRAS, 206, 377

Euclid Collaboration: Blanchard, A., Camera, S., Carbone, C., et al. 2020, A&A, 642, A191

Euclid Collaboration: Borlaff, A. S., Gómez-Alvarez, P., Altieri, B., et al. 2022, A&A, 657, A92

Euclid Collaboration: Jelic-Cizmek, G., Sorrenti, F., Lepori, F., et al. 2023, arXiv:2311.03168

Euclid Collaboration: Scaramella, R., Amiaux, J., Mellier, Y., et al. 2022, A&A, 662, A112

Euclid Collaboration: Schirmer, M., Jahnke, K., Seidel, G., et al. 2022, A&A, 662, A92

Euclid Collaboration: Tanidis, K., Cardone, V. F., Martinelli, M., et al. 2023, arXiv e-prints, arXiv:2309.00052

Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, ApJ, 426, 23

Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J. 1994, MNRAS, 266, 50–64

Fisher, K. B., Scharf, C. A., & Lahav, O. 1994, MNRAS, 266, 219

Fixsen, D. J., Cheng, E. S., Gales, J. M., et al. 1996, ApJ, 473, 576

Foglieni, M., Pantiri, M., Di Dio, E., & Castorina, E. 2023, Phys. Rev. Lett., 131, 111201

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Frigo, M. & Johnson, S. 2005, Proceedings of the IEEE, 93, 216

Gabrielli, A. 2004, Phys. Rev. E, 70, 066131

Gibelyou, C. & Huterer, D. 2012, MNRAS, 427, 1994

Glanville, A., Howlett, C., & Davis, T. M. 2021, MNRAS, 503, 3510

Goodman, J. & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759

Gott III, J. R., Jurić, M., Schlegel, D., et al. 2005, ApJ, 624, 463

Green, S. R. & Wald, R. M. 2012, Phys. Rev. D, 85, 063512

Guandalin, C., Piat, J., Clarkson, C., & Maartens, R. 2023, ApJ, 953, 144

Hahn, O. & Abel, T. 2011, MNRAS, 415, 2101

Hall, A. & Bonvin, C. 2017, Phys. Rev. D, 95

Hamana, T., Kayo, I., Yoshida, N., Suto, Y., & Jing, Y. P. 2003, MNRAS, 343, 1312

Hamilton, A. J. S. 1992, ApJ, 385, L5

Hamilton, A. J. S. 1998, in Astrophysics and Space Science Library, Vol. 231, The Evolving Universe, ed. D. Hamilton, 185

Hamilton, A. J. S. 2000, MNRAS, 312, 257

Hamilton, A. J. S. & Culhane, M. 1996, MNRAS, 278, 73

Hartlap, J., Simon, P., & Schneider, P. 2007, A&A, 464, 399

Heavens, A. F. & Taylor, A. N. 1995, MNRAS, 275, 483

Hinshaw, G., Weiland, J. L., Hill, R. S., et al. 2009, ApJS, 180, 225

Hivon, E., Gorski, K. M., Netterfield, C. B., et al. 2002, ApJ, 567, 2

Hobson, M. P., Efstathiou, G. P., & Lasenby, A. N. 2006, General Relativity: An Introduction for Physicists (Cambridge University Press)

Hockney, R. W. & Eastwood, J. W. 1988, Computer simulation using particles (CRC Press)

Howlett, C., Manera, M., & Percival, W. J. 2015, L-PICOLA: A parallel code for fast dark matter simulation

Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168

Hui, L., Gaztañaga, E., & LoVerde, M. 2007, Phys. Rev. D, 76, 103502

Hui, L., Gaztañaga, E., & LoVerde, M. 2008, Phys. Rev. D, 77, 063526

Jackson, J. C. 1972, MNRAS, 156, 1P

Jelic-Cizmek, G., Lepori, F., Bonvin, C., & Durrer, R. 2021, J. Cosmology Astropart. Phys., 04, 055

Jeong, D., Schmidt, F., & Hirata, C. M. 2012, Phys. Rev. D, 85, 023504

Joyce, M. & Marcos, B. 2007, Phys. Rev. D, 75, 063516

Kaiser, N. 1987, MNRAS, 227, 1

Karamanis, M. & Beutler, F. 2021, hankl: A lightweight Python implementation of the FFTLog algorithm for Cosmology

Kaufman, G. M. 1967, Report N. 6710, Center for Operations Research and Econometrics. Catholic University of Louvain. Heverlee, Belgium.

Keihänen, E., Kurki-Suonio, H., Lindholm, V., et al. 2019, Astronomy & Astrophysics, 631, A73

Landy, S. D. & Szalay, A. S. 1993, ApJ, 412, 64

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv:1110.3193

Lesgourgues, J. & Pastor, S. 2006, Physics Reports, 429, 307–379

Lewis, A. & Bridle, S. 2002, Phys. Rev. D, 66, 103511

Loureiro, A., Moraes, B., Abdalla, F. B., et al. 2019, MNRAS, 485, 326

Maartens, R., Clarkson, C., & Chen, S. 2018, J. Cosmology Astropart. Phys., 2018, 013

Maartens, R., Fonseca, J., Camera, S., et al. 2021, J. Cosmology Astropart. Phys., 2021, 009

Maciaszek, T., Ealet, A., Gillard, W., et al. 2022, in Space Telescopes and Instrumentation 2022: Optical, Infrared, and Millimeter Wave, ed. L. E. Coyle, M. D. Perrin, & S. Matsuura, Vol. 12180 (SPIE), 613

Matarrese, S., Mollerach, S., & Bruni, M. 1998, Phys. Rev. D, 58

Matsubara, T. 2000a, ApJ, 535, 1

Matsubara, T. 2000b, ApJ, 537, L77

McDonald, P. & Seljak, U. 2009, J. Cosmology Astropart. Phys., 2009, 007–007

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, The Journal of Chemical Physics, 21, 1087

Moore, A. W. et al. 2001, in MPA / ESO / MPE Joint Astronomy Conference: Mining the Sky, ESO Astrophysics Symposia European Southern Observatory, 71–82

Nadolny, T., Durrer, R., Kunz, M., & Padmanabhan, H. 2021, J. Cosmology Astropart. Phys., 2021, 009

Noorikuhani, M. & Scoccimarro, R. 2023, Phys. Rev. D, 107, 083528

Padmanabhan, N., Schlegel, D. J., Seljak, U., et al. 2007, MNRAS, 378, 852

Pápai, P. & Szapudi, I. 2008, MNRAS, 389, 292

Paviot, R., de la Torre, S., de Mattia, A., et al. 2022, MNRAS, 512, 1341

Peacock, J. A. 1991, MNRAS, 253, 1

Peacock, J. A., Cole, S., Norberg, P., et al. 2001, Nature, 410, 169

Peacock, J. A. & Heavens, A. F. 1985, MNRAS, 217, 805

Peebles, P. J. E. 1973, ApJ, 185, 413

Peebles, P. J. E. 1980, The Large-Scale Structure of the Universe (Princeton University Press)

Peebles, P. J. E. 2020, Principles of Physical Cosmology (Princeton University Press)

Peebles, P. J. E. 2022, Anomalies in Physical Cosmology

Percival, W. J., Burkey, D., Heavens, A., et al. 2004, MNRAS, 353, 1201

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565

Peterson, E. R., Kenworthy, W. D., Scolnic, D., et al. 2022, ApJ, 938, 112

Planck Collaboration: Aghanim, N., Akrami, Y., Ashdown, M., et al. 2020a, A&A, 641, A6

Planck Collaboration: Aghanim, N., Akrami, Y., Arroja, F., et al. 2020b, A&A, 641, A1

Pozzetti, L., Hirata, C. M., Geach, J. E., et al. 2016, A&A, 590, A3

Press, W. H. & Schechter, P. 1974, ApJ, 187, 425

Pryer, D., Smith, R. E., Booth, R., et al. 2022, J. Cosmology Astropart. Phys., 08, 019

Racca, G. D., Laureijs, R., Stagnaro, L., et al. 2016, in Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave, ed. H. A. MacEwen, G. G. Fazio, M. Lystrup, N. Batalha, N. Siegler, & E. C. Tong, Vol. 9904, International Society for Optics and Photonics (SPIE), 99040O

Raccanelli, A., Bertacca, D., Doré, O., & Maartens, R. 2014, J. Cosmology Astropart. Phys., 8, 022

Raccanelli, A., Bertacca, D., Jeong, D., Neyrinck, M. C., & Szalay, A. S. 2018, Physics of the Dark Universe, 19, 109

Raccanelli, A., Bertacca, D., Maartens, R., Clarkson, C., & Doré, O. 2016a, General Relativity and Gravitation, 48

Raccanelli, A., Montanari, F., Bertacca, D., Doré , O., & Durrer, R. 2016b, J. Cosmology Astropart. Phys., 05, 009

Raccanelli, A., Samushia, L., & Percival, W. J. 2010, MNRAS, 409, 1525

Raccanelli, A. & Vlah, Z. 2023, Power spectrum in the cave

Reimberg, P., Bernardeau, F., & Pitrou, C. 2016, J. Cosmology Astropart. Phys., 01, 048

Renk, J., Zumalacarregui, M., & Montanari, F. 2016, J. Cosmology Astropart. Phys., 2016, 040, arXiv:1604.03487 [astro-ph, physics:gr-qc]

Ross, A. J., Percival, W. J., Sánchez, A. G., et al. 2012, MNRAS, 424, 564

Sahni, V. 1995, Physics Reports, 262, 1

Samushia, L., Branchini, E., & Percival, W. J. 2015, MNRAS, 452, 3704

Samushia, L., Percival, W. J., & Raccanelli, A. 2012, MNRAS, 420, 2102

Sargent, W. L. W. & Turner, E. L. 1977, ApJ, 212, L3

Scharf, C., Hoffman, Y., Lahav, O., & Lynden-Bell, D. 1992, MNRAS, 256, 229

Schmidt, F. & Jeong, D. 2012, Phys. Rev. D, 86, 083527

Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses (Springer Berlin, Heidelberg)

Scoccimarro, R. 2015, Phys. Rev. D, 92, 083532

Secrest, N. J., von Hausegger, S., Rameez, M., Mohayaee, R., & Sarkar, S. 2022, ApJ, 937, L31

Secrest, N. J., von Hausegger, S., Rameez, M., et al. 2021, ApJ, 908, L51

Sheth, R. K. & Diaferio, A. 2001, MNRAS, 322, 901

Siewert, T. M., Schmidt-Rubart, M., & Schwarz, D. J. 2021, A&A, 653, A9

Singal, A. K. 2021, Universe, 7, 107

Szalay, A. S., Matsubara, T., & Landy, S. D. 1998, ApJ, 498, L1

Szapudi, I. 2004, ApJ, 614, 51

Tansella, V., Jelic-Cizmek, G., Bonvin, C., & Durrer, R. 2018, J. Cosmology Astropart. Phys., 10, 032

Taruya, A., Nishimichi, T., & Jeong, D. 2018, Phys. Rev. D, 98, 103532

Tassev, S., Zaldarriaga, M., & Eisenstein, D. J. 2013, J. Cosmology Astropart. Phys., 2013, 036–036

Tegmark, M., Hamilton, A. J. S., Strauss, M. A., Vogeley, M. S., & Szalay, A. S. 1998a, ApJ, 499, 555

Tegmark, M., Hamilton, A. J. S., Strauss, M. A., Vogeley, M. S., & Szalay, A. S. 1998b, ApJ, 499, 555

Tiwari, P., Kothari, R., Naskar, A., Nadkarni-Ghosh, S., & Jain, P. 2015, Astroparticle Physics, 61, 1

Villa, E., Matarrese, S., & Maino, D. 2011, J. Cosmology Astropart. Phys., 2011, 024–024

Wandelt, B. D., Hivon, E., & Górski, K. M. 2001, Phys. Rev. D, 64, 083003

Wands, D. & Slosar, A. 2009, Phys. Rev. D, 79, 123507

Weinberg, S. 1972, Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity (New York: John Wiley and Sons)

Wen, R. Y., Gebhardt, H. S. G., Heinrich, C., & Doré, O. 2024, Exact Modeling of Power Spectrum Multipole through Spherical Fourier-Bessel Basis

Werner, K. F. & Porciani, C. 2020, MNRAS, 492, 1614

Wilson, M. J., Peacock, J. A., Taylor, A. N., & de la Torre, S. 2017, MNRAS, 464, 3121

Yadav, J., Bharadwaj, S., Pandey, B., & Seshadri, T. R. 2005, MNRAS, 364, 601–606

Yahya, S., Bull, P., Santos, M. G., et al. 2015, MNRAS, 450, 2251

Yamamoto, K., Nakamichi, M., Kamino, A., Bassett, B. A., & Nishioka, H. 2006, Publications of the Astronomical Society of Japan, 58, 93

Yamamoto, K., Nishioka, H., & Suto, Y. 1999, ApJ, 527, 488

Yamamoto, K., Nishioka, H., & Taruya, A. 2000, arXiv:0012433

Yoo, J. 2010, Phys. Rev. D, 82, 083508

Yoo, J. 2023, J. Cosmology Astropart. Phys., 2023, 054

Yoo, J. & Desjacques, V. 2013, Phys. Rev. D, 88, 023502

Yoo, J., Fitzpatrick, A., & Zaldarriaga, M. 2009, Phys. Rev. D, 80, 083514

Zaroubi, S. & Hoffman, Y. 1996, ApJ, 462, 25

Zel'dovich, Y. B. 1970, A&A, 5, 84

Zhang, P., Liguori, M., Bean, R., & Dodelson, S. 2007, Phys. Rev. Lett., 99, 141302

Zonca, A., Singer, L., Lenz, D., et al. 2019, Journal of Open Source Software, 4, 1298