Detecting Magnification Lensing in Photometric Galaxy Surveys

Tiago Sinde
Master in Astronomy and Astrophysics
Departament of Physics and Astronomy
Faculty of Sciences University of Porto
2025

Detecting Magnification Lensing in Photometric Galaxy Surveys

Tiago Sinde

Dissertation carried out as part of the Master in Astronomy and Astrophysics

Departament of Physics and Astronomy

2025

Supervisor
Dr. José Fonseca, Researcher, CAUP & IA

Centro de Astrofísica da Universidade do Porto

ia instituto de astrofísica e ciências do espaço

“When we are no longer able to change a situation, we are challenged to change ourselves.”

Man’s Search for Meaning, from Viktor E. Frankl

Acknowledgements

First of all, I have to acknowledge the hugely important role my supervisor, José Fonseca, played in the making of this dissertation. The guidance and support you gave me through this entire process is what allowed this dissertation to exist. Thank you for all the times you held my hand through a problem and made me understand what I was missing. Thank you for all the dedication and enthusiasm that motivated me to believe that I could complete this project.

Secondly, I would like to thank Pedro Fanha, David Alonso and Arthur Loureiro for their small, but nonetheless relevant help in the computational part of this dissertation.

To my family, and specially to my parents, I want to thank the constant encouragement and support to pursue my goals. Even if they seemed unreachable (which, at times, a physics course might seem), you were always there, without a shred of doubt that I could reach the end of this stage of my life.

To my friends, thank you for listening to me complain about whatever was that week's problem. Whether I was studying in DFA (Departamento de Física e Astronomia) or having a break in the Physic's Bar, you were always trying to help, even though you didn't know how. Thank you for, not only allowing me to survive this academic life, but also for turning it into an amazing experience. Thank you for all the good moments at dinners, hangouts, sunsets, nights out, parties and festivals. A special thanks to all the friends with whom I shared laughs and tears while wearing a black cape on my shoulders.

Most importantly, I have to give my biggest thanks to Ana. Without her, even if this dissertation still existed, my sanity wouldn't. Thank you for always listening to my problems or technical difficulties, even though just a few words on them were, and still are, enough to make you yawn. Thank you for all the encouragement you give and for always being there for me.

Resumo

A estrutura em larga escala da matéria observada no Universo contém informações importantes, não só sobre os seus estágios primordiais, como também sobre a sua composição e evolução. Uma abordagem utilizada para compreender o agrupamento da matéria no Universo é a medição das posições angulares das suas galáxias no céu. No entanto, as nossas observações são ligeiramente distorcidas pelas interações gravitacionais que a luz sofre ao interagir com qualquer matéria ao longo do seu percurso. Estas distorções deixam marcas na distribuição estatística das galáxias observadas, fornecendo testes refinados para a teoria da gravidade de Einstein. O objetivo desta dissertação é determinar o impacto das lentes gravitacionais fracas no agrupamento angular observado de galáxias, em levantamentos fotométricos de galáxias.

Neste trabalho revemos como o espaço-tempo afeta o caminho e a energia observada dos fotões de galáxias distantes, conforme descrito pela teoria geral da gravitação de Einstein, e como isso afeta o espetro de potência observado. Para tal, usamos os pacotes de "python", `PyCAMB` [1] e `PyCCL` [2], para calcular dois estimadores do agrupamento da matéria observada: o espetro de potência angular, C_ℓ, e a função de correlação angular, $w(\theta)$, respetivamente. Para além disso, utilizamos a análise da matriz de Fisher, em conjunto com estes estimadores, para prever as incertezas em futuros levantamentos fotométricos de galáxias. Adicionalmente, prevemos também as restrições nas medições dos parâmetros cosmológicos; do enviesamento da magnificação, $s(z)$, e dum parâmetro de teste da Relatividade Geral, β. Por fim, para validar o rigor do quadro teórico, são geradas simulações realistas de catálogos do céu através do pacote `CoLoRe` [3]. Posteriormente, medimos, a partir desses catálogos, os estimadores de agrupamento C_ℓ e $w(\theta)$ com `NaMaster` [4] e `TreeCorr` [5], respetivamente.

No futuro, este trabalho poderá ser utilizado para determinar, com maior precisão, o impacto de lentes gravitacionais fracas no agrupamento de galáxias observado. Para além disso, dado que o método de cálculo destes estimadores é independente de pressupostos sobre a cosmologia (depende apenas da amostra), podemos usá-lo para testar se a teoria gravitacional padrão, a Relatividade Geral, descreve corretamente o nosso Universo, ou se devemos procurar teorias alternativas.

Palavras-chave: física, astrofísica, cosmologia, estrutura em larga escala, lentes gravitacionais fracas, espetro de potência angular, função de correlação angular

Abstract

The observed Large-Scale Structure (LSS) of matter in the Universe encodes important information about, not only its primordial stages, but also its composition and evolution. An approach used to understand the clustering of matter in the Universe is to measure the angular positions of its galaxies, in the sky. However, our observations are slightly distorted by the gravitational interactions between the source's light and any matter along its path. These distortions leave imprints on the observed statistical distribution of galaxies, providing exquisite new tests for Einstein's theory of gravity. The aim of this dissertation is to determine the impact of weak gravitational lensing on the observed angular clustering of galaxies, in photometric galaxy surveys.

In this work we review how spacetime affects the path and observed energy of photons arriving from distant galaxies, as described by Einstein's general theory of gravitation, and how it affects the observed power spectrum. For this, we will use the python packages, `PyCAMB` [1] and `PyCCL` [2], to calculate two estimators of the observed clustering of matter: the angular power spectrum, C_ℓ, and the angular correlation function, $w(\theta)$, respectively. Furthermore, we utilise Fisher matrix analysis, in conjunction with these two estimators, to forecast uncertainties for future photometric galaxy surveys. We also predict the constraints on the measurements of certain cosmological parameters; magnification bias, $s(z)$, and a General Relativity probe parameter, β. In order to validate the robustness of the theoretical framework, realistic simulations of sky catalogues are generated with the `CoLoRe` package [3]. Afterwards we measure, from those catalogues, the clustering estimators, C_ℓ and $w(\theta)$, with `NaMaster` [4] and `TreeCorr` [5], respectively.

In the future, this work can be used to more accurately measure the impact of weak gravitational lensing in the observed clustering of galaxies. Moreover, given that the method of calculation of these estimators is independent of cosmological assumptions (only dependent on the sample), we can use it to test whether standard gravitation theory, General Relativity, describes our Universe correctly or if we should look beyond it.

Keywords: physics, astrophysics, cosmology, large-scale-structure, weak gravitational lensing, angular power spectrum, angular correlation function.

Table of Contents

Acknowledgements ii

Resumo iii

Abstract iv

Table of Contents v

List of Tables viii

List of Figures ix

List of Abbreviation xii

1 Introduction 1

2 The Effect of Convergence in Cosmology 3

2.1 Basic principles of Cosmology . . . 3

2.1.1 Distances in an expanding Universe . . . 3

2.1.2 Spacetime's structure . . . 4

2.1.3 Hubble's Law and Hubble Parameter . . . 5

2.1.4 Redshift . . . 5

2.2 Weak Gravitational Lensing . . . 7

2.2.1 Deflection due to a Point Mass . . . 8

2.2.2 Deflection due to a Thin Lens . . . 8

2.2.3 The Lensing potential . . . 9

2.2.4 Convergence, Shear and Magnification . . . 10

2.3 Cosmological Perturbations . . . 13

2.3.1 Linear Metric Perturbations . . . 13

2.3.2 Testing GR through metric perturbations . . . 14

2.4 Lensing effect in measurements of Large-Scale Structure . . . 15

2.4.1 Large-Scale Structure 15
2.4.2 Measuring LSS through Galaxy Number Counts 16
2.4.3 Weak Gravitational Lensing in Cosmology 20

3 Estimators of Large-Scale Structure 23
3.1 Statistics of Large-Scale Structure 23
3.1.1 Angular Power Spectrum 24
3.1.2 Two-point Angular Correlation Function 27
3.2 Computing LSS estimators 29
3.2.1 Survey and Cosmology Specifications 29
3.2.2 Software Packages Used 31
3.2.3 Results and their Interpretation 32
3.3 Understanding the impact of Lensing with LSS estimators 33

4 Performance of LSS estimators 37
4.1 Fisher Information Matrix 37
4.1.1 Fisher Matrix of both estimators 38
4.1.2 Covariance Matrix of the Angular Power Spectrum 39
4.1.3 Covariance Matrix of the Angular Correlation Function 40
4.1.4 Parameters to constrain 42
4.1.5 Derivatives of LSS estimators 43
4.2 Comparing the performance of both estimators 44
4.2.1 Computational Procedure for Fisher Matrices 44
4.2.2 Survey Configurations and Parameter Sets 45
4.2.3 Visualising results obtained 47
4.3 Discussion 54

5 Estimating the effect of Weak Lensing Magnification from Simulations 58
5.1 Generate Galaxy Catalogue Simulations 58
5.2 Create Observed Galaxy Catalogue 61
5.3 Measuring Simulated LSS with Clustering Estimators 65
5.3.1 Measuring the Angular Power Spectrum 65
5.3.2 Measure Angular Correlation function 68
5.4 Discussion 70
5.4.1 Challenges and Limitations encountered 70
5.4.2 Interpretation of the Clustering Measurements 70
5.4.3 Conclusions on Measuring Magnification Lensing 71

6 Conclusions **72**

References **74**

A Theoretical error bars **83**

A.1 Error bars of the Angular Power Spectrum 83

A.2 Error bars of the Angular Correlation Function 83

B Analytic derivatives **85**

B.1 Analytic derivative of $\partial C_\ell/\partial\beta$ 85

B.2 Analytic derivative of $\partial C_\ell/\partial A_s$ 87

List of Tables

4.1 Fiducial values of the Cosmological parameters relevant to our analysis. . . 42
4.2 Survey configurations used to test the performance of the C_ℓ and $w(\theta)$ estimators. Valid for both parameter sets. 47

List of Figures

2.1 Sketch of a typical gravitational lens system. Adapted from Bartelmann and Schneider [10]. 9
2.2 Effect of the components of convergence, κ, and shear, γ, on a circular source represented by the purple circle. Here, the κ represents the change in size of the observed source, while the distortions in shape of the observed source are divided into the real, Re$(\gamma) = \gamma_1$, and the imaginary, Im$(\gamma) = \gamma_2$, components of γ (see eq.(2.26)). 12
2.3 Observed Galaxy Distribution from SDSS. A slice of the observed Universe, revealing the filamentary LSS. The statistical properties of this distribution are directly compared to simulations to measure cosmological parameters (Figure Credit: M. Blanton and SDSS) [18]. 16
2.4 Simulated Growth of Structure. As time increases from left ot right, the cosmological simulation shows the gravitational collapse of initial randomly distributed matter into a filamentary structure. (Figure Credit: Andrey Kravtsov, Anatoly Klypin, National Center for Supercomputer Applications (NCSA) [17]). 16

3.1 Representation of multiple components of spherical harmonics, from the monopole $\ell = 0$ to the multipole$\ell = 3$ [27]. 25
3.2 Two sky simulations: (a) is a uniform 10° grid of 25 low-surface-brightness clusters on a random background; and (b) is a quadrupole background with 25 randomly placed clusters. Measured $w(\theta)$ appear below each, (c) and (d), respectively. Adapted from Practical Statistics for Astronomers [25]. . . 28
3.3 DES-like galaxy redshift distribution (black dashed line) and the five Gaussian photometric window functions $W(z)$ used in this analysis. The window functions were not normalised for visualisation purposes 30
3.4 (Left) The angular power spectrum, C_ℓ, for the DES-like sample. The horizontal dashed line represents the expected shot noise level P_N. (Right) The corresponding angular correlation function, $w(\theta)$. This figure shows the contributions from the M+RSD+L signal without error bars. 32
3.5 Autocorrelation of the redshift bin centred at $z_c = 0.585$. The y-axis for $w(\theta)$ is scaled as $\theta w(\theta) \times 10^4$ to better distinguish the contributions of the different components. Bellow each plot we also show their Signal to Noise ratio. As expected, the local matter density term dominates. 34
3.6 Cross-correlation of redshift bins centred at $z_i = 0.585$ and $z_j = 1.175$. The y-axis for $w(\theta)$ is scaled as $\theta w(\theta) \times 10^4$. Bellow each plot we also show their Signal to Noise ratio. As predicted, the non-local lensing term becomes the dominant contribution, illustrating how tomographic cross-correlations isolate the magnifying effect of foreground matter on background sources. . 35

4.1 Corner plot showing the Fisher forecast constraints (error ellipses) for the cosmological parameters in the baseline model, calculated along the $s(z)$ parameters. 48
4.2 Corner plot showing the Fisher forecast constraints for the parameters $s(z_1)$, $s(z_2)$, $s(z_3)$, $s(z_4)$ in the baseline model, calculated along the cosmological parameters. 49
4.3 Relative uncertainties ($\sigma(\lambda_\mu)/\lambda_\mu \times 100\%$) for all parameters in Set $s(z)$ for the Baseline model. 49
4.4 Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set $s(z)$ when moving from the Baseline model to Model 1 with $f_{sky} = 1$, $\theta_{\text{max}} = 18^\circ$, and $\ell_{\text{min}} = 10$. 50
4.5 Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set $s(z)$ when moving from the Baseline model to Model 2 with $f_{sky} = 0.33$, $\theta_{\text{max}} = 18^\circ$, and $\ell_{\text{min}} = 10$. 50
4.6 Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set $s(z)$ when moving from the Baseline model to Model 3 with $f_{sky} = 1$, $\theta_{\text{max}} = 18^\circ$, and $\ell_{\text{min}} = 20$. 50
4.7 Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set $s(z)$ when moving from the Baseline model to Model 4 with $f_{sky} = 1$, $\theta_{\text{max}} = 18^\circ$, and $\ell_{\text{min}} = 30$. 51
4.8 Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set $s(z)$ when moving from the Baseline model to Model 5 with $f_{sky} = 0.33$, $\theta_{\text{max}} = 18^\circ$, and $\ell_{\text{min}} = 30$. 51
4.9 Corner plot showing the Fisher forecast constraints (error ellipses) for parameter β and the cosmological parameters, in the baseline model. 52
4.10 Relative uncertainties ($\sigma(\lambda_\mu)/\lambda_\mu \times 100\%$) for all parameters in Set β for the Baseline model. 52
4.11 Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set β when moving from the Baseline model to Model 1 with $f_{sky} = 1$, $\theta_{\text{max}} = 18^\circ$, and $\ell_{\text{min}} = 10$. 53
4.12 Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set β when moving from the Baseline model to Model 2 with $f_{sky} = 0.33$, $\theta_{\text{max}} = 18^\circ$, and $\ell_{\text{min}} = 10$. 53
4.13 Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set β when moving from the Baseline model to Model 3 with $f_{sky} = 1$, $\theta_{\text{max}} = 18^\circ$, and $\ell_{\text{min}} = 20$. 53
4.14 Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set β when moving from the Baseline model to Model 4 with $f_{sky} = 1$, $\theta_{\text{max}} = 18^\circ$, and $\ell_{\text{min}} = 30$. 54
4.15 Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set β when moving from the Baseline model to Model 5 with $f_{sky} = 0.33$, $\theta_{\text{max}} = 18^\circ$, and $\ell_{\text{min}} = 30$. 54

5.1 Distribution of galaxies, per redshift, that was generated through the Luminosity function in equation (5.1) 60
5.2 Comparison between the true cosmological redshift (z_{True}) and the observed photometric redshift (z_{phot}) for galaxies in the simulation. Each point represents a galaxy. The photometric scatter and RSD effects cause z_{phot} to cover a broader range (0.01 to 2.5) than z_{True} (0.05 to 2.0). 62

5.3 The Schechter Luminosity Function, which was evaluated at the medium redshift of the catalogue, $z = 0.97$, was used as the probability distribution for assigning "True" magnitudes (m_{True}) to galaxies. $M_{\text{lim, cut-off}}$ is obtained through the use of eq.(5.3) with $m_{\text{lim, cut-off}}$ and represents the magnitude limit of a survey. 63
5.4 Distribution of observed magnitudes (m_{obs}) for galaxies in the catalogue after applying the magnitude cut-off at $m_{\text{lim,cut-off}} = 24.0$. 64
5.5 Redshift distribution of galaxies after selection into photometric redshift bins of width. The bins are defined by slicing the observed z_{phot} values (Right panel). The resultant "True" z-bins have gaussian distributions due to the photometric errors (Left panel). 65
5.6 `HEALPix` maps of the galaxy over-density, $\delta(\hat{\mathbf{n}})$, for two different redshift bins, in which each pixel's colour represents the excess density of galaxies (see eq.(5.12)). The maps include the effects of photometric errors, RSD, and lensing as applied to the observed catalogue. 66
5.7 Measured angular power spectra from the simulated galaxy catalogue (green points) compared to the theoretical prediction with lensing (orange line) and without it (red line). The measurement is performed using the `NaMaster` pseudo-C_ℓ estimator. 68
5.8 Measured angular correlation function, $w(\theta)$, from the simulated catalogue (points) using the `TreeCorr` implementation of the Landy-Szalay estimator. The solid line shows the theoretical prediction. Each point is the average of 10 measurements with different random point realisations. 69

6.1 Strategy to constrain gravitational potentials Φ and Ψ. 74

List of Abbreviations

LSS	Large Scale Structure
FLRW	Friedmann-Lemaître-Robertson-Walker
GR	General Relativity
SDSS	Sloan Digital Sky Survey
NCSA	National Center for Supercomputer Applications
RSD	Redshift-Space Distortions
PDF	Probability Distribution Function
CMB	Cosmic Microwave Background
DES	Dark Energy Survey
ΛCDM	Lambda-Cold Dark Matter
RA	Right Ascension
Dec	Declination
RAM	Random-Access Memory
MCMC	Markov Chain Monte Carlo

1. Introduction

The cosmological principle states that the Universe is statistically isotropic and homogeneous when considered at *large enough scales* [6]. An estimate for an upper limit for which this principle is applicable is $260\,h^{-1}$ Mpc, for the flat ΛCDM model [7].

However, this is just a simplification of reality since galaxies, and clusters of galaxies, exist. Besides, even on large scales, the distribution of matter is not perfectly uniform [8]. The Large-Scale Structure (LSS) of the Universe represents local departure from homogeneity. Thus, these deviations contain important information about the Universe's composition and evolution. As a means to study the effect of these perturbations on the energy content of the Universe, we examine the distribution of its matter. While ordinary baryonic matter is directly observable, the dominant component, dark matter, is not. Fortunately, both baryonic and dark matter are concentrated in galaxies, or in their vicinity [9]. As a consequence of that, the distribution of galaxies serves as the primary tracer for the underlying total matter distribution.

The logical next step is to measure the position of galaxies and extract information from it, although a challenge emerges in this procedure. We observe light that has travelled across the Universe, and, if this light passes through the gravitational fields of matter, its path might be bent due to the distortion of spacetime. This effect, known as weak gravitational lensing, causes the observed images of galaxies to be distorted, in size and/or shape, when compared to their true forms [10]. To overcome this hurdle, cosmologists have to model the impact of the lensing effect in the observed cluster of galaxies. Out of those tools, the two most prevalent estimators, for two-dimensional projections of matter on the celestial sphere, are the Angular Power Spectrum, C_ℓ, and the Angular Correlation Function, $w(\theta)$.

This dissertation aims to study the impact of weak magnification lensing on photometric galaxy surveys and analyse in which circumstances it can be detected. To achieve this, we will compare the performance of the two aforementioned estimators when measuring galaxy clustering in a photometric survey, and attempt to detect this effect in simulated

data.

Following this introduction, we will advance to Chapter 2, which presents the principles of cosmology and weak gravitational lensing. It then discusses cosmological perturbations and concludes by examining the impact of weak lensing magnification on LSS measurements. Chapter 3 reviews the statistics of LSS and provides a detailed explanation on the C_ℓ and $w(\theta)$ estimators. It also describes their computation process and utilises them to study the impact of lensing convergence. Next, Chapter 4 is where we will compare the performance of the LSS estimators through the use of Fisher Information Matrices. We then derive the necessary equations, compute the matrices for specific survey configurations, and present a comparative analysis of the results. Afterwards, Chapter 5 involves generating a realistic simulated galaxy catalogue. We introduce observational biases to mimic a photometric survey, measure the galaxy clustering using both estimators, and discuss the results obtained. To finish, Chapter 6 will summarise our findings, discuss their implications, and suggest directions for future work.

2. The Effect of Convergence in Cosmology

Before delving into the details of my work, it is essential to understand the fundamental concepts on which it is based upon. This chapter will review the essentials about distances in Cosmology (section 2.1) and Weak Gravitational Lensing (section 2.2), with a particular focus on magnification lensing and convergence. Afterwards, we will use cosmological perturbations to show the usefulness of this work (section 2.3). Finally, we describe how the weak magnification lensing effect impacts our measurements of LSS (section 2.4).

2.1. Basic principles of Cosmology

2.1.1. Distances in an expanding Universe

In Cosmology, the Universe, when considered at large enough scales ($260\,h^{-1}$ Mpc, for flat ΛCDM [7]), is often approximated to a continuous fluid, in such a way that each element of this fluid can be assigned with a spatial coordinate, $\vec{r}$. Accordingly, a point in spacetime corresponds to a portion of fluid element passing through that point at **proper time** t, measured by a clock which is moving with the fluid (a comoving clock).

Furthermore, as time increases in an expanding universe, so does the "real" distance between any two points. Hence, if we consider that $\vec{d}$ represents this distance, the **proper distance** between these two points, then [11]

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

where $a(t)$ is the **cosmic scale factor** and $\vec{r}$ is the **comoving distance**. The comoving distance factors out the expansion of the universe, giving a distance that does not change with time. Meanwhile, the scale factor describes the size/scale of the Universe according to it's expansion through time. Additionally, this factor is defined as unity for the present time, $a(t_0) \equiv 1$, which implies that $\vec{d}(t_0) = \vec{r}$.

2.1.2. Spacetime's structure

The metric is a tensor, $g_{\mu\nu}$, that allows us to define distances and angles on the 4-dimensional fabric of the Universe. The invariant line element, ds^2, takes advantage of the metric tensor to turn coordinate distances into physical distances. Therefore, by using Einstein notation with the metric signature convention $(-,+,+,+)$, the spacetime distance between two points can be written as [11]

$$\begin{aligned} ds^2 &= -c^2dt^2 + a^2(t)d\vec{r}^2\,, \quad \text{with} \quad \vec{r} = r^i\,, \\ &= g_{\mu\nu}dr^\mu dr^\nu\,, \qquad\qquad \text{with} \quad r^\mu = (t,\vec{r})\,. \end{aligned} \tag{2.2}$$

There exist a multitude of metrics, with the most well-known being the Euclidean metric and the flat Minkowski metric. The metric one chooses to employ in their studies is related to the conditions/assumptions that are considered about the structure of the Universe.

From now on we will assume that the Universe is **homogeneous** (has constant density), **isotropic** (equal in all directions), **expanding** and has **curvature**. A Universe under these conditions can be described by the Friedmann-Lemaître-Robertson-Walker (FLRW) metric [11],

$$ds^2 = -c^2dt^2 + a^2(t)\left[\frac{dr^2}{1-Kr^2/R_0^2} + r^2d\Omega^2\right]\,. \tag{2.3}$$

Here, $d\Omega^2 \equiv d\theta^2 + \sin^2\theta d\phi^2$ is defined as the metric of the unit two-sphere, and K is the curvature, for which $K>0$, $K=0$ and $K<0$ represent positive, flat and negative curvature, respectively.

In this dissertation we will always consider that $K=0$. So, for a flat Universe, the metric in equation (2.3) simplifies to [11]

$$\boxed{ds^2 = -c^2dt^2 + a^2(t)\left[dr^2 + r^2d\Omega^2\right]}\,. \tag{2.4}$$

In an additional note, it is often useful to introduce the **conformal time**, τ, which is related to the proper time by

$$dt = a(t)d\tau\,. \tag{2.5}$$

Then, applying this relation to equation (2.4) gives [11]

$$ds^2 = a^2(\tau)\left[-c^2d\tau^2 + dr^2 + r^2d\Omega^2\right]\,. \tag{2.6}$$

2.1.3. Hubble's Law and Hubble Parameter

Hubble's Law refers to the observational discovery that galaxies are receding from us with a proper velocity v that is proportional to their proper distance d:

$$\vec{v}(t) = H_0 \, \vec{d}(t) \, , \tag{2.7}$$

where the constant of proportionality, H_0, is known as the **Hubble constant**. However, it has later been found that the linear relation in equation (2.7) is only a reliable approximation when considering small redshifts ($z \ll 1$).

In order to describe the expansion of the Universe, at any point in time, we can generalise the Hubble constant into the **Hubble parameter**, $H(t)$, with the help of the scale factor. Then, the Hubble parameter is defined as the relative expansion rate of the scale factor,

$$H(t) \equiv \frac{1}{a(t)} \frac{\mathrm{d}a}{\mathrm{d}t} \, , \tag{2.8}$$

where $\frac{\mathrm{d}a}{\mathrm{d}t}$ is the proper time derivative of the scale factor. Physically, $H(t)$ quantifies the **fractional rate of expansion of the Universe** at a given proper time t. Today, its value is the Hubble constant, $H(t_0) = H_0$.

2.1.4. Redshift

The expanding Universe is not the only obstacle researchers encounter when measuring cosmological distances. Due to its large size we cannot use the parallax method to measure the distance to any body beyond the vicinity of our own galaxy [12], therefore another method is used for very large distances, redshift measurements. As photons travel astronomical distances they lose energy, due to the expansion of the Universe. For visible light, this shifts the spectrum toward the red. Therefore, this phenomenon is called cosmological redshift, z, and is quantified by the relative change in wavelength [11],

$$z \equiv \frac{\lambda_{obs} - \lambda_{emit}}{\lambda_{emit}} \, , \tag{2.9}$$

where λ_{emit} is the wavelength of light when it was emitted and λ_{obs} is the wavelength observed.

Furthermore, there's a link between redshift and the scale factor. To understand it let's consider that a short signal, of conformal duration $\delta\tau$, is emitted at conformal time τ_{emit} and observed at τ_{obs}. The conformal duration of this signal is measured equally both by the source and the observer, however, the physical time intervals are different at the points

of emission and detection. This is,

$$\delta t_{\text{emit}} = a(t_{\text{emit}})\delta\tau \qquad \text{and} \qquad \delta t_{\text{obs}} = a(t_{\text{obs}})\delta\tau\,.$$

If δt_{emit} is the period of the light wave, then the light is emitted with wavelength $\lambda_{\text{emit}} = c\delta t_{\text{emit}}$, but is observed with wavelength $\lambda_{\text{obs}} = c\delta t_{\text{obs}}$. Therefore, the ratio of wavelengths can be given by [11]

$$\frac{\lambda_{\text{obs}}}{\lambda_{\text{emit}}} = \frac{a(t_{\text{obs}})}{a(t_{\text{emit}})}\,. \tag{2.10}$$

Recalling that, by convention, we define $a(t_{\text{obs}}) \equiv 1$ (see section 2.1.1), then

$$\frac{\lambda_{\text{obs}}}{\lambda_{\text{emit}}} = \frac{1}{a(t_{\text{emit}})}\,. \tag{2.11}$$

This leads us to the relation between redshift and the expansion of the Universe [11],

$$1 + z = \frac{1}{a}\,. \tag{2.12}$$

From this relation we infer that, today, the redshift is $z = 0$, while higher redshifts correspond to earlier times, when the Universe was smaller ($a < 1$).

In order to calculate a photon's redshift, it is necessary to know its wavelength both when it was emitted and when it was observed. Additionally, since redshift is a direct measure of the scale factor, it can be related to the distance to the source. Thus, the comoving distance, r, to an object at redshift z can be calculated by integrating over the expansion history, from the time the light was emitted to when it is observed. This results in the following expression [11]:

$$\begin{aligned} r(z) &= \int_{t_{\text{emit}}}^{t_{\text{obs}}} \frac{c}{a(t')}\,\mathrm{d}t' \\ &= \int_0^z \frac{c}{H(z')}\,\mathrm{d}z' \quad \left(\text{using the relation } \mathrm{d}z = -(1+z)H(z)\,\mathrm{d}t\right)\,. \end{aligned}$$

From now on we will consider natural units, $c = 1$, which simplifies the previous equation into

$$r(z) = \int_0^z \frac{1}{H(z')}\,\mathrm{d}z'\,. \tag{2.13}$$

This integral shows that, to calculate distances from redshift, we need to know: not only the metric that describes spacetime (see eq.(2.4)), but also how the expansion rate H has evolved through time. Moreover, in the previous equation, the Hubble parameter, $H(z)$, was introduced as a function of redshift. Due to its evolution depending on the energy

content of the Universe, we can write it, for a standard cosmological model, as [11]

$$H(z) = H_0\sqrt{\Omega_{\gamma,0}(1+z)^4 + \Omega_{m,0}(1+z)^3 + \Omega_{K,0}(1+z)^2 + \Omega_{\Lambda,0}}\ , \tag{2.14}$$

where:

H_0: The Hubble constant (the present-day expansion rate);

$\Omega_{\gamma,0}$: The current density parameter for **radiation** (relativistic particles, photons). Its density dilutes as $(1+z)^4$ due to volume expansion and energy loss (redshift) of individual photons;

$\Omega_{m,0}$: The current density parameter for **matter** (both baryonic and dark matter). Its density dilutes as $(1+z)^3$ due to volume expansion alone.

$\Omega_{\Lambda,0}$: The current density parameter for **dark energy**, often described by a cosmological constant. Its energy density is constant and does not dilute with expansion, $(1+z)^0$;

$\Omega_{K,0}$: The **curvature density parameter**. It describes the geometry of the universe (flat, open, or closed) and evolves as $(1+z)^2$.

Furthermore, the sum of these density parameters, by definition (see equation 2.14), equal to one: $\Omega_{\gamma,0} + \Omega_{m,0} + \Omega_{\Lambda,0} + \Omega_{K,0} = 1$.

So, by measuring the redshift of a source and knowing the values of the cosmological parameters ($H_0, \Omega_{\gamma,0}, \Omega_{m,0}, \ldots$), we can compute the integral of $r(z)$ and determine the distance to that source.

2.2. Weak Gravitational Lensing

Gravitational lensing is a phenomenon in which a body (or bodies) acts on a passing light ray as a lens, bending its trajectory due to gravity. The concept was theoretically predicted before its observational measurement and even before the theory of General Relativity (GR) was proposed[13].

In 1801, Soldner applied Newton's corpuscular theory of light, to solve a ballistic scattering problem. Interpreting light through the perspective of many, high velocity, but very small masses that feel other bodies' gravity, one can say that light particles suffer a small deviation on their motion, creating a scattering angle. Ignoring the motion of the scatterer body (bigger mass) and considering a small deflection, by Newtonian arguments,

the angle of deflection is [13]

$$\alpha_N = \frac{2GM}{dc^2}\,. \tag{2.15}$$

In this equation, d represents the distance of closest approach (or the impact parameter) between scatterer and scattered bodies, G is the gravitational constant, M is the scatterer mass and c is the speed of light. In spite of this proposal being a great attempt of describing the problem, Newton's theory can't be applied to light since photons have no mass. Therefore, we need a better theory, such as General Relativity. In GR it's considered that light travels in spacetime through the shortest path. Due to the fact that mass distorts spacetime around it, then the shortest possible path that light can take will also be distorted.

2.2.1. Deflection due to a Point Mass

Let us start by seeing how GR tackles the deflection of light by a point mass, M, which is acting as a gravitational lens. For this we have to consider light rays that propagate outside the strong gravitational field close to the event horizon (otherwise the light ray would fall into the lens itself). Specifically, their impact parameter, ξ, has to be much larger than the Schwarzschild radius of the lens, $\xi \gg R_S \equiv 2GMc^{-2}$. In these conditions, GR predicts that the deflection angle, $\hat{\alpha}$, is [10]

$$\hat{\alpha} = \frac{4GM}{c^2\xi}\,. \tag{2.16}$$

This angle is twice the one obtained in equation (2.15) with Newtonian gravity. Additionally, according to the condition $\xi \gg R_S$, the deflection angle has to be small, $\hat{\alpha} \ll 1$.

2.2.2. Deflection due to a Thin Lens

In the thins lens approximation we consider that the mass of the lens, instead of being a point mass, is now distributed throughout a plane perpendicular to the incoming light ray. This indicates that the mass density projected onto a plane is just the surface mass density and, therefore, can be defined as [10]

$$\Sigma\left(\vec{\xi}\right) \equiv \int \rho\left(\xi_1, \xi_2, r_3\right)\, dr_3\,. \tag{2.17}$$

Then, the deflection angle can be written as [10]

$$\hat{\vec{\alpha}}\left(\vec{\xi}\right) = \frac{4G}{c^2}\int \Sigma\left(\vec{\xi}'\right)\frac{\vec{\xi}-\vec{\xi}'}{\left|\vec{\xi}-\vec{\xi}'\right|^2}\, d^2\xi'\,. \tag{2.18}$$

This expression is valid as long as the light ray's deviation from a straight, undeflected, path is small when compared to the scale over which the mass distribution changes significantly.

In order to better understand the geometry of this deflection angle we can direct our attention to Figure 2.1. In it, we can observe a typical situation in gravitational lensing, where a mass concentration on the lens plane at redshift z_d (angular diameter distance D_d) deflects light rays from a source plane at redshift z_s (angular diameter distance D_s). Under the assumptions that there are no significant deflectors near the line-of-sight and that the physical extent of the lens mass, along the line-of-sight, is negligible compared to both D_d and the lens-source distance D_{ds}, we may invoke the *thin-lens approximation*. This powerful simplification allows us to model the smoothly curved trajectory of a light ray as two straight line segments connected by an instantaneous change in direction, or a "kink", at the lens plane. The magnitude and orientation of this kink are characterised by the deflection angle $\hat{\alpha}$, which is a function of the surface mass density Σ of the deflector and the impact vector $\vec{\xi}$ of the incident light ray.

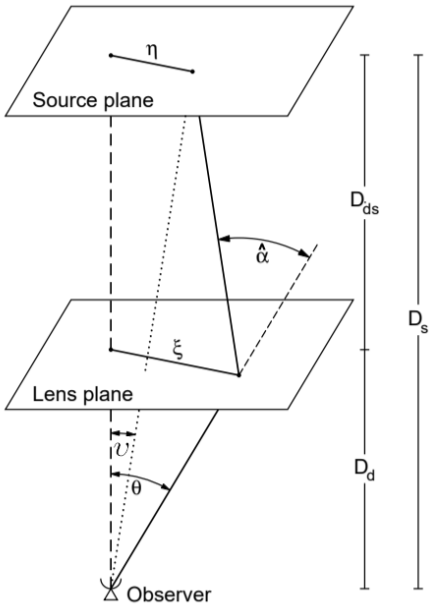


FIGURE 2.1: Sketch of a typical gravitational lens system. Adapted from Bartelmann and Schneider [10].

2.2.3. The Lensing potential

The lensing potential (also called deflection potential) encapsulates the information about the mass distribution of the lens and the geometry of the lens-source system. It is

also the fundamental quantity from which key observables like deflection angles, image distortions, and magnification lensing are derived [10].

We begin by defining the dimensionless surface mass density, also called convergence, as [10]

$$\kappa\left(\vec{\theta}\right) = \frac{\Sigma(D_d\vec{\theta})}{\Sigma_{cr}}\,, \qquad \text{with} \quad \Sigma_{cr} = \frac{c^2}{4\pi G}\frac{D_s}{D_d D_{ds}}\,. \tag{2.19}$$

Here, Σ_{cr} is the critical surface mass density (which depends on the redshifts of both the source and lens). Hence, Σ_{cr} is a characteristic value for the surface mass density which distinguishes between "weak" and "strong" lenses. Lenses with $\kappa \gtrsim 1$ (i.e., $\Sigma \gtrsim \Sigma_{\rm cr}$) are capable of producing strong lensing effects, such as multiple images or giant arcs, while weak lensing deals with the regime of $\kappa \ll 1$ [10].

The scaled deflection angle, $\vec{\alpha}(\vec{\theta})$, can be written in terms of κ, by integrating over the convergence field as [10]

$$\vec{\alpha}(\vec{\theta}) = \frac{1}{\pi}\iint_{\mathbb{R}^2} \kappa(\vec{\theta'})\frac{\vec{\theta}-\vec{\theta'}}{|\vec{\theta}-\vec{\theta'}|^2}\, d^2\theta'\,. \tag{2.20}$$

Since the deflection field $\vec{\alpha}(\vec{\theta})$ is a vector gradient field, it can be expressed as the gradient of a scalar potential,

$$\vec{\alpha}(\vec{\theta}) = \nabla\psi(\vec{\theta}). \tag{2.21}$$

Substituting this into the integral form of $\vec{\alpha}$ reveals the definition of the lensing potential [10]:

$$\psi(\vec{\theta}) = \frac{1}{\pi}\iint_{\mathbb{R}^2} d^2\theta'\ \kappa(\vec{\theta'})\ln|\vec{\theta}-\vec{\theta'}|. \tag{2.22}$$

Moreover, this potential is the two-dimensional analogue of the Newtonian gravitational potential, meaning that it satisfies the Poisson equation [10]

$$\nabla^2_\Omega\psi(\vec{\theta}) = 2\kappa(\vec{\theta})\,, \tag{2.23}$$

in which ∇^2_Ω is the angular Laplacian. This equation is a fundamental field equation of lensing, since it directly links the mass distribution, κ, to the potential, ψ.

2.2.4. Convergence, Shear and Magnification

We have previously seen that gravitational lensing distorts the observed position of a source. However, when we're observing bodies that are more complex than point-sources, this effect can also change the observed size and/or shape of the source.

The unlensed angular position $\vec{v}$ and the lensed angular position $\vec{\theta}$ are shown in Figure 2.1. The connection between the unlensed coordinate systems to the lensed one is given by the following Jacobian [10, 14]:

$$\begin{aligned}\mathcal{J}(\vec{\theta}) &= \frac{\partial \vec{v}}{\partial \vec{\theta}} \\ &= \left(\delta_{ij} - \frac{\partial^2 \psi(\vec{\theta})}{\partial\theta_i \partial\theta_j}\right) \equiv \left(\delta_{ij} - \psi_{,ij}(\vec{\theta})\right) \\ &= \begin{pmatrix} 1-\psi_{,11} & -\psi_{,12} \\ -\psi_{,21} & 1-\psi_{,22} \end{pmatrix} . \end{aligned} \tag{2.24}$$

From here we can introduce two concepts:

1. The **lensing convergence**, $\kappa\left(\vec{\theta}\right)$, which is responsible for the change in the trace part of the Jacobian matrix $\left(\text{tr}(\mathcal{J}) = 2(1-\kappa)\right)$ and is given by [14]

$$\kappa \equiv \frac{1}{2}(\psi_{,11} + \psi_{,22}) = \frac{1}{2}\nabla^2_\Omega \psi \,. \tag{2.25}$$

2. The **complex shear**, γ, which is defined as [14]

$$\gamma \equiv \gamma_1 + i\gamma_2 = |\gamma| e^{2i\phi} \,; \tag{2.26}$$

$$\gamma_1 = \frac{1}{2}(\psi_{,11} - \psi_{,22}) = |\gamma| \cos(2\phi) \,; \tag{2.27}$$

$$\gamma_2 = \frac{1}{2}\left(\psi_{,12} + \psi_{,21}\right) = \psi_{,12} = |\gamma| \sin(2\phi) \,. \tag{2.28}$$

This allow us to rewrite the Jacobian as

$$\mathcal{J}(\vec{\theta}) = (1-\kappa)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - |\gamma| \begin{pmatrix} \cos(2\phi) & \sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{pmatrix} . \tag{2.29}$$

The first term in equation (2.29), dependent on κ, represents an isotropic transformation (rescale source by a constant factor, in all directions) and the second one, related to γ, stretches the image of the source along one specific direction. In other words, while the overall size change is due to convergence, the shape distortion is due to the shear (tidal gravitational field). For a better understanding of this lensing effects see Figure 2.2.

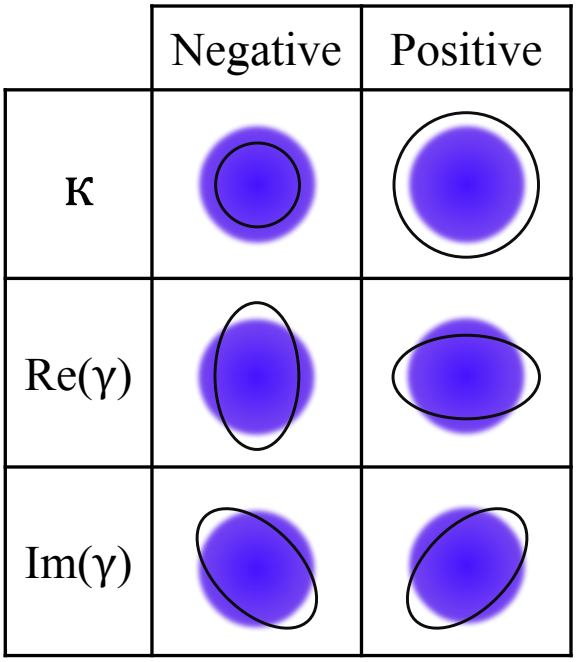


FIGURE 2.2: Effect of the components of convergence, κ, and shear, γ, on a circular source represented by the purple circle. Here, the κ represents the change in size of the observed source, while the distortions in shape of the observed source are divided into the real, Re$(\gamma) = \gamma_1$, and the imaginary, Im$(\gamma) = \gamma_2$, components of γ (see eq.(2.26)).

Moreover, from these two concepts we can define another one, the **magnification lensing**, μ. From Liouville's theorem we know that lensing conserves the surface brightness of a background source [10]. On the other hand, lensing causes focusing of light rays, resulting in an amplification of the image flux through the local solid-angle distortion. Thus, lensing magnification is given by the ratio between the lensed to the unlensed image solid angle as $\mu = \delta\Omega^{Image}/\delta\Omega^{Source} = 1/\det\{\mathcal{J}\}$, with [14]

$$\mu \equiv \frac{1}{det\mathcal{J}} = \frac{1}{(1-\kappa)^2 - |\gamma|^2}\,. \tag{2.30}$$

Furthermore, in the **weak-lensing limit**, we have that $|\gamma| \ll 1$ and $|\kappa| \ll 1$. So, the magnification factor, approximated to first order, becomes

$$\boxed{\mu \simeq 1 + 2\kappa}\,. \tag{2.31}$$

In this dissertation we will work in the weak-lensing regime and through equation (2.31) we see that the shear contributions can be disregarded. Therefore, our analysis will be focused on magnification lensing, μ, **considering only the convergence parameter**, κ.

2.3. Cosmological Perturbations

While the standard cosmological model assumes a homogeneous and isotropic Universe on large scales, this homogeneity is purely statistical. It accommodates small-scale inhomogeneities, such as galaxies and clusters, which are treated as perturbations against a smooth background. Even a structure as vast as the Milky Way represents a mere local perturbation within this framework.

This section examines the nature of these cosmic perturbations, also known as metric perturbations, which manifest as fluctuations in the fabric of spacetime itself.

2.3.1. Linear Metric Perturbations

Let us go beyond the background metric of equation (2.4) and consider that this metric is affected by first-order linear perturbations. Then, these perturbations will generate a new metric:

$$g_{\mu\nu}\left(t,\vec{r}\right) = \bar{g}_{\mu\nu} + \delta g_{\mu\nu}\left(t,\vec{r}\right) \ , \tag{2.32}$$

where $\bar{g}_{\mu\nu}$ is the flat FLRW background metric (see eq.(2.4)). Thus, the perturbed spacetime can be written as [11]

$$ds^2 = a^2(\tau)\left[-(1+2A)d\tau^2 + 2B_i dx^i d\tau + (\delta_{ij}+2E_{ij})dx^i dx^j\right] \ , \tag{2.33}$$

where the scalar A, the vector $\vec{B}$ and the tensor with components E_{ij} depend on space and conformal time. It is known that any vector can be decomposed into its scalar quantity and its unit vector (which describes its direction), such as [11]

$$B_i = \underbrace{\partial_i B}_{\text{scalar}} + \underbrace{\hat{B}_i}_{\text{vector}} \ . \tag{2.34}$$

We can apply the same decomposition to a tensor, so that [11]

$$E_{ij} = \underbrace{C\delta_{ij} + \partial_{\langle i}\partial_{j\rangle}E}_{\text{scalar}} + \underbrace{\partial_{(i}\hat{E}_{j)}}_{\text{vector}} + \underbrace{\hat{E}_{ij}}_{\text{tensor}} \ . \tag{2.35}$$

For the purposes of this dissertation we will only consider the scalar part of both B_i and E_{ij}, see Baumann [11] for more details.

The reparametrisation in equation (2.33) contains more degrees of freedom than the physical degrees of freedom in gravity. In other words, we have more variables with which to define the metric tensor than there are Einstein field equations. This gives us the liberty to choose which Gauge to use, i.e. we have the freedom to choose the coordinates

with which we'll define the perturbed spacetime. An effective way to handle this freedom is to construct **gauge-invariant variables**, which do not transform under a change of coordinates and, thus, represent physical perturbations. The Bardeen variables, Φ and Ψ, are the fundamental gauge-invariant quantities that describe metric perturbations [15]:

$$\Phi \equiv A + \mathcal{H}\left(B - \frac{\mathrm{d}E}{\mathrm{d}\tau}\right) + \frac{\mathrm{d}}{\mathrm{d}\tau}\left(B - \frac{\mathrm{d}E}{\mathrm{d}\tau}\right) ; \tag{2.36}$$

$$\Psi \equiv -C + \frac{1}{3}\nabla^2 E - \mathcal{H}\left(B - \frac{\mathrm{d}E}{\mathrm{d}\tau}\right) . \tag{2.37}$$

These are the "real" spacetime perturbations, since they cannot be removed by any gauge transformation. Note that, when defining the Bardeen potentials, we introduced the co-moving Hubble parameter, $\mathcal{H}$. Similarly to the proper distance, in equation (2.1), this parameter is given by [11]

$$\mathcal{H}(z) \equiv \frac{1}{a(t)}\frac{\mathrm{d}a(t)}{\mathrm{d}\tau} = a(t)H(z) . \tag{2.38}$$

An alternative approach is to fix the gauge and only then include the Bardeen potentials. Different gauges are used for different purposes and, in our situation, the most intuitive gauge is the Newtonian gauge. This is due to the fact that, in this gauge, the Bardeen potentials are analogous to the Newtonian gravitational potential. This gauge fixes $B = E = 0$, allowing us to rewrite A and C as $A \equiv \Psi$ and $C \equiv -\Phi$ [11]. Therefore, equation (2.33) is reduced to

$$ds^2 = a^2(\tau)\left[-(1+2\Psi)d\tau^2 + (1-2\Phi)\delta_{ij}dx^i dx^j\right] . \tag{2.39}$$

This is the **perturbed flat FLRW metric**. In this form, the Bardeen potentials Ψ and Φ represent the **gravitational potential** and the **curvature potential**, respectively.

2.3.2. Testing GR through metric perturbations

The Bardeen potentials can be used to test the validity of GR due to the fact that they are gauge invariant. To do so, we apply the perturbed flat FLRW metric, equation(2.33), to the purely spatial part of both the matter perturbations of the energy-momentum tensor, $T^i_{\ j}$ and the Einstein equation, $G^i_{\ j} = 8\pi G T^i_{\ j}$. From this we obtain the following expression [11]:

$$\partial_{\langle i}\partial_{j\rangle}\left(\Phi - \Psi\right) = 8\pi G a^2 \Pi_{ij} . \tag{2.40}$$

In equation (2.40) the anisotropic stress, Π_{ij}, originates from interactions of matter, photons or neutrinos.

As mentioned in section 2.1.1, when considering large scales, the behaviour of dark matter and baryonic matter can be approximated to the one of perfect fluids[11]. These fluids have negligible anisotropic stress. Photons only start to develop an anisotropic stress component during the matter-dominated era, but their energy density is already non-dominant and will keep decreasing [11]. The only relevant source of anisotropic stress remaining are neutrinos. However, due to their effect being relatively small, we will discard their contribution [11]. Therefore, under these assumptions, $\Pi_{ij} \approx 0$, which implies that equation (2.40) becomes

$$\Phi - \Psi \approx 0 \qquad \text{or} \qquad \frac{\Psi}{\Phi} \approx 1\,. \tag{2.41}$$

The assumptions employed in deriving equation (2.41) are analogous to those utilised in General Relativity. This is due to the fact that, a Universe with no anisotropic stress is also uniformly isotropic. Thus, if one can measure the relation between Φ and Ψ, then it is possible to determine if GR is the correct model for describing spacetime, or if we need to modify it.

2.4. Lensing effect in measurements of Large-Scale Structure

2.4.1. Large-Scale Structure

We already mentioned that the Universe is not homogeneous on all scales. Galaxies are arranged in a vast cosmic web of filaments, clusters, and voids known as Large-Scale Structure. This structure formed from the gravitational amplification of tiny primordial density fluctuations. Analysing the clustering of matter of the LSS provides powerful constraints on the composition of the Universe and its laws of gravity.

As a means to describe the formation and evolution of the LSS of the Universe, a fundamental quantity was defined : the density contrast. This quantity, denoted by δ, represents the deviation in local energy density in comparison with the mean background energy [11]:

$$\delta(t,\vec{r}) \equiv \frac{\rho(t,\vec{r}) - \bar{\rho}(t)}{\bar{\rho}(t)}\,. \tag{2.42}$$

In this equation, $\rho(t,\vec{r})$ is the matter density at a general position $\vec{r}$ and a proper time t; and $\bar{\rho}(t)$ is the average background energy density of the Universe [13].

Moreover, the evolution of the LSS has been studied through a combination of large observational surveys and computer simulations. By comparing quantitative measures of clustering from surveys like the Sloan Digital Sky Survey (SDSS) [16], Figure 2.3, with the predictions of simulations [17], Figure 2.4, we can test our cosmological model.

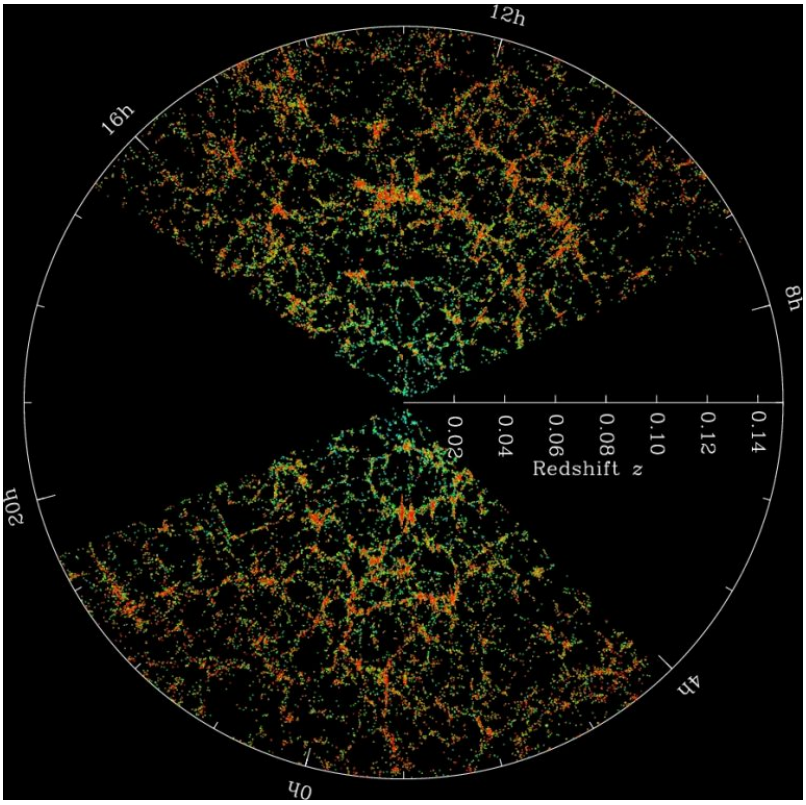


FIGURE 2.3: Observed Galaxy Distribution from SDSS. A slice of the observed Universe, revealing the filamentary LSS. The statistical properties of this distribution are directly compared to simulations to measure cosmological parameters (Figure Credit: M. Blanton and SDSS) [18].

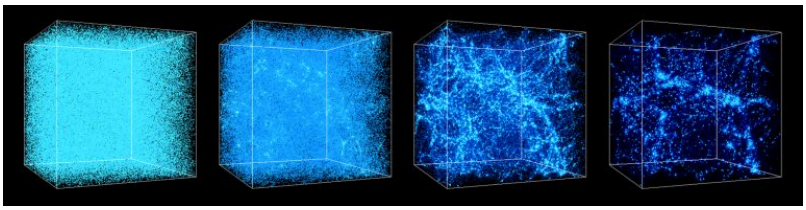


FIGURE 2.4: Simulated Growth of Structure. As time increases from left ot right, the cosmological simulation shows the gravitational collapse of initial randomly distributed matter into a filamentary structure. (Figure Credit: Andrey Kravtsov, Anatoly Klypin, National Center for Supercomputer Applications (NCSA) [17]).

2.4.2. Measuring LSS through Galaxy Number Counts

The introduction to concepts such as redshift (see section 2.1.4) and density contrast (section 2.4.1), allows us to interpret galaxy survey observations and gain insight into the perturbations of the Universe's energy content.

In photometric redshift galaxy surveys, what the observer measures are the **number of galaxies** in a certain direction $\hat{n}$, within a redshift interval centred at z. This measurement is written as $N(\hat{n}, z)d\Omega_{\hat{n}}dz$, which, when averaged over all angles in the sky, defines

the mean redshift distribution $\bar{N}(z)dz$. From this, one can obtain the redshift density perturbation [19]:

$$\begin{aligned}\delta_z(\hat{n},z) = \frac{\rho(\hat{n},z)-\bar{\rho}(z)}{\bar{\rho}(z)} &= \frac{\frac{N(\hat{n},z)}{V(\hat{n},z)}-\frac{\bar{N}(z)}{V(z)}}{\frac{\bar{N}(z)}{V(z)}} \\ &= \frac{N(\hat{n},z)-\bar{N}(z)}{\bar{N}(z)} - \frac{\delta V(\hat{n},z)}{V(z)}\,,\end{aligned} \tag{2.43}$$

where $V(\hat{n},z)$ is the volume element and δV its perturbation.

However, we don't actually observe the redshift density perturbation. The truly observed quantity is the perturbation in the number density of galaxies, Δ, commonly referred to as the **galaxy number counts**. In other words, instead of observing the fluctuations in matter density, for a certain region of space, we only observe the fluctuations in the total number of galaxies, in that region. Then, the galaxy number counts is defined as [19]

$$\Delta(\hat{n},z) \equiv \frac{N(\hat{n},z)-\bar{N}(z)}{\bar{N}(z)}\,, \tag{2.44}$$

which, when connected with equation (2.43), gives the relation

$$\Delta(\hat{n},z) = \delta_z(\hat{n},z) - \frac{\delta V(\hat{n},z)}{V(z)}\,. \tag{2.45}$$

Moreover, if we explicitly compute δ_z and the volume perturbations, $\delta V/V$, then the perturbation of total number density of galaxies is given by [19]

$$\begin{aligned}\Delta(\hat{n},z) =& D_g + \Phi + \Psi + \frac{1}{\mathcal{H}}\left[\dot{\Phi} + \partial_r(\vec{v}\cdot\hat{n})\right] \\ &+ \left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2} + \frac{2}{r(z)\mathcal{H}}\right)\left(\Psi + \vec{v}\cdot\hat{n}\int_0^{r(z)}(\dot{\Phi}+\dot{\Psi})\,dr'\right) \\ &+ \frac{1}{r(z)}\int_0^{r(z)}\left[2 - \frac{r(z)-r'}{r'}\nabla^2_\Omega\right](\Phi+\Psi)\,dr'\,.\end{aligned} \tag{2.46}$$

In equation (2.46) we have that:

- Φ and Ψ are the Bardeen gravitational and curvature potentials (see section 2.3.1);
- D_g is the density fluctuation in comoving gauge [20],
- $\vec{v}$ is the gauge-invariant quantity associated with the peculiar velocity of galaxies [20];
- $\mathcal{H}(z) = aH(z)$ represents the comoving Hubble parameter (see eq.(2.38));
- ∇^2_Ω is the angular Laplacian.

Additionally, the dots denote derivatives with respect to conformal time τ; the perturbations in the first two lines are evaluated at the unperturbed coordinates $(\tau(z), -r(z)\hat{n})$; and what is inside the integrals is evaluated at conformal time $(\tau_0 - r)$ and comoving radius r.

Furthermore, as it stands, equation (2.46) gives the perturbation of the total number density of galaxies. In practice, however, we cannot observe all galaxies, but only those with a larger flux than a certain limit. Usually this limit is given in terms of a limiting magnitude $m_\star$ related to a limiting flux $F_\star$ by $m = -2.5 \log_{10} F + \text{const.}$ (the constant depends on the units in which we measure the flux). Then, from Montanari and Durrer [21] we obtain the galaxy number counts expression[1] that is constrained by a magnitude cut-off of $m < m_\star$:

$$\begin{aligned}\Delta(\hat{n}, z, m_\star) = b(z)\delta_M &+ \frac{1}{\mathcal{H}}\left[\dot{\Phi} - i\left(\vec{k}\cdot\hat{n}\right)\left(\vec{v}(\vec{k})\cdot\hat{n}\right)\right] \\ &+ (2-5s)\left[\int_0^{r(z)} \frac{1}{r(z)}(\Phi+\Psi)dr' - \kappa\right] \\ &+ \mathcal{H}(b_{evo}-3)i\frac{\vec{v}(\vec{k})}{\vec{k}} + (5s-2)\Phi + \Psi \\ &+ \left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2} + \frac{2-5s}{r(z)\mathcal{H}} + 5s - b_{evo}\right)\left(\Psi - \vec{v}(\vec{k})\cdot\hat{n} + \int_0^{r(z)}(\dot{\Phi}+\dot{\Psi})dr'\right)\end{aligned} \quad . \tag{2.47}$$

It is important to acknowledge that equation (2.47) was written in Fourier space. Moreover, this equation incorporates the galaxy density fluctuation term D_g, from equation (2.46), which in Fourier space is defined by $D_g = b(z)\,\delta_M + \mathcal{H}(b_{evo}-3)i\frac{\vec{v}(\vec{k})}{\vec{k}}$. Here, the δ_M denotes the matter density contrast (check eq. (2.42)). Additionally, to make equation (2.47) more realistic, three bias parameters were included:

1. **Galaxy bias,** $b(z)$

   The galaxy bias term, $b(z)$, was introduced, in the last equation, to account for the fact that, instead of measuring the perturbations in overall matter density, we're only measuring those perturbations in baryonic matter, through galaxies;

2. **Evolution bias,** b_{evo}

   In equation (2.46), we neglect a possible evolution of the number counts, i.e. we assume that $a^3\bar{n}$ is constant ($\bar{n}$ denotes the background number density). However, the fact that, as time goes on, new galaxies form and galaxies merge, implies the number density evolution is not as simple as $(1+z)^3$ (see description of eq.(2.14)).

[1] For the derivation of eq.(2.47), starting from eq.(2.46), check: Dio et al. [20] and Challinor and Lewis [1]

To account for this, we introduce the **evolution bias** term [21]:

$$b_{evo}(z, L_\star) = \frac{\partial \ln \left(a^3 \bar{n} \left(z, L > L_\star \right) \right)}{\partial \ln a} \, ; \tag{2.48}$$

3. **Magnification bias,** $s(z)$

   The magnification bias parameter $s(z, m_\star)$ characterises how gravitational lensing affects galaxy number counts in flux-limited surveys. As discussed in section 2.2, lensing modifies observed galaxy densities through two competing effects: flux amplification, which brings fainter galaxies above the detection threshold, and solid-angle distortion, which reduces apparent galaxy number densities [22]. This parameter is defined as the logarithmic derivative of background number density, with respect to magnitude [21],

   $$s(z, m_\star) \equiv \left. \frac{\partial \log_{10} \bar{n}(z, m < m_\star)}{\partial m} \right|_{m_\star} \tag{2.49}$$

   Then, the impact of $s(z, m_\star)$ depends on the slope of the luminosity function, at the survey's magnitude limit $m_\star$. Furthermore, the sensitivity of the survey to magnification lensing is encoded in $s(z)$, i.e. a steeper luminosity function (larger s) implies stronger magnification-induced fluctuations in observed galaxy counts.

Additionally, the expression for galaxy number counts, limited by a magnitude threshold $m_\star$, see equation (2.47), can be decomposed into five distinct physical contributions [21],

$$\Delta(\hat{n}, z, m_\star) = \Delta^M + \Delta^{RSD} + \Delta^L + \Delta^V + \Delta^P \, . \tag{2.50}$$

These components represent the following effects:

1. **Matter Density Fluctuation:**

   $$\begin{aligned} \Delta^M(\hat{n}, z) &= D_g \\ &= b(z) \, \delta_M + \mathcal{H}(b_{evo} - 3) i \frac{\vec{v}(\vec{k})}{\vec{k}} \, . \end{aligned} \tag{2.51}$$

   This term represents the local fluctuation in the number of galaxies due to the underlying matter density contrasts δ_M, biased by the galaxy bias factor $b(z)$.

2. **Redshift-Space Distortions (RSD):**

   $$\Delta^{RSD}(\hat{n}, z) = -i \mathcal{H}^{-1} \left(\vec{k} \cdot \hat{n} \right) \left(\vec{v}(\vec{k}) \cdot \hat{n} \right) \, . \tag{2.52}$$

This term accounts for the apparent distortion in the galaxy distribution along the line of sight caused by the peculiar velocities of galaxies ($\vec{v}$), which add to the cosmological redshift.

3. **Lensing Magnification:**

$$\boxed{\Delta^L(\hat{n},z) = -\big(2-5s(z)\big)\kappa(z)} \,. \tag{2.53}$$

This term describes the change in observed number counts due to gravitational lensing. The convergence κ magnifies (or demagnifies) the brightness of galaxies, shifting them across the magnitude limit $m_\star$ and thus modulating their detectability. The parameter s is the slope of the luminosity function.

4. **Doppler Terms:**

$$\Delta^V(\hat{n},z) = \left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2} + \frac{2-5s}{r(z)\mathcal{H}} + 5s - b_{evo}\right)\left(-\vec{v}(\vec{k})\cdot\hat{n}\right).$$

These are additional corrections arising from peculiar velocities and their evolution.

5. **Potential Terms:**

$$\begin{aligned}\Delta^P(\hat{n},z) = \frac{1}{\mathcal{H}}\dot{\Phi} + (2-5s)\int_0^{r(z)} \frac{1}{r(z)}(\Phi+\Psi)\,dr' + (5s-2)\Phi + \Psi \\ + \left(\frac{\dot{\mathcal{H}}}{\mathcal{H}^2} + \frac{2-5s}{r(z)\mathcal{H}} + 5s - b_{evo}\right)\left(\Psi + \int_0^{r(z)} (\dot{\Phi}+\dot{\Psi})dr'\right).\end{aligned}$$

These are integrated Sachs-Wolfe and other general relativistic effects arising from the time-varying gravitational potentials (Φ, Ψ) along the line of sight.

For the analysis of this work, the primary focus will be on the matter density, Δ^M, redshift-space distortion, Δ^{RSD}, and - most importantly - the **lensing magnification**, Δ^L. The lensing term is of particular interest as it provides a unique and powerful probe of the matter distribution between the observer and the source population, allowing us to constrain cosmological parameters through its integrated effect. This work will specifically investigate the impact and detectability of this lensing contribution.

2.4.3. Weak Gravitational Lensing in Cosmology

The lensing potential provides a crucial link between observed light deflections and the underlying matter distribution. While typically expressed in terms of angular coordinates $\vec{\theta}$ on the sky, for cosmological applications it is more practical to formulate the

lensing potential as a function of comoving distance $\vec{r}$.

Following Lewis and Challinor [23], we relate angular positions to comoving distances via angular diameter distances. Since we are considering a flat Universe, that expression simplifies to

$$\vec{\alpha}(\hat{n}, r_\star) = -2 \int_0^{r_\star} \frac{r_\star - r}{r_\star\, r} \nabla\Psi(r\hat{n}; \tau_0 - r)\, dr\,. \tag{2.54}$$

Moreover, if we define the lensing potential so that the deflection angle is $\vec{\alpha} = \nabla\psi$, then [23]

$$\psi(\hat{n}, r_\star) \equiv -2 \int_0^{r_\star} \frac{r_\star - r}{r_\star\, r} \Psi(r\hat{n}; \tau_0 - r)\, dr\,. \tag{2.55}$$

Recalling the relation between κ and ψ in equation (2.25) might lead one to think of using this relation, in conjunction with equation (2.55), in order to obtain an expression for κ. However, this would result in a less intuitive interpretation of convergence. Since the lensing potential of equation (2.55) is negative, we should add a minus sign to the expression of κ. This ensures that it retains the physically intuitive meaning of being positive when over-densities cause convergence of light rays : magnification. Therefore, by employing this convention [21, 23],

$$\kappa = -\frac{1}{2}\nabla^2_\Omega \psi\,, \tag{2.56}$$

which implies that

$$\kappa(\hat{n}, r_\star) = \int_0^{r_\star} \frac{r_\star - r}{r_\star\, r} \nabla^2_\Omega \Psi(r\hat{n}; \tau_0 - r)\, dr\,. \tag{2.57}$$

This formula directly connects the observable weak lensing convergence to the underlying gravitational potential, Ψ, of the Large-Scale Structure of the Universe.

Until now we've discussed Gravitational Lensing in GR, nevertheless, what if GR is not the best model to describe this phenomenon? Because of this uncertainty, we will introduce a simple scaling factor denoted by β. This parameter has the objective of probing for deviations from GR, and from the standard Lambda-Cold Dark Matter (ΛCDM) cosmological model. Consequently, the lensing potential, $\psi(z_i, \hat{n})$, will be modified by the addition of parameter β [21],

$$\psi(z_i, \hat{n}) = \beta\psi^{\Lambda\text{CDM}}(z_i, \hat{n})\,. \tag{2.58}$$

In the standard ΛCDM model the fiducial value of β is 1.

While deviations from $\beta = 1$ could arise from dark energy clustering or other extensions to ΛCDM without modifying GR [24], such anomalies typically indicate modified

gravity theories, making β a valuable probe of fundamental physics [21].

An important distinction to make is that equation (2.55) assumes the Shwarzchild metric, which implies that the gravitational potentials are approximately equal, $\Phi \approx \Psi$. This is why in the lensing potential equation there's a 2Ψ term but no Φ term. Contrary to that, **in Cosmology** (flat FLRW metric), we must consider both potentials. Therefore [21],

$$\psi^{\Lambda\text{CDM}}(\hat{n}, z) = -\int_0^{r(z)} \frac{r(z) - r'}{r(z)r'} \left(\Phi + \Psi\right)\left(r'\hat{n}, \tau_0 - r'\right) dr' \,. \tag{2.59}$$

If we now introduce equation (2.58), the previous equation becomes

$$\psi(\hat{n}, z) = -\int_0^{r(z)} \frac{r(z) - r'}{r(z)r'} \; \beta \; \left(\Phi + \Psi\right)\left(r'\hat{n}, \tau_0 - r'\right) dr' \,. \tag{2.60}$$

Finally, by applying equation (2.56) to equation (2.60) we obtain the expression for the magnification convergence, in Cosmology,

$$\boxed{\kappa(\hat{n}, z) = \frac{1}{2}\int_0^{r(z)} \frac{r(z) - r'}{r(z)r'} \; \beta \; \nabla^2_\Omega \left(\Phi + \Psi\right) \; dr'} \,. \tag{2.61}$$

This equation can be utilised to constrain the value of $\Phi + \Psi$, which, as mentioned in section 2.3.2, can be exploited to test the validity of GR. However, this subject goes beyond the scope of this dissertation. So, the overall procedure to constrain this relation is only outlined in Chapter 6, Figure 6.1, where we discuss the possible continuations of this project.

3. Estimators of Large-Scale Structure

Having established the fundamental principles relevant to this work in the previous Chapter, the next step is to utilise these principles to determine the clustering of galaxies and thus measure LSS. That will be the goal of this chapter. In order to achieve this, it is first necessary to explain the two LSS estimators that will be utilised. These are the Angular Power Spectrum, C_ℓ and the Angular Correlation function, $w(\theta)$, and will be explained in section 3.1.

Following this introductory step, we can then compute the LSS estimators. The process employed for the calculation will be explained in section 3.2. Additionally, an intuitive interpretation of the results will also be provided.

To finish, the results obtained with these estimators will be connected with the focal point of this work, i.e. the analysis of magnification lensing convergence, κ. In section 3.3 we'll demonstrate the impact of κ in our observations of galaxy clustering.

3.1. Statistics of Large-Scale Structure

As was introduced in section 2.4.1, there is important information encoded in the LSS of the Universe. In order to decode this information into measurable quantities, such as the amount of matter clustering, astronomers have tried multiple statistical methods. This section discusses two of these methods, namely the Angular Power Spectrum and the Angular Correlation Function.

Even though we mentioned, in section 2.4.2, that we observe galaxy number counts, Δ, the reality is that we can't measure them directly. Since their statistical average is, by definition, zero, $\langle\Delta\rangle = 0$, what we actually measure are the deviations of number counts from this average. Or, in statistical terms, instead of measuring the observable, we measure its variance. This is relevant since the Probability Distribution Function (PDF) of these matter fluctuations (the deviations from the average) are the means by which we gain insight into the Cosmology of the Universe.

In this instance, the information on the position of each galaxy, in isolation, is not pertinent. The relevant information is the correlation between the positions of multiple galaxies. In order to measure these correlations, we will utilise statistical estimators.

Since we're analysing three-dimensional data, it is evident that the most intuitive statistical estimator to employ is a $3D$ one, such as the well-known $3D$ matter power spectrum. However, this method requires converting observed angular positions and redshifts into comoving distances, a process that inherently depends on an assumed cosmological model. This introduces a model-dependent bias into the analysis. Fortunately, this issue does not arise with $2D$ statistics, since the observed angular positions of galaxies allow us to operate directly on the observed celestial sphere. Even though this method implies an averaging over the considered redshift bin, for the purpose of this thesis, the effect we want to study is not washed away in this averaging. Therefore, by using statistical estimators such as C_ℓ and $w(\theta)$, our analysis will be independent of the underlying cosmology.

3.1.1. Angular Power Spectrum

The Angular Power Spectrum, denoted by C_ℓ, is a $2D$ statistical estimator used to quantify the clustering of matter/galaxies on a two dimensional map of the sky.

Since the structure of the Universe is distributed throughout space one can divide this three dimensional space into several spherical shells, centred on Earth. Then, the fluctuations of matter, on the $2D$ surface of all those shells, can be quantified with C_ℓ.

Mathematically, this estimator is derived from a Fourier analysis on a sphere using spherical harmonics - the $2D$ analogues of sine and cosine waves. In it, the multipole moment ℓ represents an angular scale on the sky, approximately given by $\theta \sim 180°/\ell$; and the value of C_ℓ itself quantifies the amplitude of the fluctuations at that specific scale [25].

As a means to obtain the spherical representation of a certain function f we start from the equation of Laplace, which in spherical coordinates is given by [26]

$$\nabla^2 f = 0 \tag{3.1}$$

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial\varphi^2} = 0\,. \tag{3.2}$$

This equation can be solved through separation of variables by separating the radial solution, $R(r)$, from the angular solution, $Y(\theta,\phi)$. That can be done by applying the following

ansatz to the Laplace equation [26]:

$$f(r,\theta,\phi) = R(r)\, Y(\theta,\phi)\,. \tag{3.3}$$

After solving it we obtain the expression of the angular solution, which is the Spherical Harmonic [26],

$$\begin{aligned} Y_l^m(\theta,\varphi) &= \Theta_l^m(\theta)\,\Phi_m(\varphi) \\ &= \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}\, P_l{}^m(\cos\theta) e^{im\varphi}\,. \end{aligned} \tag{3.4}$$

Spherical harmonics describe the surface of a sphere. It doesn't matter how complex the function that describes the surface of the sphere is, it can always be described by a sum of spherical harmonics. From this description we can expand the number counts function, $\Delta(\hat{n},z)$, into spherical harmonics, since it lays on the surface of a sphere with radius, z. So, [20]

$$\Delta(\hat{n},z) = \sum_{\ell=0}^{\ell}\sum_{m=-\ell}^{m=\ell} a_{\ell,m}(z) Y_{\ell,m}(\hat{n})\,, \qquad \text{with} \quad a_{\ell m} = \int \Delta(\hat{n},z) Y^*_{\ell m}(\hat{n}) d\Omega\,. \tag{3.5}$$

Here, $a_{\ell,m}(z)$ is the amplitude of the perturbations $Y_{\ell,m}(\hat{n})$ on the surface of a sphere with radius z. In Figure 3.1, we can observe a visualisation of the spherical harmonics for the first 4 multipoles. Furthermore, the colours on the surface of the spheres represent the $a_{\ell m}$s.

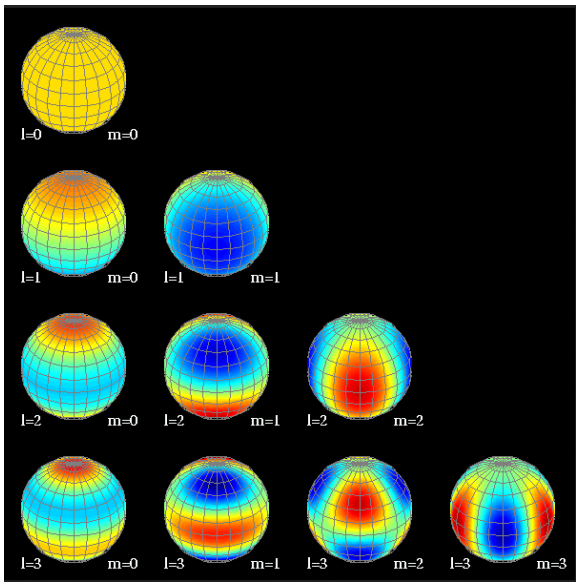


FIGURE 3.1: Representation of multiple components of spherical harmonics, from the monopole $\ell = 0$ to the multipole$\ell = 3$ [27].

Similarly to what was previously mentioned, in the beginning of this chapter, the statistical average of $a_{\ell m}$ is $\langle a_{\ell m} \rangle = 0$, for an unmasked surface, implying that we can not measure the amplitudes of the fluctuations of matter directly. Therefore, we will use **its** variance [21]:

$$\langle a_{\ell m}(z) a^*_{\ell' m'}(z') \rangle \equiv \delta_{\ell \ell'} \delta_{m m'} C_\ell(z, z') \,. \tag{3.6}$$

Here, C_ℓ is the Angular Power Spectrum and it represents the variance of the amplitude of matter fluctuations. To clarify the notation, the assumption of isotropy is what allows us to infer that $\langle a_{\ell m}(z) a^*_{\ell' m'}(z') \rangle$ is a function of ℓ alone, and not m. [25] This is why we write C_ℓ instead of $C_{\ell m}$.

In general, the angular power spectrum, $C_\ell(z_i, z_j)$, between two central redshifts, z_i and z_j, can be written as [20]

$$C_\ell(z_i, z_j) = 4\pi \int_0^{+\infty} \frac{1}{k} \mathcal{P}(k) \Delta_\ell(z_i, k) \Delta_\ell(z_j, k) \, dk \,. \tag{3.7}$$

From here, one should notice two details:

1. The primordial matter power spectrum, $\mathcal{P}(k)$, which represents the initial fluctuations of matter in the Universe. It is given by [28]

$$\mathcal{P}(k) = A_s \left(\frac{k}{k_0} \right)^{n_s - 1} \,, \tag{3.8}$$

   where the pivot scale is $k_0 = 0.05$ Mpc^{-1}, n_s is the spectral index and A_s a normalisation amplitude.

2. The number counts are no longer being expressed in order of $\hat{n}$, but are now in order of k, $\Delta(\hat{n}, z) \rightarrow \Delta_\ell(z, k)$. The full expression of $\Delta_\ell(z, k)$ can be seen in the paper from Dio et al. [20].

Moreover, equation 3.7 is valid for a pair of infinitely thin spherical shells located at redshift z_i and z_j. However, when using realistic photometric data, the redshift bins considered have finite thickness, which is why they can be described by window selection functions, $W(z, z_c)$. An example of this function is a Gaussian centred at some redshift z_i, denoted by $W(z, z_c = z_i)$, and scaled to a given number density of galaxies per redshift interval dN/dz. We define the number counts of a thick shell as [20]

$$\Delta_{\ell, i}(k) \equiv \int \frac{dN}{dz} W(z, z_i) \Delta_\ell(z, k) \, dz \,. \tag{3.9}$$

Therefore, the expression we'll use to analyse the angular power spectrum is [20]

$$C_{\ell,ij} = \int \frac{1}{k}\mathcal{P}(k)\Delta_{\ell,\,i}(k)\Delta_{\ell,\,j}(k)\; dk \quad . \tag{3.10}$$

When $i = j$, the redshift bins considered are the same, so the estimator becomes $C_{\ell,ii}$ and it represents the autocorrelation, of the angular power spectrum, of the spherical shell with radius z_i. Notwithstanding, for the case that $i \neq j$, the estimator represents the angular power spectrum of the cross-correlation between two separate shells.

In an additional note, we can also see that, when the window functions are Dirac distributions, $W(z, z_c) = \delta(z_c - z)$, we recover equation (3.7), since in that case, $C_{\ell,\,ij} = C_\ell(z_i, z_j)$.

3.1.2. Two-point Angular Correlation Function

We've just seen that the Angular Power Spectrum, C_ℓ, quantifies clustering on shells of the celestial sphere, through the use of spherical harmonics . A complementary statistic is the **Two-point Angular Correlation Function**, denoted by $w(\theta)$, which also quantifies clustering of galaxies but in an angular configuration space. This function describes the probability of having an excess of galaxy-pairs (this is why it's called Two-point), within a certain angle θ, relative to the expected value (the mean density of galaxies). The angle θ indicates the size of the patch of the sky being observed, within the celestial sphere.

Formally, this estimator is the fractional increase (relative to a random distribution) in the probability, δP, of finding objects in two solid angle elements, $\delta\Omega_1$ and $\delta\Omega_2$, separated by θ [25]:

$$\delta P = \zeta^2 \left[1 + w(\theta)\right] \delta\Omega_1 \delta\Omega_2 \,, \tag{3.11}$$

where ζ is the object surface density, that is, the density of galaxies in a certain redshift shell.

The advantage $w(\theta)$ has over C_ℓ is that it's considered to be simple and intuitive [25]. As galaxy clustering increases, so does the number of galaxy pairs within smaller angles, and $w(\theta)$ quantifies this increase as a function of galaxy separation, θ [25]. In other words, a positive value of $w(\theta)$ indicates clustering, a value of zero implies no correlation, and a negative value suggests anti-clustering (a repelling effect).

Let us see how this estimator can be utilised to quantify clustering. If we consider the observable Universe to be a sphere of radius z_{CMB}, the redshift of the Cosmic Microwave Background (CMB), then we can divide it into multiple spherical shells of radius z. The distribution of galaxies within a certain shell can be analysed to compute $w(\theta)$ for that specific shell/map of the sky. This is illustrated in Figure 3.2, which shows us two generated sky distributions: one (upper left) simulating low-contrast galaxy clusters in a regular grid on a random background; the other (upper right) simulating galaxy clusters on a background with LSS in the form of a quadrupole.

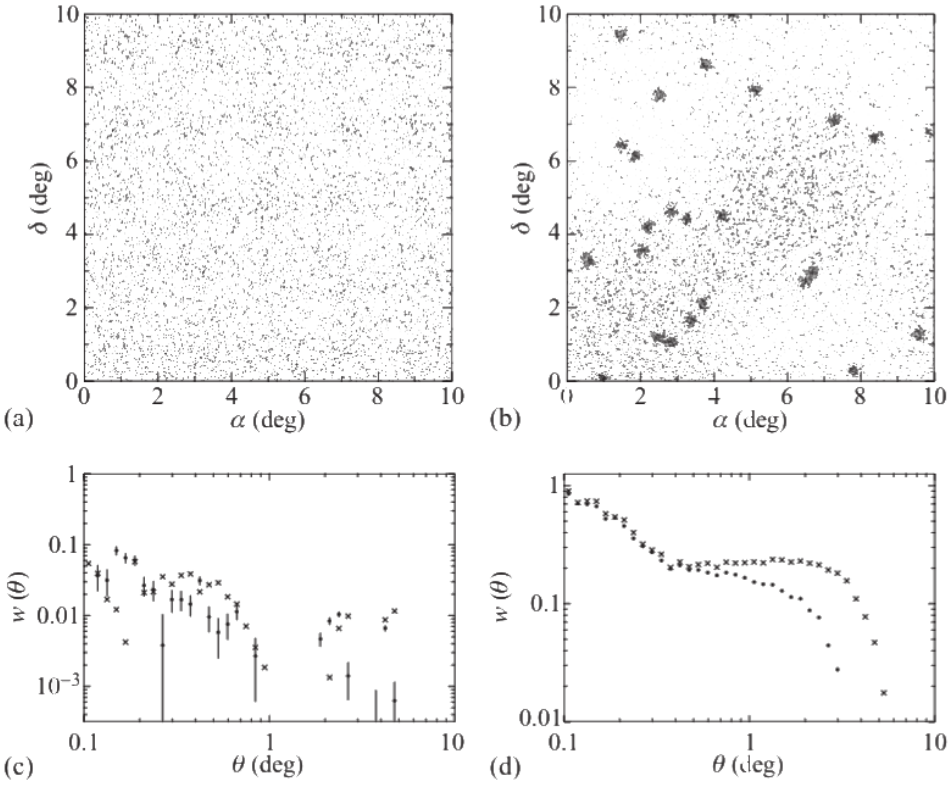


FIGURE 3.2: Two sky simulations: (a) is a uniform 10° grid of 25 low-surface-brightness clusters on a random background; and (b) is a quadrupole background with 25 randomly placed clusters. Measured $w(\theta)$ appear below each, (c) and (d), respectively. Adapted from Practical Statistics for Astronomers [25].

Fundamentally, $w(\theta)$ is a projection of the three-dimensional spatial correlation function $\xi(r, z)$ along the line of sight [29]. However, for this analysis, we exploit the fact that $w(\theta)$ and the angular power spectrum, C_ℓ, provide equivalent descriptions of the galaxy

distribution and are statistically related. These two quantities are connected by the well-known relation [8]

$$w_{ij}(\theta) = \sum_{\ell \geq 0}^{\ell_{max}} \left(\frac{2\ell + 1}{4\pi} P_\ell(\cos\theta) C_{\ell,\,ij} \right) \,, \tag{3.12}$$

where P_ℓ represent the Legendre polynomials. Here we introduced $C_{\ell,ij}$ from equation (3.10), which gives us the tomographic expression of $w(\theta)$ for the correlation between galaxies in redshift bins i and j. This allows us to compute the angular correlation function directly from the C_ℓ spectra obtained from our number count analysis.

3.2. Computing LSS estimators

This section details the practical calculation of both C_ℓ and $w(\theta)$, for a photometric galaxy survey designed to mimic the Dark Energy Survey (DES) [30]. First, we define the initial conditions in common for the calculation of both estimators. These are: the galaxy redshift distribution, the photometric window selection functions and the fiducial cosmology. Then, we outline the computational methods used to calculate each estimator. To finish, we present the results and explain how to interpret them.

3.2.1. Survey and Cosmology Specifications

The galaxy distribution we'll use is very similar to the one used in Fonseca et al. [30], so we'll consider it to be a DES-like distribution. We will define the normalised distribution of galaxy sources, for a certain redshift, $p(z)$, as [30]

$$p(z) = \left(\frac{z}{z_0} \right)^{1.04} \exp\left[- \left(\frac{z}{z_0} \right)^{1.74} \right] \,, \tag{3.13}$$

in which $z_0 = 0.57$ [30].

Moreover, the galaxy number density distribution, per redshift and per steradian (the same concept that was mentioned in the beginning of section 2.4.2) is given by

$$\frac{dN}{dz d\Omega} = \frac{N_{gal}}{\Omega} \times p(z) \,. \tag{3.14}$$

Here, we will consider the scale factor (the value of number density of galaxies, per arcmin^2) to be $N_{gal}/\Omega = 22.36$ [galaxies/arcmin^2] $= 80496$ [galaxies/deg^2] [30].

The aforementioned Window Selection function, $W(z)$, (see section 3.1.1) divides our data into multiple redshift bins, which in our case will be photometric. Although spectroscopic surveys provide higher-quality data, the sheer number of known galaxies is too large for these measurements to be feasible. This is because spectroscopic surveys are much more time-consuming than photometric surveys [31]. Therefore, instead of using Dirac delta functions for exact spectroscopic redshifts measurements, we have to rely on photometric measurements, which have Gaussian errors associated with them [31] . Hence, our window selection function will be a Gaussian distribution, centred at a certain redshift z_c, applied to the total distribution of galaxies (defined in eq.(3.14)), in order to adjust it:

$$W(z) = \frac{dN}{dzd\Omega} \times \frac{1}{\sigma_z\sqrt{2\pi}} \exp\left[-\frac{(z-z_c)^2}{2\sigma_z^2}\right] , \tag{3.15}$$

with $\sigma_z = 0.05(1+z)$ as the photometric redshift error [32].

In our DES-like distribution we will consider 5 photometric z-bins with the characteristics we just described; their central redshifts are $z_c = \{0.585\ ,\ 0.66\ ,\ 0.775\ ,\ 0.89\ ,\ 1.175\}$[30].

For the purpose of a better understanding of equations (3.14) and (3.15) we produced a visual aid, which is Figure 3.3.

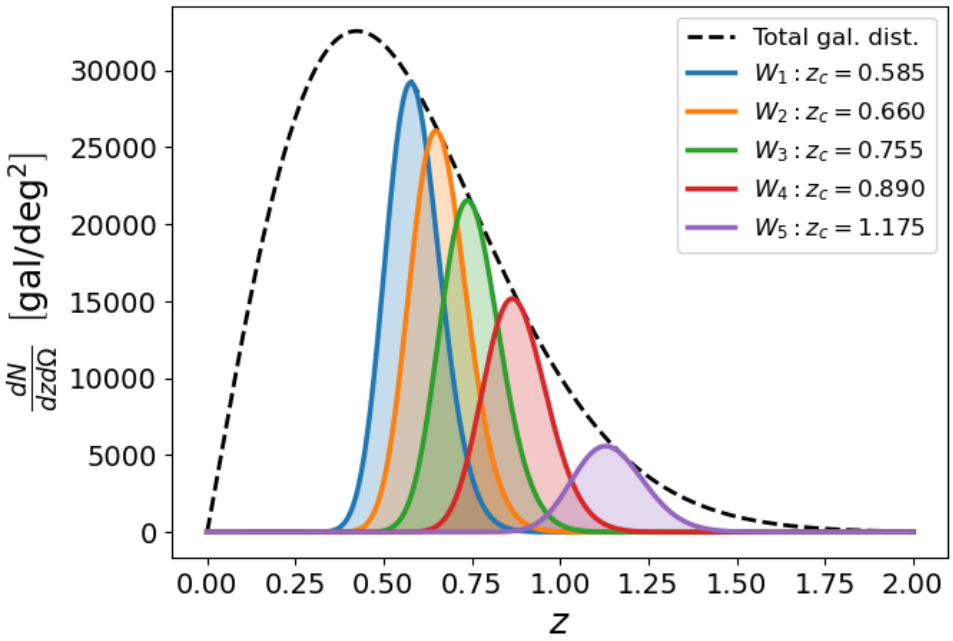


FIGURE 3.3: DES-like galaxy redshift distribution (black dashed line) and the five Gaussian photometric window functions $W(z)$ used in this analysis. The window functions were not normalised for visualisation purposes

Furthermore, the estimators calculations assume a fiducial ΛCDM cosmology with parameters consistent with Planck 2018 collaboration [33] (see section 4.1.4 for details). These parameters are used by the downstream packages to compute growth functions and matter power spectra.

3.2.2. Software Packages Used

The angular power spectrum, C_ℓ, and the angular correlation function, $w(\theta)$, are computed using the `PyCAMB` [1] and `PyCCL` [2] packages, respectively. The calculation proceeds in two steps.

First, the angular power spectrum $C_\ell(z_i, z_j)$ is computed for each pair of redshift bins (z_i, z_j) using `PyCAMB`. The required inputs are:

- The redshift window functions, $W(z_i)$ and $W(z_j)$;
- A fiducial cosmology (see section 4.1.4);
- A model for galaxy bias. We adopt a simple linear bias model [30]:

$$b(z_i) = 1 + 0.84 z_i \,; \tag{3.16}$$

- The desired multipole range, $\ell \in [\ell_{min}, \ell_{max}]$.

Second, the angular correlation function $w(\theta)$ is calculated from the C_ℓ values using `PyCCL`. The inputs for this step are:

- The C_ℓ values computed in the previous step;
- The fiducial cosmology;
- The angular separation range, $\theta \in [\theta_{min}, \theta_{max}]$.

Furthermore, the range of values chosen for the ℓs and θs is based on the following relation [25]:

$$\theta \simeq \frac{180^\circ}{\ell} \,, \tag{3.17}$$

which was roughly mentioned in the beginning of section 3.1.1.

We decided to consider a range that includes the entirety of the observable sky, the full sky, so the maximum value of θ is $\theta_{max} = 90^\circ$, which corresponds to $\ell_{min} = 2$. Then, we set $\ell_{max} = 1000$ in order to include both the linear and the non-linear part of the angular power spectrum, which corresponds to $\theta_{min} = 0.18^\circ$.

3.2.3. Results and their Interpretation

The results from our considerations in sections 3.2.1 and 3.2.2 can be seen in Figure 3.4. This Figure shows the resulting auto-correlation angular power spectrum and its corresponding angular correlation function for the redshift bin centred at $z_c = 0.585$, including only the matter density, Δ^M, redshift-space distortions Δ^{RSD}, and lensing Δ^L contributions to the number counts (see section 2.4.2).

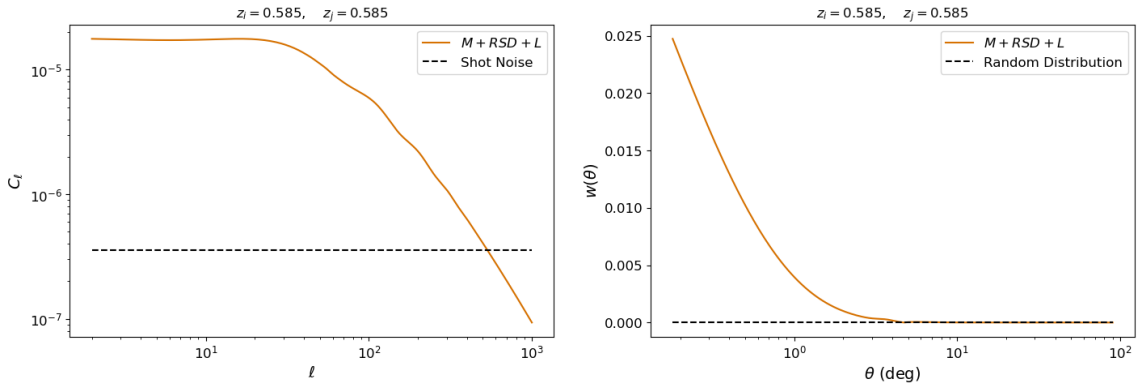


FIGURE 3.4: (Left) The angular power spectrum, C_ℓ, for the DES-like sample. The horizontal dashed line represents the expected shot noise level P_N.
(Right) The corresponding angular correlation function, $w(\theta)$.
This figure shows the contributions from the M+RSD+L signal without error bars.

Let's start by interpreting the $w(\theta)$ plot in Figure 3.4. For example, in a perfectly homogeneous, and random, universe where galaxies are distributed uniformly without any clustering (i.e. a purely random Poisson distribution) the Angular Correlation function would be zero for all angular separations: $w(\theta) = 0$, for all θ. However in a Universe with matter clustering, galaxies are more likely to be found near each other than they would be in a random distribution, therefore $w(\theta) \geq 0$, which is what we observe in the graph.

Still in Figure 3.4, the plot of C_ℓ can be better understood when we consider the relation in Peebles [8],

$$C_\ell = 2\pi \int_{-1}^{+1} w(\theta) P_\ell(\cos\theta)\, d(\cos\theta)\,. \tag{3.18}$$

With this relation in mind, and remembering the previously mentioned example for $w(\theta)$ (a perfectly homogeneous universe with infinite galaxies distributed uniformly without any clustering) one might conclude that, as the value $w(\theta) = 0$ for all θs, so would the value of C_ℓ for all ℓs. However, since the real universe has a finite amount of galaxies, the value of C_ℓ changes, even though $w(\theta)$ remains 0. In this case, there would be a finite number of randomly placed discrete galaxies, with Poisson distribution, which would imply

that $C_{\ell,ij} = P_{N,ij}$ for all ℓs. This change in value is due to shot noise, which represents fluctuations originated by random Poisson sampling, of galaxies. The Shot Noise is given by [20]

$$P_{N,ij} = \frac{\delta_{ij}}{n_i}\,, \qquad \text{where} \qquad n_i = \int W(z_i)\frac{dN}{dz}dz\,, \tag{3.19}$$

in which n_i is the number density of galaxies (gal/str) inside each bin z_i and z_j.

Therefore, by analysing Figure 3.4 again, we can conclude that everything above the horizontal line with amplitude $P_{N,ij}$, for C_ℓ, and amplitude 0 for $w(\theta)$, represents the clustering of galaxies. These signals contain the LSS information we seek!

3.3. Understanding the impact of Lensing with LSS estimators

Given our current understanding of the statistical estimators employed to analyse the LSS of the Universe, we are now prepared to analyse the impact of lensing convergence in our observations. In this process, the discrepancy between autocorrelations and cross-correlations of z-bins will be our primary focus.

Let us begin with the observed galaxy over-density, Δ^G, which we model as a simplified version of equation (2.50):

$$\Delta^G = \Delta^M + \Delta^{RSD} + \Delta^L\,. \tag{3.20}$$

The relative magnitudes of the autocorrelations of these terms are determined by the physical processes they represent [19]:

- **Matter Density and RSD are Local:** The density perturbation and the velocity field are highly correlated for galaxy pairs at the same cosmic time. This strong, local $3D$ clustering signal dominates the autocorrelation.
- **Lensing is Non-Local:** The magnification convergence κ, for a source is an integral of the matter density along the line-of-sight (see equation 2.57). While the autocorrelation $\langle \Delta_i^L \Delta_i^L \rangle$ is non-zero, it is an integrated effect dispersed over a large radial distance. This makes it the subdominant term since, in autocorrelations, we're only considering the local clustering signal in a single redshift bin.

From this, it follows that the **autocorrelation** of matter fluctuations are approximated by

$$\langle \Delta_i^G \Delta_i^G \rangle \simeq \left\langle \Delta_i^M \Delta_i^M \right\rangle + \left\langle \Delta_i^{RSD} \Delta_i^{RSD} \right\rangle + 2\left\langle \Delta_i^M \Delta_i^{RSD} \right\rangle\,. \tag{3.21}$$

This is the expected behaviour obtained in the tests made by Bonvin and Durrer [19], which we confirm in our calculations. Using the initial conditions mentioned in section 3.2.1, the autocorrelation of the bin centred at $z_c = 0.585$ is shown in Figure 3.5. In this figure we can observe that, for both C_ℓ and $w(\theta)$, the Matter density term is the one **that** dominates, while the lensing term is negligible.

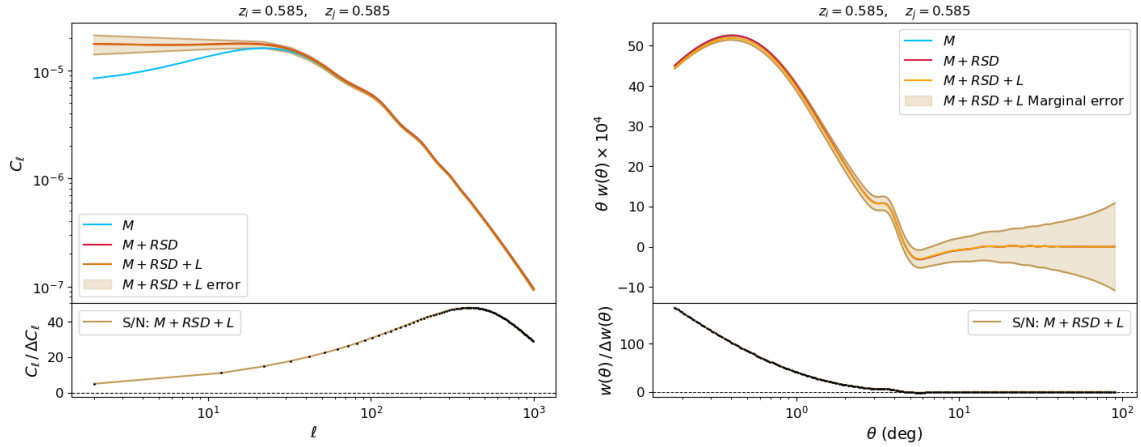


FIGURE 3.5: Autocorrelation of the redshift bin centred at $z_c = 0.585$. The y-axis for $w(\theta)$ is scaled as $\theta w(\theta) \times 10^4$ to better distinguish the contributions of the different components. Bellow each plot we also show their Signal to Noise ratio. As expected, the local matter density term dominates.

There is, however, a drawback in using angular autocorrelations, which is the loss of information as all galaxies, in the bin, are projected onto a single spherical shell. To recover this information, we must incorporate the cross-correlations between different redshift bins. This tomographic approach allows us to probe the $3D$ structure of the Universe by measuring correlations as a function of the radial separation between bins [34]. This is the reason why we will include **cross-correlations** in our analysis.

For bins with sufficient redshift separation, the relative importance of the terms changes drastically. In this instance [35, 36]:

- **Local Terms Vanish:** The correlation between local density fields at different cosmic times, $\langle \Delta_i^M \Delta_j^M \rangle$, becomes vanishingly small due to the rapid decay of the $3D$ matter correlation function with comoving separation. Correlations involving RSD terms also become negligible;

- **Lensing-Density Correlation Dominates:** The dominant surviving term is $\langle \Delta_i^M \Delta_j^L \rangle$. This term is physically interpreted as the correlation between the matter field at the lower redshift z_i (which acts as a lens) and the magnification of sources at the higher

redshift z_j. The reverse term, $\langle \Delta_i^L \Delta_j^M \rangle$, is negligible since high-redshift matter cannot lens lower-redshift sources.

Therefore, for well-separated bins, the cross-correlation between two distinct bins i and j (with $z_i < z_j$) is dominated by the lensing signal:

$$\begin{aligned} \langle \Delta(z_i)\Delta(z_j) \rangle &\simeq \langle \Delta_i^M \Delta_j^L \rangle \\ &= -b(z_i)(2 - 5s(z_j)) \big\langle \delta_M(z_i)\kappa(z_j) \big\rangle \,. \end{aligned} \tag{3.22}$$

Again, this expression agrees with the results obtained in Bonvin and Durrer [19].

Once more, our calculations confirm our predictions. Figure 3.6 shows the cross-correlation between the bins centred at $z_i = 0.585$ and $z_j = 1.175$. In it, the lensing term, which was subdominant in the auto-correlation, now becomes the primary contributor to the signal on most scales. This clear dominance provides a direct and measurable signature of the magnification effect, for both estimators, even when taking into consideration the estimated error bars (see Appendix A).

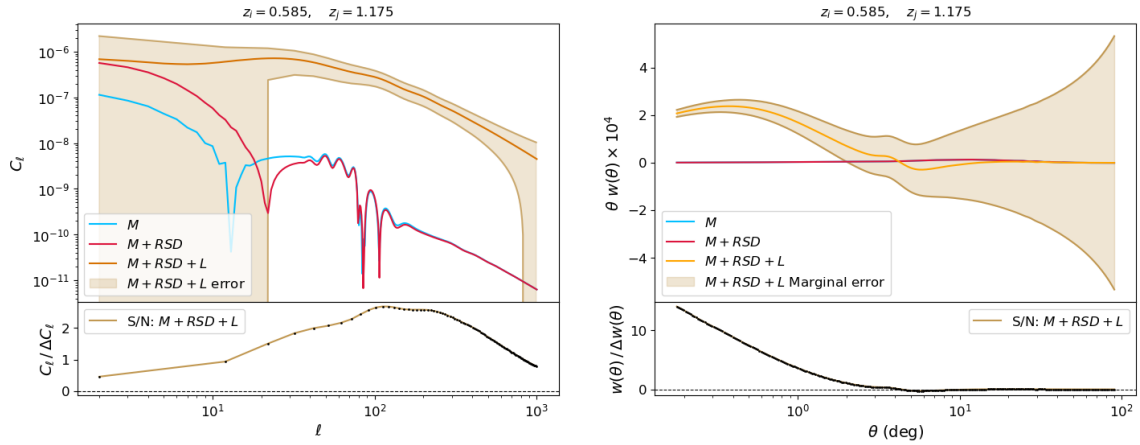


FIGURE 3.6: Cross-correlation of redshift bins centred at $z_i = 0.585$ and $z_j = 1.175$. The y-axis for $w(\theta)$ is scaled as $\theta w(\theta) \times 10^4$. Bellow each plot we also show their Signal to Noise ratio. As predicted, the non-local lensing term becomes the dominant contribution, illustrating how tomographic cross-correlations isolate the magnifying effect of foreground matter on background sources.

In conclusion, this analysis demonstrates that galaxies are not independent tracers of the underlying matter density. The observed galaxy field is modified by weak lensing magnification, an effect that is negligible in single-bin analyses but becomes the dominant feature in the cross-correlation of distant redshift bins. Therefore, a complete tomographic analysis on the LSS of the Universe must account for this non-local effect in order to avoid

biases. Moreover, this effect can be used, in theory, to extract cosmological information from the lensing signal itself.

4. Performance of LSS estimators

This chapter evaluates the precision of the C_ℓ and $w(\theta)$ estimators in constraining cosmological parameters, with particular focus on two parameters: the magnification bias, $s(z)$, and β.

In order to achieve this goal, we introduce the theoretical framework that is necessary to employ Fisher matrices (section 4.1). Then we describe the procedure of computing the Fisher matrices and using their forecasts to obtain the minimum measurement errors achievable with each LSS estimator (section 4.2). Afterwards, we will compare the results for two scenarios: one where the multipole ℓ and angular scale θ are coupled, and another where they are decoupled. This comparison will clarify which estimator performs better for measuring LSS (section 4.3).

4.1. Fisher Information Matrix

For a preliminary analysis on the performance of these two estimators, we will use Fisher Information Matrices. In essence, this method can be utilised to ascertain what is the minimum possible observational error that one would obtain by measuring some cosmological parameter, λ, with these two estimators [37].

The advantage of Fisher matrix calculation is that, given some data, X, it provides an essentially instantaneous result (when compared to the large computational time of Markov Chain Monte Carlo (MCMC) method), with no stochastic noise [37]. This is the reason why we are going to use Fisher matrices to compare the usefulness of these two estimators.

The general expression for each component of the Fisher Matrix is given by [37]

$$M_{\mu\nu} = \bar{\mathbf{X}}^T_{,\mu}\mathbf{C}^{-1}\bar{\mathbf{X}}_{,\nu} + \frac{1}{2}\mathrm{Tr}\left[\mathbf{C}^{-1}\mathbf{C}_{,\mu}\mathbf{C}^{-1}\mathbf{C}_{,\nu}\right] . \tag{4.1}$$

Where the mean of the data vector, $\mathbf{X}$, is a vector containing all the expected values of $\mathbf{X}$, which is given by $\bar{\mathbf{X}} = \langle\mathbf{X}\rangle$. Additionally, the Covariance matrix of the data is $\mathbf{C} = \mathrm{Cov}(\mathbf{X},\mathbf{X}) = \langle(\mathbf{X}-\bar{\mathbf{X}})(\mathbf{X}-\bar{\mathbf{X}})^T\rangle$; and the indices μ,ν represent the cosmological

parameter, λ_μ, that we are considering.

4.1.1. Fisher Matrix of both estimators

In this section we'll explain how we are going to determine each component of the Fisher Matrix, $M_{\mu\nu}$, for both LSS estimators: C_ℓ and $w(\theta)$.

Let's begin by considering the following two cases: that our data vector is the angular power spectrum, $\mathbf{X} = \vec{C}_{\ell,ij}$; and that $\mathbf{X}' = \vec{w}_{ij}(\theta)$. Here, the vector $\vec{C}_{\ell,ij}$ contains a value for each ℓ element that we are considering in a certain redshift bin combination, z_{ij}, where this notation represents $z_{ij} = \{z_i, z_j\}$. The same goes for vector $\vec{w}_{ij}(\theta)$ and each value of θ.

In order to obtain a Fisher matrix we need the mean of the data vector and the covariance matrix of the data. In this cases, the mean of the data vectors are themselves:

$$\bar{\mathbf{X}} = \langle \vec{C}_{\ell,ij} \rangle = \vec{C}_{\ell,ij}\,, \tag{4.2}$$

$$\bar{\mathbf{X}}' = \langle \vec{w}_{ij}(\theta) \rangle = \vec{w}_{ij}(\theta)\,, \tag{4.3}$$

and their covariance matrices are denoted by

$$\text{Cov}\left(\vec{C}_{\ell,ij}, \vec{C}_{\ell',pq}\right) \equiv \Gamma_{\ell\ell',ijpq}\,, \tag{4.4}$$

$$\text{Cov}\Big(\vec{w}_{ij}(\theta), \vec{w}_{pq}(\theta')\Big) \equiv \Gamma_{\theta\theta',ijpq}\,. \tag{4.5}$$

Moreover, recalling equation (4.1) we see that the expression for the Fisher matrix has two components. The first term, $\bar{\mathbf{X}}^T_{,\mu}\mathbf{C}^{-1}\bar{\mathbf{X}}_{,\nu}$, describes how an observable $\mathbf{X}$ depends on parameter λ, while the second term tells us how the error bar of the observable depends on λ. In the case where our observable is either of the LSS estimators, we conclude that the fist term has a greater impact on the Fisher matrix expression, than the second one. Therefore, we will use the following approximation [38]:

$$\bar{\mathbf{X}}^T_{,\mu}\mathbf{C}^{-1}\bar{\mathbf{X}}_{,\nu} \gg \frac{1}{2}\text{Tr}\left[\mathbf{C}^{-1}\mathbf{C}_{,\mu}\mathbf{C}^{-1}\mathbf{C}_{,\nu}\right]\,. \tag{4.6}$$

Then, the expression for the Fisher matrix, of both estimators, reduces to

$$M_{\mu\nu} \simeq \bar{\mathbf{X}}^T_{,\mu}\mathbf{C}^{-1}\bar{\mathbf{X}}_{,\nu}\,. \tag{4.7}$$

Substituting by more explicit notation, we can rewrite the Fisher matrix expression of each estimator as

$$M_{\mu\nu} \simeq \frac{\partial \vec{C}^T_{\ell,ij}}{\partial \lambda_\mu} \left(\Gamma_{\ell\ell',ijpq}\right)^{-1} \frac{\partial \vec{C}_{\ell',pq}}{\partial \lambda_\nu} \,, \tag{4.8}$$

$$M'_{\mu\nu} \simeq \frac{\partial \vec{w}^T_{ij}(\theta)}{\partial \lambda_\mu} \left(\Gamma_{\theta\theta',ijpq}\right)^{-1} \frac{\partial \vec{w}_{pq}(\theta')}{\partial \lambda_\nu} \,. \tag{4.9}$$

So, by interpreting equations (4.8) and (4.9) we conclude that, to obtain each component of $M_{\mu\nu}$, all we need to calculate are: the derivatives $\partial C_{\ell,ij}/\partial\lambda_\mu$ and the covariance matrix $\Gamma_{\ell\ell',ijpq}$ (this logic also applies to $M'_{\mu\nu}$ and $w(\theta)$).

Afterwards, we can use the inverse of the Fisher matrix to provide the lower bound on the variance of the parameters considered. Then, the marginalised error of a certain parameter λ_μ is given by the Cramér-Rao bound [37, 39],

$$\sigma(\lambda_\mu) \geq \sqrt{(M^{-1})_{\mu\mu}} \,. \tag{4.10}$$

4.1.2. Covariance Matrix of the Angular Power Spectrum

The expression for the covariance matrix of $\vec{C}_{\ell,ij}$ is deduced in the appendix of Fonseca and Liguori [28] and is given by

$$\begin{aligned}
\Gamma_{\ell\ell',ijpq} &\equiv \mathsf{Cov}\left(C_{\ell,ij}, C_{\ell',pq}\right) \\
&= \left\langle (C_{\ell,ij} - \bar{C}_{\ell,ij})(C_{\ell',pq} - \bar{C}_{\ell',pq}) \right\rangle \\
&= \left\langle C_{\ell,ij} C_{\ell',pq} \right\rangle - \left\langle C_{\ell,ij} \right\rangle \left\langle C_{\ell',pq} \right\rangle \\
&= \frac{\delta_{\ell\ell'}}{2\ell+1} \left[C_{\ell,iq} C_{\ell,pj} + C_{\ell,ip} C_{\ell,qj} \right] \,.
\end{aligned} \tag{4.11}$$

If we now introduce shot noise from equation (3.19), $P_{N,ij}$, and the fraction of sky that its being observed, f_{sky}, then the equation becomes

$$\boxed{\begin{aligned}
\Gamma_{\ell\ell',ijpq} = \frac{\delta_{\ell\ell'}}{2\ell+1} \frac{1}{f_{sky}} \Big[&\left(C_{\ell,iq} + P_{N,iq}\right) \left(C_{\ell,pj} + P_{N,pj}\right) \\
&+ \left(C_{\ell,ip} + P_{N,ip}\right) \left(C_{\ell,qj} + P_{N,qj}\right) \Big]
\end{aligned}} \,. \tag{4.12}$$

Here, the range of values of f_{sky} is constrained to $f_{sky} = [0, 1]$, where 0 means we have no information and 1 is the case where we have information on the entire sky. Intuitively, we see that the higher the fraction of sky, the better our results will be.

4.1.3. Covariance Matrix of the Angular Correlation Function

Contrary to C_ℓ we didn't find any shortcut for the expression of the covariance matrix of $\vec{w}_{ij}(\theta)$, so we add to deduce it ourselves:

$$\begin{aligned}\Gamma_{\theta\theta',ijpq} &\equiv \mathrm{Cov}\Big(w_{ij}(\theta), w_{pq}\left(\theta'\right)\Big) \\ &= \left\langle (w_{ij}(\theta) - \bar{w}_{ij}(\theta))(w_{pq}(\theta') - \bar{w}_{pq}(\theta'))\right\rangle \\ &= \left\langle \left[\sum_\ell \frac{2\ell+1}{4\pi} P_\ell(\cos\theta)(C_{\ell,ij} - \bar{C}_{\ell,ij})\right]\left[\sum_{\ell'} \frac{2\ell'+1}{4\pi} P_{\ell'}(\cos\theta')(C_{\ell',pq} - \bar{C}_{\ell',pq})\right]\right\rangle \\ &= \sum_\ell \sum_{\ell'} \frac{(2\ell+1)(2\ell'+1)}{(4\pi)^2} P_\ell(\cos\theta) P_{\ell'}(\cos\theta') \left\langle (C_{\ell,ij} - \bar{C}_{\ell,ij})(C_{\ell',pq} - \bar{C}_{\ell',pq})\right\rangle \,.\end{aligned} \tag{4.13}$$

By substituting the covariance of $\vec{C}_{\ell,ij}$ with equation (4.11) we obtain

$$\Gamma_{\theta\theta',ijpq} = \sum_\ell \sum_{\ell'} \frac{(2\ell+1)(2\ell'+1)}{16\pi^2} P_\ell(\cos\theta) P_{\ell'}(\cos\theta') \frac{\delta_{\ell\ell'}}{2\ell+1} \left[C_{\ell,iq}C_{\ell,pj} + C_{\ell,ip}C_{\ell,qj}\right] \,.$$

This simplifies to the general expression of the covariance matrix of $w(\theta)$:

$$\boxed{\Gamma_{\theta\theta',ijpq} = \sum_\ell \frac{(2\ell+1)}{16\pi^2} P_\ell(\cos\theta) P_\ell(\cos\theta') \left[C_{\ell,iq}C_{\ell,pj} + C_{\ell,ip}C_{\ell,qj}\right]} \,. \tag{4.14}$$

I will just note that, for more specific cases, this general expression simplifies to more commonly used expressions [32], such as

$$\Gamma_{\theta\theta',ij} = \sum_\ell \frac{(2\ell+1)}{16\pi^2} P_\ell(\cos\theta) P_\ell(\cos\theta') \left[C_{\ell,ij}^2 + C_{\ell,ii}C_{\ell,jj}\right] \,, \quad \text{when } i=p, j=q\,; \tag{4.15}$$

$$\Gamma_{\theta\theta'} = \sum_\ell \frac{(2\ell+1)}{16\pi^2} P_\ell(\cos\theta) P_\ell(\cos\theta') \left[2C_\ell^2\right] \,, \qquad \text{when } i=j=p=q\,. \tag{4.16}$$

We already have the general expression but, just like with C_ℓ there are still two parameters missing in it: the Shot Noise, P_N, and the fraction of the observed sky, f_{sky}. So, let's introduce these two parameters in equation (4.14), in order to obtain

$$\begin{aligned}\Gamma_{\theta\theta',ijpq} = \frac{1}{f_{sky}} \sum_\ell \Bigg\{ &\frac{(2\ell+1)}{16\pi^2} P_\ell(\cos\theta) P_\ell(\cos\theta') \\ &\Big[\left(C_{\ell,iq} + P_{N,iq}\right)\left(C_{\ell,pj} + P_{N,pj}\right) \\ &+ \left(C_{\ell,ip} + P_{N,ip}\right)\left(C_{\ell,qj} + P_{N,qj}\right) \Big] \Bigg\}\end{aligned} \,. \tag{4.17}$$

Moreover, when we calculate the covariance matrix of $w(\theta)$ we encounter a **discretisation problem** that doesn't appear in the C_ℓ analysis. This is due to the fact that, in

the covariance matrix of C_ℓ we have defined values for each multipole ℓ, since $\ell \in \mathbb{N}$. However, the angles have a wider range of possibilities, since $\theta \in \mathbb{R}^+$. This is problematic since, for a fixed range of values ℓ, there is an equivalent range of angles that contain infinite values in it. Because of this, we will add a correction term to $\Gamma_{\theta\theta',ijpq}$, such that it will account for this discretisation problem. We will create this term based on the correction added in the calculation of the mean. In other words, the same way that the mean of the data x has a discretisation correction of N:

$$\bar{x} = \frac{1}{N}\sum_i^N x_i \,. \tag{4.18}$$

We can also add this term to the covariance of x with y:

$$\mathsf{Cov}(x,y) = \Big\langle \big(x_i - \bar{x}\big)\big(y_i - \bar{y}\big) \Big\rangle \tag{4.19}$$

$$= \frac{1}{N}\sum_i^N \big(x_i - \bar{x}\big)\big(y_i - \bar{y}\big)\,, \tag{4.20}$$

which, for $x = w(\theta)$ and $y = w(\theta')$ gives us

$$\mathsf{Cov}\big(w(\theta), w(\theta')\big) = \Big\langle \big[w_i(\theta) - \bar{w}(\theta)\big]\big[w_i(\theta') - \bar{w}(\theta')\big] \Big\rangle \tag{4.21}$$

$$= \frac{1}{N_\theta}\sum_i^N \big[w_i(\theta) - \bar{w}(\theta)\big]\big[w_i(\theta') - \bar{w}(\theta')\big]\,. \tag{4.22}$$

Here, N_θ represents the number of bins that contain the angular information. So, its value is the separation between the central value of each bin, $\Delta\theta$ divided by a scaling factor, which in this case will be the minimum value of θ considered:

$$N_\theta = \frac{\Delta\theta}{\theta_{min}}\,. \tag{4.23}$$

Therefore, the transformation added to expression in equation (4.17) becomes

$$\Gamma_{\theta\theta',ijpq} \longrightarrow \frac{1}{N_\theta}\Gamma_{\theta\theta',ijpq}\,, \tag{4.24}$$

which finally gives us the expression of the Covariance Matrix that we will use in our analysis:

$$\Gamma_{\theta\theta',ijpq} = \frac{1}{N_\theta f_{sky}} \sum_\ell \left\{ \frac{(2\ell+1)}{16\pi^2} P_\ell(\cos\theta) P_\ell(\cos\theta') \left[\left(C_{\ell,iq} + P_{N,iq}\right) \left(C_{\ell,pj} + P_{N,pj}\right) + \left(C_{\ell,ip} + P_{N,ip}\right) \left(C_{\ell,qj} + P_{N,qj}\right) \right] \right\} . \tag{4.25}$$

4.1.4. Parameters to constrain

The goal of this chapter is to investigate how good C_ℓ and $w(\theta)$ are at evaluating certain parameters. Therefore, we need to choose some parameters to analyse.

The structure and energy contents of the Universe, commonly referred to as the Cosmology of the Universe, can be defined by multiple parameters. These are called the **cosmological parameters**. Even though we are not certain about their values, multiple studies have been done to try to constrain them, so, we have an approximate idea of what values they have. In this work we need to define the cosmology of a fiducial Universe, so we used the cosmological parameters results from Planck 2018 collaboration [33] to guide ourselves. With this in mind, the fiducial values used for each cosmological parameter are defined in Table 4.1.

	H_0	w	$\Omega_{cdm,0}$	$\Omega_{b,0}$	h	A_s	n_s
Fiducial values	67.5	-1	0.27	0.045	0.675	2×10^{-9}	0.965

TABLE 4.1: Fiducial values of the Cosmological parameters relevant to our analysis.

Despite the fact that most of the cosmological parameters, in Table 4.1, have already been mentioned (H_0, in eq.(2.14) and the A_S and n_S in eq.(3.8)) a few of them are still new to us. The w represents the relation between pressure and energy density of the Universe, when these is approximated to a perfect fluid, $w \equiv p/\rho$. Moreover, the previously mentioned density parameter for matter, the $\Omega_{m,0}$ in eq.(2.14), can be decomposed into baryonic matter, $\Omega_{b,0}$, and cold dark matter, $\Omega_{cdm,0}$, through $\Omega_{m,0} = \Omega_{cdm,0} + \Omega_{b,0}$. Furthermore, h is the dimensionless Hubble constant and is used to simplify calculations since one can use h instead of H_0, by using the relation $H_0 = 100h\,[\text{km s}^{-1}\,\text{Mpc}^{-1}]$.

In addition to these parameters, we will also consider the astrophysical parameter, **magnification bias**, $s(z)$, and **parameter** β (to recall their explanation see eq.(2.49) and

section 2.4.3, respectively). The fiducial value of β for the ΛCDM model is already known to be 1. For the fiducial value of the magnification bias, $s(z)$ we'll use the following equation [21]:

$$s(z) = s_0 + s_1 z + s_2 z^2 + s_3 z^3 \,, \tag{4.26}$$

in which the coefficients used are $s_0 = 0.1194$, $s_1 = 0.2122$, $s_2 = -0.0671$, $s_3 = 0.1031$ [21].

Therefore, the vector that contains all the parameters that we will analyse is

$$\vec{\lambda} = \{s(z), \beta, H_0, w, \Omega_b, \Omega_{cdm}, A_s, n_s\} \,.$$

4.1.5. Derivatives of LSS estimators

There are two ways to tackle this problem: analytic derivatives or numerical ones. Analytic derivatives are faster computationally, but they require a useful analytic expression, which doesn't always exist. On the other hand, numerical derivatives take longer to compute but we don't need to know their expression. So, the use of analytic or numerical derivatives will depend on the parameter, λ_μ (see section 4.1.4), that we are considering.

To begin we will analyse the **derivatives of** C_ℓ. For this observable we'll use analytic derivatives for parameters β and A_s; the derivatives of all the other parameters will be solved numerically:

- **Analytic derivative of** $\partial C_{\ell,ij}/\partial\beta$:

$$\frac{\partial C_{\ell,ij}}{\partial \beta} = C^G_{\ell,ij} + C^L_{\ell,ij} - C^{D+RSD}_{\ell,ij} \,. \tag{4.27}$$

  To see the detailed derivation and explanation of this equation check appendix B.1.

- **Analytic derivative of** $\partial C_{\ell,ij}/\partial A_s$:

$$\frac{\partial C_{\ell,ij}}{\partial A_s} = \frac{C_{\ell,ij}}{A_s} \,, \tag{4.28}$$

  see derivation in appendix B.2.

- **Numerical derivatives**:

  To calculate the numerical derivatives we will use the five-point stencil approximation

$$f'(x) \approx \frac{-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)}{12h} \,. \tag{4.29}$$

Applying it to the $C_{\ell,ij}$ derivatives we obtain

$$\frac{\partial C_{\ell,ij}}{\partial \lambda_\mu} \approx \frac{-C_{\ell,ij}(\bar{\lambda}_\mu + 2h_\mu) + 8C_{\ell,ij}(\bar{\lambda}_\mu + h_\mu) - 8C_{\ell,ij}(\bar{\lambda}_\mu - h_\mu) + C_{\ell,ij}(\bar{\lambda}_\mu - 2h_\mu)}{12h_\mu} .$$

Here, we denoted $\bar{\lambda}_\mu$ as the fiducial value of the cosmological parameter that we are considering, and h_μ as the space between points. As a rule of thumb, the value of h_μ, for a specific cosmological parameter μ, is given by $h_\mu \simeq 0,01\ \bar{\lambda}_\mu$.

Moving on to the **derivatives of the $w(\theta)$ estimator**, we can take advantage of equation (3.12) to find that

$$\begin{aligned} \frac{\partial w_{ij}(\theta)}{\partial \lambda_\mu} &= \frac{\partial}{\partial \lambda_\mu} \sum_{\ell \geq 0}^{\ell_{max}} \left(\frac{2\ell+1}{4\pi} P_\ell(\cos\theta) C_{\ell,\,ij} \right) \\ &= \sum_\ell \left(\frac{2\ell+1}{4\pi} P_\ell(\cos\theta)\ \frac{\partial C_{\ell,ij}}{\partial \lambda_\mu}\bigg|_{\lambda_\mu = \bar{\lambda}_\mu} \right) . \end{aligned} \tag{4.30}$$

Therefore, in order to obtain the values of $\partial w_{ij}(\theta)/\partial\lambda_\mu$ we can just use the values of $\partial C_{\ell,ij}/\partial\lambda_\mu$, which we already explored this in section.

4.2. Comparing the performance of both estimators

This section details the numerical procedure for comparing the performance of the C_ℓ and $w(\theta)$ estimators. We implement the theoretical expressions from the previous sections in a Python code to compute Fisher matrix forecasts. The methodology is as follows: first, we outline the computational steps for constructing the Fisher matrices; second, we define the specific parameter sets and survey configurations used in the analysis; third, we present the resulting forecasts.

4.2.1. Computational Procedure for Fisher Matrices

We begin by defining the galaxy distribution for our analysis. We adopt the DES-like distribution described in section 3.2.1, with the caveat that we will only consider the first four redshift bins:

$$z_1 = 0.585, \quad z_2 = 0.66, \quad z_3 = 0.755, \quad z_4 = 0.89\,.$$

In order to streamline the computation, we will redefine our data as

$$\mathbf{Y} = \vec{C}_{\ell,I}\,, \qquad \text{for every value of } I \text{ and } \ell\,, \tag{4.31}$$

$$\mathbf{Y}' = \vec{w}_I(\theta)\,, \quad \text{for every value of } I \text{ and } \theta\,, \tag{4.32}$$

where the index I denotes a specific combination of redshift bins. In this analysis we will consider two sets of bin combinations:

- **Autocorrelations**: $I_{\text{auto}} = \{z_{11}, z_{22}, z_{33}, z_{44}\}$,
- **Full-correlations** (auto + cross): $I_{\text{full}} = \{z_{11}, z_{12}, z_{13}, z_{14}, z_{22}, z_{23}, z_{24}, z_{33}, z_{34}, z_{44}\}$,

where we used the notation z_{ij} to represent the correlation of redshift bin z_i with bin z_j.

So, for a specific bin combination, I, and a set of initial parameters, λ_μ, the Fisher matrix is calculated as follows:

1. **Covariance Matrices:** Compute the covariance matrices for $\vec{C}_\ell$ and $\vec{w}(\theta)$ with dimensions $I\ell \times I\ell$ and $I\theta \times I\theta$, respectively, following the methods in Sections 4.1.2 and 4.1.3.
2. **Parameter Derivatives:** Calculate the derivatives of the data vectors, $\partial\vec{C}_\ell/\partial\lambda_\mu$ and $\partial\vec{w}(\theta)/\partial\lambda_\mu$ as outlined in Section 4.1.1. These vectors have dimensions $I\ell$ and $I\theta$, respectively.
3. **Matrix Inversion:** Invert the covariance matrices, $\Gamma \to \Gamma^{-1}$, and transpose the derivatives as $\bar{\mathbf{X}}_{,\mu} \to \bar{\mathbf{X}}^T_{,\mu}$.
4. **Fisher Matrix Construction:** Finally, assemble the Fisher matrices for each estimator using equations (4.8) and (4.9).

4.2.2. Survey Configurations and Parameter Sets

A theoretical analysis of the sky has access to complete information, including full-sky coverage and angular correlations across all scales. However, observational studies face significant limitations in data accessibility. To systematically investigate these constraints, we define 6 survey configurations, each targeting specific aspects of observational limitations. Furthermore, due to parameter degeneracies, we analyse the bias parameter $s(z)$ and parameter β in separate sets.

Our analysis begins with a Baseline configuration representing an ideal scenario: full-sky coverage ($f_{\text{sky}} = 1$) with complete angular scale access ($\ell \in [2, 1000]$ and $\theta \in [0.18°, 90°]$). To approach realistic survey conditions, we then consider reduced angular ranges. Model 1 adopts $\theta \in [0.18°, 18°]$ (equivalent to $\ell \in [10, 1000]$), consistent with a DES-like survey coverage [30]. Since DES covers approximately 1000 to 5000 deg^2 ($\theta_{\text{max}} \in [31.6°, 70.7°]$), we must also account for a limited sky coverage.

The sky fraction f_{sky} corresponding to a maximum angular scale θ_{max} is derived as follows:

$$\text{Total Sky Area} = 4\pi r^2$$

$$\text{Partial Sky Area} = 2\pi r^2(1 - \cos\theta_{\text{max}})$$

$$f_{\text{sky}}(\theta_{\text{max}}) = \frac{2\pi r^2(1 - \cos\theta_{\text{max}})}{4\pi r^2} \tag{4.33}$$

$$\therefore \quad f_{\text{sky}}(\theta_{\text{max}}) = \frac{1}{2}(1 - \cos\theta_{\text{max}}) \tag{4.34}$$

Thus, according to this relation, $\theta_{max} = 71^\circ$ corresponds to a sky fraction of $f_{\text{sky}} \approx 0.33$, which we will adopt for configurations with reduced sky coverage.

Furthermore, we examine two distinct scenarios for scale treatment between the C_ℓ and $w(\theta)$ estimators:

1. **Coupled ℓ and θ**: The multipole and angular ranges are linked by the relation $\theta \approx 180/\ell$ (Equation 3.17), representing the ideal case where both estimators probe identical physical scales.

2. **Decoupled ℓ and θ**: The ℓ-range is reduced while maintaining the original θ-range, breaking the direct coupling. This addresses the challenge that small-ℓ measurements are particularly susceptible to survey masks and systematic foreground errors, as these modes require information distributed across large sky areas. Consequently, we expect C_ℓ to be more sensitive to masking effects than $w(\theta)$.

Then, Table 4.2 details the six configurations used in our Fisher forecasts. The Baseline model serves as our ideal reference scenario, while subsequent models systematically introduce observational limitations:

- **Model 1**: Adopts a realistic maximum angular scale ($\theta_{\text{max}} = 18^\circ$) comparable to DES-like surveys while maintaining full sky coverage.

- **Model 2**: Incorporates both the reduced angular scale of Model 1 and limited sky coverage ($f_{\text{sky}} = 0.33$).

- **Model 3**: Returns to full sky coverage but introduces a low-multipole cut-off at $\ell_{\text{min}} = 20$, removing information from the largest angular scales in the C_ℓ estimator. This represents the challenge of measuring small-ℓ modes, which require full-sky information that is compromised by observational masks.

- **Model 4**: as the same goal as Model 3, but implements a more restrictive low-multipole cut-off at $\ell_{\text{min}} = 30$.
- **Model 5**: Combines the limitations of Models 2 and 4, featuring both reduced sky coverage and the high ℓ_{min} cut-off, representing our most realistic scenario.

Therefore, the progression from Baseline to Model 5 thus transitions from an ideal, full-information scenario to our most observationally constrained case.

	Coupling	θ **Range**	ℓ **Range**	f_{sky}
Baseline	Coupled	$[0.18°, 90°]$	[2, 1000]	1
Model 1	Coupled	$[0.18°, 18°]$	[10, 1000]	1
Model 2	Coupled	$[0.18°, 18°]$	[10, 1000]	0.33
Model 3	Decoupled	$[0.18°, 18°]$	[20, 1000]	1
Model 4	Decoupled	$[0.18°, 18°]$	[30, 1000]	1
Model 5	Decoupled	$[0.18°, 18°]$	[30, 1000]	0.33

TABLE 4.2: Survey configurations used to test the performance of the C_ℓ and $w(\theta)$ estimators. Valid for both parameter sets.

A key challenge in our analysis arises from the strong degeneracy between parameters $s(z)$ and β. Including both in a single Fisher matrix would result in a singular (non-invertible) matrix, preventing the use of equation (4.10). To circumvent this issue, we analyse them in separate parameter sets:

- **Set** $s(z)$: $\lambda = \{s(z_1), s(z_2), s(z_3), s(z_4), H_0, w, \Omega_b, \Omega_{\text{cdm}}, A_s, n_s\}$
- **Set** β: $\lambda = \{\beta, H_0, w, \Omega_b, \Omega_{\text{cdm}}, A_s, n_s\}$

All survey configurations in Table 4.2 are applicable to both parameter sets.

4.2.3. Visualising results obtained

To better interpret the Fisher matrix results, we visualise the forecasted constraints using error ellipses and histograms of the relative uncertainties, rather than just listing the numerical values of $\sigma(\lambda_\mu)$, see equation (4.10).

First, we present the baseline model's constraints, which represent the combination of parameters that demonstrate the greatest performance in comparison to all other succeeding models. Figures 4.1, 4.2, and 4.9 show the corner plots with the error ellipses for each parameter set.

The relative uncertainty for a parameter λ_μ is defined as

$$\frac{\sigma(\lambda_\mu)}{\lambda_\mu} \times 100\% \,, \tag{4.35}$$

where $\sigma(\lambda_\mu)$ is the marginalised error from the Fisher matrix (see eq. 4.10). The relative uncertainties for all parameters in the baseline model are compared in the histograms in Figures 4.3 and 4.10.

To quantify how the constraints degrade in other models relative to this baseline, we calculate the fractional increase in relative uncertainty:

$$\left[\frac{\sigma(\lambda_\mu)}{\lambda_\mu} - \frac{\sigma(\lambda_{\mu,\text{Base}})}{\lambda_{\mu,\text{Base}}}\right] \times 100\% \,. \tag{4.36}$$

Here, a value of 0 indicates performance identical to the baseline, while a positive value corresponds to a percent increase in the parameter's relative uncertainty.

The performance comparisons for the Set $s(z)$ models are shown in Figures 4.4 to 4.8. The corresponding comparisons for Set β are shown in Figures 4.11 to 4.15.

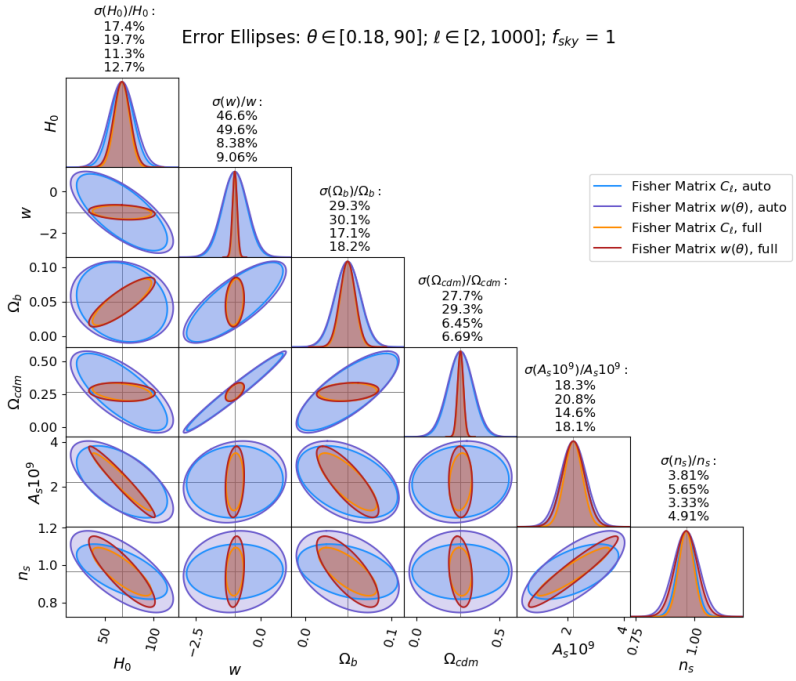


FIGURE 4.1: Corner plot showing the Fisher forecast constraints (error ellipses) for the cosmological parameters in the baseline model, calculated along the $s(z)$ parameters.

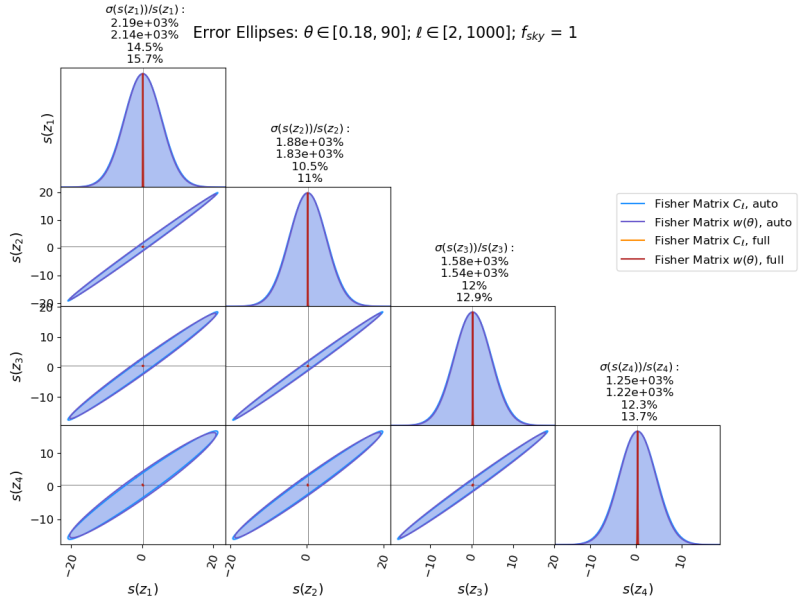


FIGURE 4.2: Corner plot showing the Fisher forecast constraints for the parameters $s(z_1)$, $s(z_2)$, $s(z_3)$, $s(z_4)$ in the baseline model, calculated along the cosmological parameters.

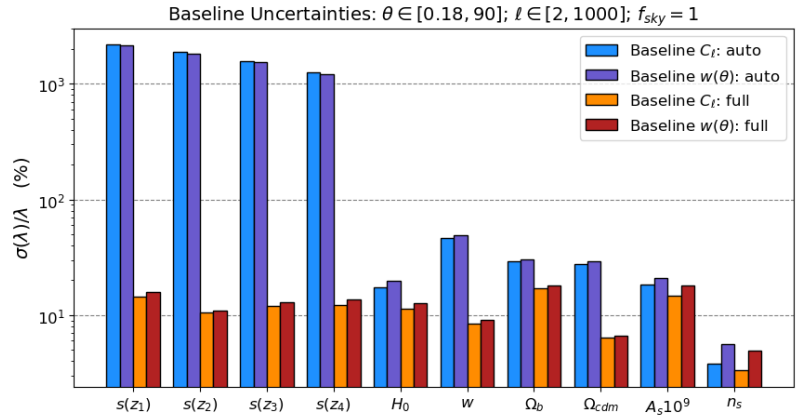


FIGURE 4.3: Relative uncertainties ($\sigma(\lambda_\mu)/\lambda_\mu \times 100\%$) for all parameters in Set $s(z)$ for the Baseline model.

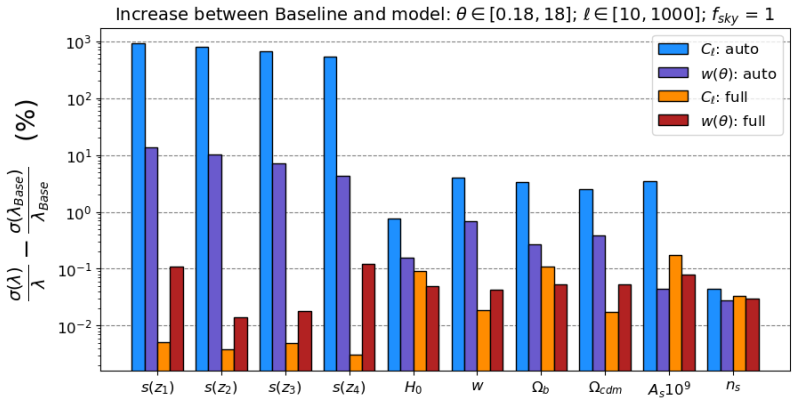


FIGURE 4.4: Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set $s(z)$ when moving from the Baseline model to Model 1 with $f_{sky} = 1$, $\theta_{\text{max}} = 18°$, and $\ell_{\text{min}} = 10$.

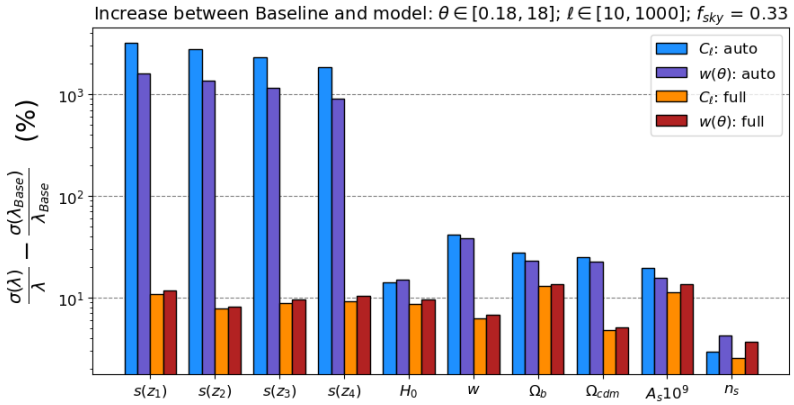


FIGURE 4.5: Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set $s(z)$ when moving from the Baseline model to Model 2 with $f_{sky} = 0.33$, $\theta_{\text{max}} = 18°$, and $\ell_{\text{min}} = 10$.

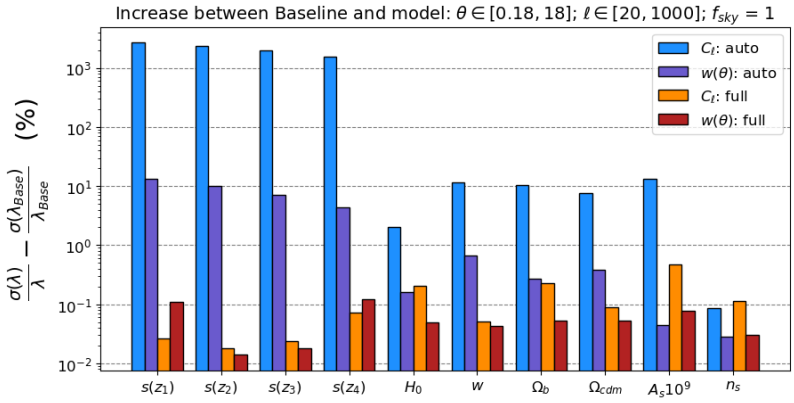


FIGURE 4.6: Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set $s(z)$ when moving from the Baseline model to Model 3 with $f_{sky} = 1$, $\theta_{\text{max}} = 18°$, and $\ell_{\text{min}} = 20$.

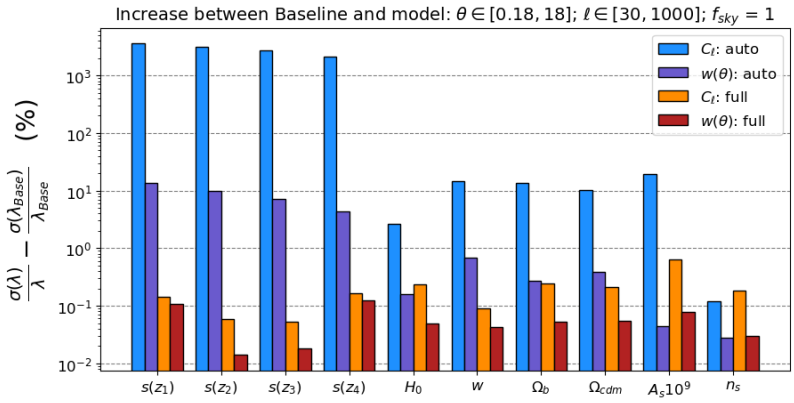


FIGURE 4.7: Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set $s(z)$ when moving from the Baseline model to Model 4 with $f_{sky} = 1$, $\theta_{\text{max}} = 18°$, and $\ell_{\text{min}} = 30$.

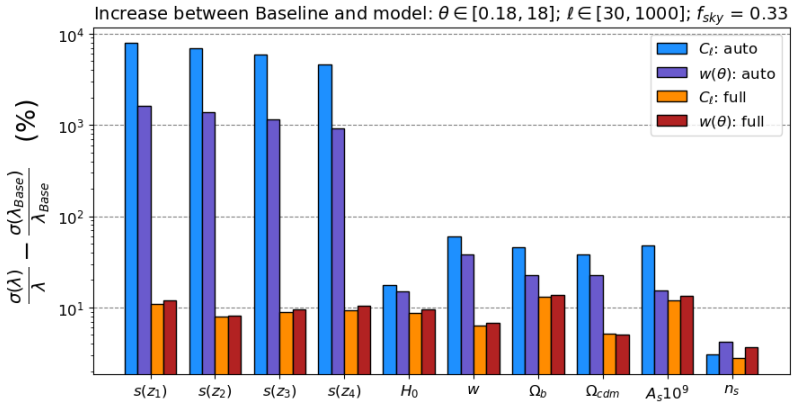


FIGURE 4.8: Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set $s(z)$ when moving from the Baseline model to Model 5 with $f_{sky} = 0.33$, $\theta_{\text{max}} = 18°$, and $\ell_{\text{min}} = 30$.

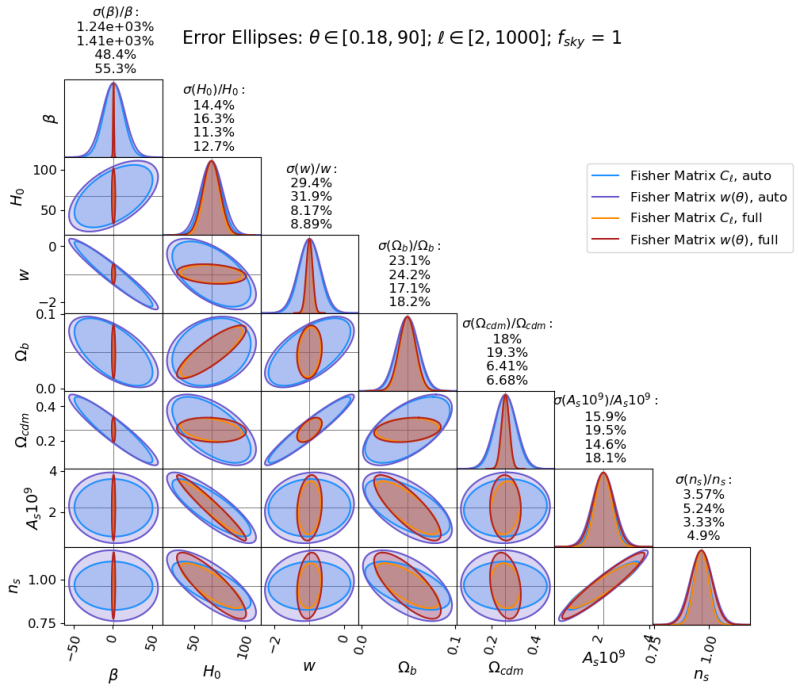


FIGURE 4.9: Corner plot showing the Fisher forecast constraints (error ellipses) for parameter β and the cosmological parameters, in the baseline model.

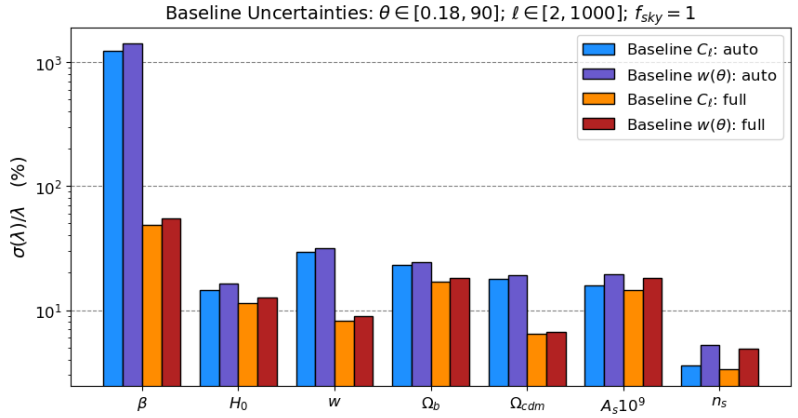


FIGURE 4.10: Relative uncertainties ($\sigma(\lambda_\mu)/\lambda_\mu \times 100\%$) for all parameters in Set β for the Baseline model.

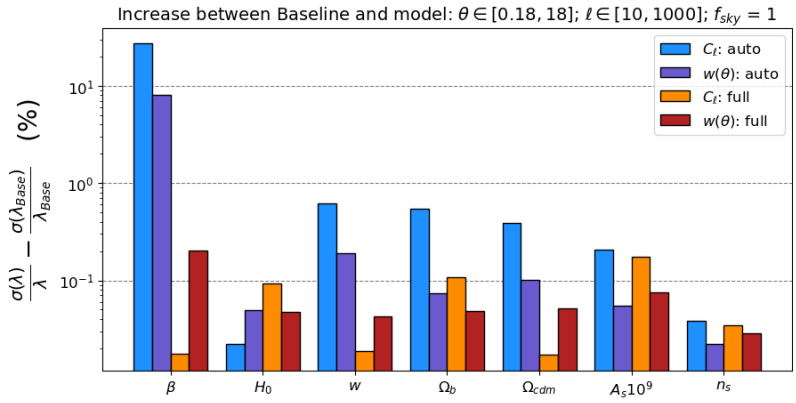


FIGURE 4.11: Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set β when moving from the Baseline model to Model 1 with $f_{sky} = 1$, $\theta_{\text{max}} = 18°$, and $\ell_{\text{min}} = 10$.

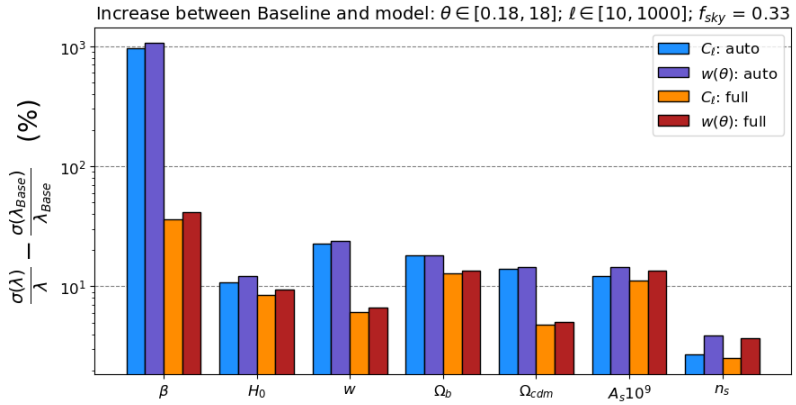


FIGURE 4.12: Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set β when moving from the Baseline model to Model 2 with $f_{sky} = 0.33$, $\theta_{\text{max}} = 18°$, and $\ell_{\text{min}} = 10$.

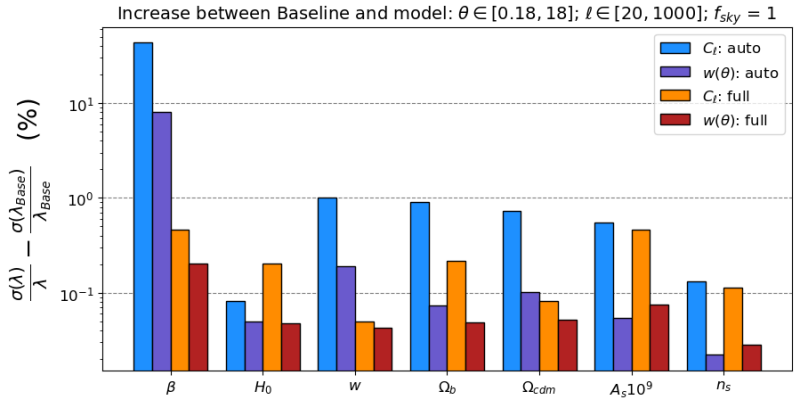


FIGURE 4.13: Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set β when moving from the Baseline model to Model 3 with $f_{sky} = 1$, $\theta_{\text{max}} = 18°$, and $\ell_{\text{min}} = 20$.

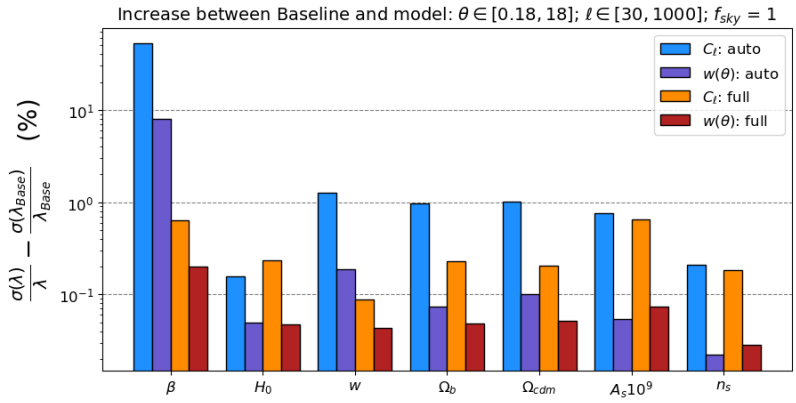


FIGURE 4.14: Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set β when moving from the Baseline model to Model 4 with $f_{sky} = 1$, $\theta_{\text{max}} = 18°$, and $\ell_{\text{min}} = 30$.

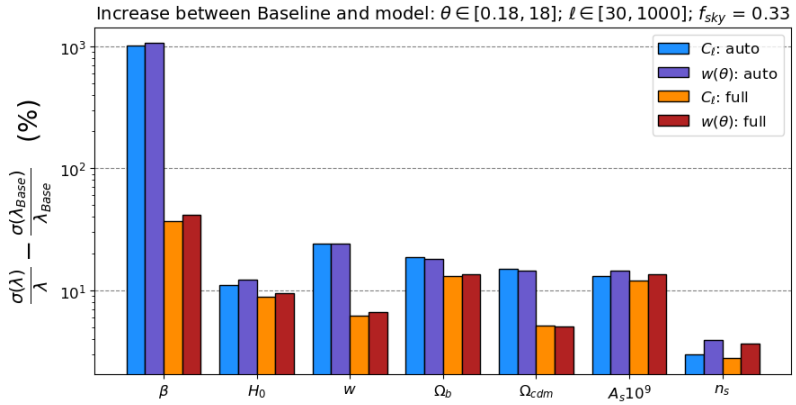


FIGURE 4.15: Absolute increase in relative uncertainty, eq.(4.36), for parameters in Set β when moving from the Baseline model to Model 5 with $f_{sky} = 0.33$, $\theta_{\text{max}} = 18°$, and $\ell_{\text{min}} = 30$.

4.3. Discussion

In this chapter we introduced Fisher Matrix forecasts along with the necessary theory to calculate them. Then, the aforementioned framework was used to compute Fisher forecasts, with the objective of comparing the performance of C_ℓ and $w(\theta)$ when measuring multiple parameters. The following discussion will concern the results obtained and the main conclusions to be taken from them.

Overall, the results obtained in the previous section show that both parameter sets (set $s(z)$ and set β) are consistent across each model, thereby indicating robust trends within those models. Moreover, for all the models studied, parameters $s(z)$ and β exhibit the **best results for the full-correlations case** (autocorrelations and cross-correlations),

since this case's relative uncertainties are smaller than the ones obtained with autocorrelations. These results demonstrate the connection between parameters $s(z)$ and β, and magnification lensing. This is due to the fact that the cross-correlations between widely separated redshift bins are dominated by the lensing signal (see section 3.3).

A comparison with Montanari and Durrer [21] via corner plots (Figures 4.1, 4.2 and 4.9) indicates their constraints are more precise, likely because they employed a higher number of redshift bins (5 and 10, compared to our 4). This suggests a potential avenue for improving our results.

As expected, the **best results are the Baselines**, see Figures 4.3 and 4.10. In this particular model, the most effective approach for measuring both $s(z)$ and β is through the full-correlations of C_ℓ. Nevertheless, this method exhibits only a marginal enhancement when compared to $w(\theta)$ in full-correlations.

The following analysis is applicable to both sets of parameters and will focus on the comparison of the Baseline model with the other five models.

- **Angular Range (Baseline vs. Model 1):** Reducing the angular range from the Baseline (Figures 4.3 and 4.10) to Model 1 (Figures 4.4 and 4.11) has the greatest negative impact on the C_ℓ auto-correlation, and the least on the C_ℓ full-correlation, for both $s(z)$ and β.

- **Sky Coverage (Model 2):** Reducing the sky coverage (f_{sky}) increases the uncertainties of measuring any parameters with any estimators by, approximately, an order of magnitude (see Figures 4.5 and 4.12). An increase much greater than the one seen from the reduction in angular scale of Model 1. This reveals the importance of a large sky coverage.

- **Angular Multipole Range (Models 1, 3, and 4):** As we transition form Model 1 to Model 3 and then 4, the angular multipole range is progressively restricted (removing low-ℓs), making the uncertainty of the C_ℓ estimators increases significantly. In Model 4 (Figures 4.7 and 4.14), this uncertainty increase as reached the point for which all parameters have higher uncertainties with C_ℓ than with $w(\theta)$.

- **Realistic Scenario (Model 5):** This model (see Figures 4.8 and 4.15), which combines a limited angular scale ($\theta_{\text{max}} = 18°$), a restricted multipole range ($\ell_{\text{min}} = 30$) with reduced sky coverage ($f_{\text{sky}} = 0.33$), is our most realistic case. In this scenario,

the performance of the full-correlation C_ℓ and $w(\theta)$ estimators becomes nearly identical, with C_ℓ slightly better. This finding contradicts the trend observed in Models 1, 3 and 4. Notably, Model 4 shows $w(\theta)$ outperforming C_ℓ. The key factor to be considered is the loss of large-scale information due to partial sky coverage. This loss negates the primary advantage $w(\theta)$ enjoyed in Model 4, as it can no longer utilise the large-angles information. Meanwhile, the C_ℓ estimator, which was already deprived of these large scales (deprived of low-ℓ modes) in Model 4, exhibits a smaller performance decline. The net effect is that the two estimators converge in performance under these realistic observational constraints, due to the fact that, even though their angular range is decoupled, in practice they have access to the same amount of information.

Comparing the ideal Baseline model with the realistic Model 5 reveals a significant increase in uncertainty. For $s(z)$, the relative uncertainty, using full-correlations, increases from approximately 15% (Baseline) to 25% (Model 5). For β, the uncertainty increases more drastically, from about 50% (Baseline) to 100% (Model 5). This implies that, while a moderately precise measurement of $s(z)$ might be feasible, accurately constraining β under realistic survey conditions is not attainable due to its high uncertainty. Two primary strategies can mitigate this issue:

1. **Increase the number of redshift bins,** as previously noted, to extract more information from the cross-correlations.

2. **Increase the sky coverage (f_{sky}).** This is the most critical factor. The results from Model 4 demonstrate that with full sky coverage ($f_{\text{sky}} = 1$), even with a limited angular range, the increase in uncertainty for both $s(z)$ and β is below 1% compared to the Baseline. Additionally, under these conditions, the $w(\theta)$ estimator would outperform C_ℓ.

In conclusion, the optimal LSS estimator doesn't exist. The choice between the C_ℓ and $w(\theta)$ estimators depends critically on the specific survey configuration, particularly the sky coverage and the available angular scales. Additionally, our findings indicate that the sky fraction (f_{sky}) has a substantially greater impact on estimator performance than the specific angular range.

Furthermore, accurately constraining parameters like $s(z)$ or β requires the inclusion of cross-correlations between redshift bins. Without these cross-correlations, the analysis lacks access to the lensing signal, which is essential for probing these parameters.

5. Estimating the effect of Weak Lensing Magnification from Simulations

Previously we've studied the behaviour of LSS estimators when the observed data is affected by magnification lensing (see Chapter 3).

Now that we know how to determine the profiles of C_ℓ and $w(\theta)$ under these conditions, we will go further and test if the lensing signal can actually be differentiated from the other components when we use realistic data. To achieve this we'll start by generating a simulation of galaxies distributed in a specified redshift range (section 5.1). Then, we will modify this simulated data according to the restrictions we encounter when making real observations (section 5.2). Next, we'll use this simulation of the observed data to measure galaxy clustering, both with C_ℓ (section 5.3.1) and $w(\theta)$ (section 5.3.2). To conclude we'll discuss the meaning of the results obtained and their usefulness (section 5.4).

5.1. Generate Galaxy Catalogue Simulations

The first step in our analysis is to generate a realistic three-dimensional distribution of galaxies. For this purpose, we use the `CoLoRe` package [3], which is designed to efficiently produce synthetic realisations of cosmological surveys. Hence, this section describes the configuration and inputs used to generate a single galaxy survey realisation.

The `CoLoRe` program generates matter density fluctuations within a cubic volume with side length L_{box}. The size of the box must be large enough to contain a sphere of comoving radius $r(z_{\text{max}})$, where z_{max} is the maximum redshift of the simulation. Following Ramírez-Pérez et al. [3], we set $L_{\text{box}} \simeq 2r(z_{\text{max}})$. For our simulation we adopt $z_{\text{max}} = 2.0$.

Furthermore, the spatial resolution is determined by N_{grid}, the number of grid cells along each dimension. The resulting cell size is $\Delta x = L_{\text{box}}/N_{\text{grid}}$. We choose $N_{\text{grid}} = 2^9$ (512), implying that the box will be divided into $N_{grid}^3 \simeq 1.3 \times 10^7$ cells.

The underlying LSS, or the "cosmic web," is defined by fiducial cosmology and the linear matter power spectrum, $P(k)$, considered. We adopt the fiducial cosmology described

in section 4.1.4, and compute $P(k)$ for wavenumbers $k \in [10^{-5}, 10]$ Mpc^{-1} using the `PyCCL` package. This power spectrum serves as a key input for `CoLoRe` to generate the Gaussian random field representing the initial density fluctuations.

Having given `CoLoRe` the needed information to construct the background Universe, the next step is to populate it with galaxies. To achieve this, `CoLoRe` requires a model for the galaxy redshift distribution, $dN/dz/d\Omega$, and their bias, $b(z)$. Since an observed distribution (such as the DES-like distribution from section 3.2.1) is itself biased, we instead generate an intrinsic distribution from a Luminosity function.

We begin with an r'-band Schechter Luminosity Function given by Gabasch et al. [40] and parametrised in terms of absolute magnitude M [21],

$$\phi(M,z)\,dM = \frac{\ln(10)}{2,5}\phi_\star(z)\left(10^{(M_\star(z)-M)/2.5}\right)^{\alpha+1}\exp\left[-10^{(M_\star(z)-M)/2.5}\right]dM\,, \tag{5.1}$$

where $M_\star(z) = -21.49 - 1.25\log(1+z)$ is a characteristic magnitude; $\phi_\star(z) = (2.59 - 0.136z - 0.081z^2)10^{-3}\left[\text{Mpc}^{-3}\right]$ is a characteristic number density (number of galaxies per volume element) and $\alpha = -1.33$ is a constant faint-end slope, that describes the abundance of high-M galaxies [30]. In order to mimic the specifications of the DES survey we consider the approximation of the previously mentioned r'-band luminosity function to an r-band one, $r' \simeq r$ [30]. Therefore, the number density of galaxies with smaller magnitude than M_{lim} (brighter than $M_{lim}(z)$), is given by the Cumulative Luminosity function

$$n_L(M \leq M_{lim}, z) = \int_{-\infty}^{M_{lim}(z)} \phi(M,z)dM\,. \tag{5.2}$$

However, since we don't observe the absolute magnitude, M, we write it in relation to the apparent magnitude in the observer-frame band, m, through the following equation: [41]

$$M(m,z) = m - 25 - 5\log_{10}\frac{d_L(z)}{\text{Mpc}} - k(z)\,, \tag{5.3}$$

in which the d_L is the luminosity distance and the k-correction considered is $k(z) \simeq 1.5z$ [41]. Following this procedure we will set the magnitude limit of our simulation to $m_{\text{lim,sim}} = 24.5$. To convert this comoving number density, with units $\left[n_L(z)\right] = [\text{galaxies/Mpc}^3]$, into an observable distribution per unit redshift and solid angle, we consider the volume element $dV = r^2(z)(dr/dz)dzd\Omega$, where $r(z)$ is the comoving distance (see eq.(2.13)) and $dr/dz = c/H(z)$ its derivative. Then, considering that $n_L(z) = dN/dV$, we can obtain a

distribution of galaxies based on the Luminosity function by doing the following:

$$\begin{aligned} dN &= n_L(M,z)dV \\ &= n_L(M,z)r^2(z)\frac{dr}{dz}dzd\Omega \\ &= n_L(M,z)r^2(z)\frac{c}{H(z)}dzd\Omega\,, \end{aligned} \tag{5.4}$$

$$\frac{dN}{d\Omega dz} = n_L(M,z)r^2(z)\frac{c}{H(z)}\,. \tag{5.5}$$

Finally, we convert the solid angle from steradians to square degrees for compatibility with `CoLoRe`. So, equation (5.5) becomes

$$\frac{dN}{d\Omega dz} = \left(\frac{\pi}{180}\right)^2 n_L(M,z)r^2(z)\frac{c}{H(z)}\,. \tag{5.6}$$

The galaxy distribution in equation (5.6) is the one we will give to `CoLoRe` as input, and it can be visualised in Figure 5.1.

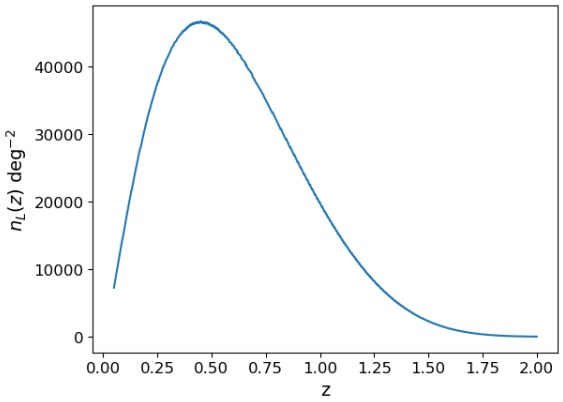


FIGURE 5.1: Distribution of galaxies, per redshift, that was generated through the Luminosity function in equation (5.1)

In conjunction with this distribution, we provide `CoLoRe` with the linear galaxy bias model, $b(z)$, from equation (3.16). This parameter dictates how galaxies trace the underlying matter distribution.

With all inputs defined, we configure `CoLoRe` to generate sources according to the specified galaxy distribution and bias. We also include the effects of magnification lensing from a gravitational lens positioned at the mean redshift of our catalogue. Executing `CoLoRe` with this configuration produces the simulated galaxy catalogue that we will use in the following sections of this chapter.

5.2. Create Observed Galaxy Catalogue

The simulated galaxy catalogue generated by `CoLoRe` represents a "perfect" observation of the Universe. In order to create a synthetic observational dataset, we must apply several observational effects that remove information from this "ideal" data. This section describes the process of transforming the "True" catalogue into an "Observed" catalogue by incorporating photometric redshift errors, redshift-space distortions (RSD), magnification lensing effects, and survey selection slices.

The comoving distance to a galaxy is determined by its cosmological redshift, z_{True}. However, observational surveys that measure redshifts with photometry have Gaussian uncertainties associated with them (see section 3.2.1). Furthermore, the observed redshift is perturbed by the peculiar velocities of each galaxy, the previously mentioned RSD effect (see section 2.4.2). Therefore, we model the probability of measuring a photometric redshift z_{phot} for a galaxy with true redshift z_{True} as a Gaussian distribution,

$$P(z_{\text{phot}}|z_{\text{True}}) = \frac{1}{\sigma_z\sqrt{2\pi}} \exp\left[-\frac{(z_{\text{phot}} - \bar{z})^2}{2\sigma_z^2}\right] . \tag{5.7}$$

Here, $\bar{z} = z_{\text{True}} + \delta z$; δz is the redshift perturbation due to RSD; and the photometric redshift error is given by $\sigma_z = 0.05(1 + z_{\text{True}})$.

For each galaxy in the original catalogue, we draw a random observed photometric redshift, z_{phot}, from this probability distribution function (PDF). Figure 5.2 illustrates the result of this transformation, showing the scatter between z_{True} and z_{phot}.

Weak gravitational lensing, particularly magnification, alters observed galaxy properties in two key aspects: it perturbs the observed angular positions and amplifies the observed fluxes [3]. `CoLoRe` provides the necessary deflection angles ($d\text{RA}, d\text{Dec}$) and the convergence (κ) for each galaxy, which we use to apply these effects. `CoLoRe` assigns angular positions, Right Ascension (RA) and Declination (Dec) coordinates, to each galaxy. In an effort to obtain their observed quantities we have to apply the deflection angles to their true positions:

$$\text{RA}_{\text{obs}} = \text{RA}_{\text{True}} + d\text{RA}\,, \tag{5.8}$$

$$\text{Dec}_{\text{obs}} = \text{Dec}_{\text{True}} + d\text{Dec}\,. \tag{5.9}$$

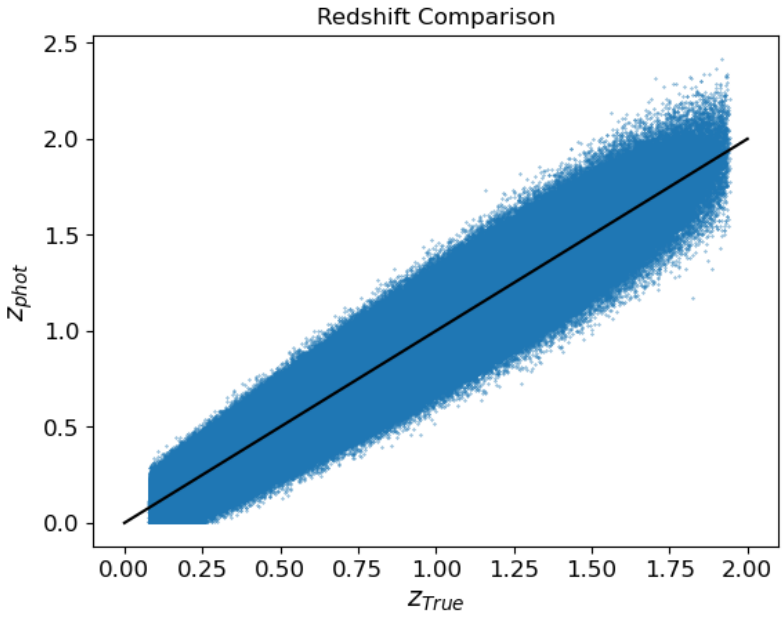


FIGURE 5.2: Comparison between the true cosmological redshift (z_{True}) and the observed photometric redshift (z_{phot}) for galaxies in the simulation. Each point represents a galaxy. The photometric scatter and RSD effects cause z_{phot} to cover a broader range (0.01 to 2.5) than z_{True} (0.05 to 2.0).

The next step is to perturb the flux values of each source. From equation (2.31) and Ramírez-Pérez et al. [3] we know that the observed flux, F_{obs}, is related to the true flux, F_{True}, through the convergence:

$$F_{\text{obs}} = F_{\text{True}}(1 + 2\kappa)\,. \tag{5.10}$$

Despite that, it is not trivial what expression we should use to assign a flux to each source. Consequently, we will avoid this issue by using the apparent magnitude, as expressed in eq.(5.3) Therefore, by employing the relation between flux and magnitude,

$$\begin{aligned} m_1 - m_2 &= -\frac{5}{2}\log_{10}\left(\frac{F_1}{F_2}\right) \\ \implies \frac{F_1}{F_2} &= 10^{-\frac{2}{5}(m_1 - m_2)}\,, \end{aligned}$$

and considering F_2 to be our reference flux, the magnification effect becomes:

$$\begin{aligned}
\frac{F_{\text{obs}}}{F_{\text{ref}}} &= \frac{F_{\text{True}}}{F_{\text{ref}}}(1+2\kappa) \\
F_{\text{obs}} &= F_{\text{True}}(1+2\kappa) \\
10^{-\frac{2}{5}(m_{\text{obs}}-m_\star)} &= 10^{-\frac{2}{5}(m_{\text{True}}-m_\star)}(1+2\kappa) \\
10^{-\frac{2}{5}m_{\text{obs}}} &= 10^{-\frac{2}{5}m_{\text{True}}}(1+2\kappa) \\
-\frac{2}{5}m_{\text{obs}} &= -\frac{2}{5}m_{\text{True}} + \log_{10}(1+2\kappa) \\
m_{\text{obs}} &= m_{\text{True}} - \frac{5}{2}\log_{10}(1+2\kappa)\,.
\end{aligned} \tag{5.11}$$

In order to apply this expression, we must first assign a "True" magnitude, m_{True}, to each galaxy. We draw these magnitudes randomly from the Schechter Luminosity Function (in eq.(5.1)), which, for simplicity, was evaluated at the medium redshift of the catalogue. This is done to ensure that the physically motivated distribution is correlated with the underlying galaxy population. Figure 5.3 shows the PDF used for the sampling of magnitudes.

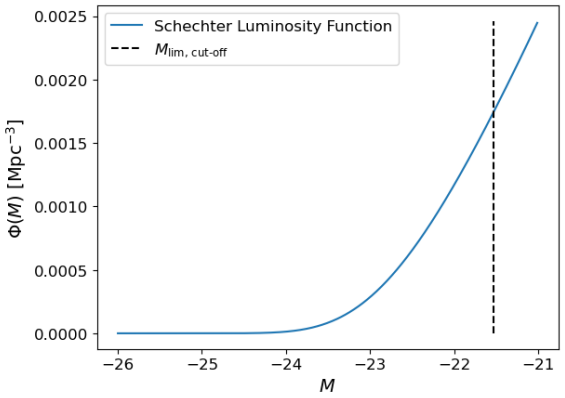


FIGURE 5.3: The Schechter Luminosity Function, which was evaluated at the medium redshift of the catalogue, $z = 0.97$, was used as the probability distribution for assigning "True" magnitudes (m_{True}) to galaxies. $M_{\text{lim, cut-off}}$ is obtained through the use of eq.(5.3) with $m_{\text{lim, cut-off}}$ and represents the magnitude limit of a survey.

Moreover, we are now able to use equation (5.11) to compute the observed magnitude, m_{obs}, for each galaxy.

A real survey has a finite detection limit, and in order to simulate this, we impose a magnitude cut-off, removing all galaxies fainter than the survey's limiting magnitude of $m_{\text{lim, cut-off}} = 24.0$. We choose this value since DES will observe galaxies with magnitude $r < 24$ [30]. Moreover, we simulate galaxies magnitudes up to $m_{\text{lim,sim}} = 24.5$, for two reasons: the first is that we want to guarantee that we are not including edge effects. The second one relates to the fact that when we perturb the m_{True} into m_{obs} their values

change based on κ (see eq.(5.11)) in the pixel, which can be both positive and negative. We therefore want to ensure that any faint galaxy that would be magnified into our sample is simulated. Figure 5.4 shows the histogram of observed magnitudes after this removal. The shape of the distribution is consistent with the input luminosity function (see Figure 5.3), confirming that the magnitude assignment was performed correctly. After applying the magnitude cut-off, we update the RA, Dec, and redshift arrays with the purpose of removing the excluded galaxies.

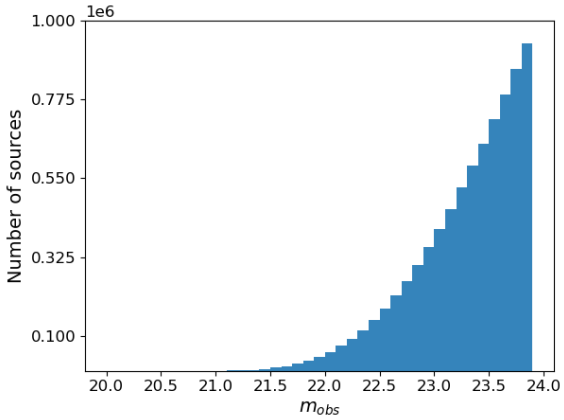


FIGURE 5.4: Distribution of observed magnitudes (m_{obs}) for galaxies in the catalogue after applying the magnitude cut-off at $m_{\text{lim,cut-off}} = 24.0$.

The final step in preparing the catalogue for analysis is to divide the galaxies, from the sphere created by `CoLoRe`, into two shells (two redshift bins). To achieve that, we use the observed photometric redshifts, z_{phot}, as would be done with real data (due to the fact that we wouldn't have access to the z_{True} data). So, for the same set of central redshifts used in section 3.2.1, we select all galaxies satisfying $z_{c,\text{phot}} - \Delta z/2 \leq z_{c,\text{phot}} \leq z_{c,\text{phot}} + \Delta z/2$ with $\Delta z = 0.1$, which is a typical bin width in photometric galaxy surveys. The resulting redshift distribution for each bin is shown in Figure 5.5.

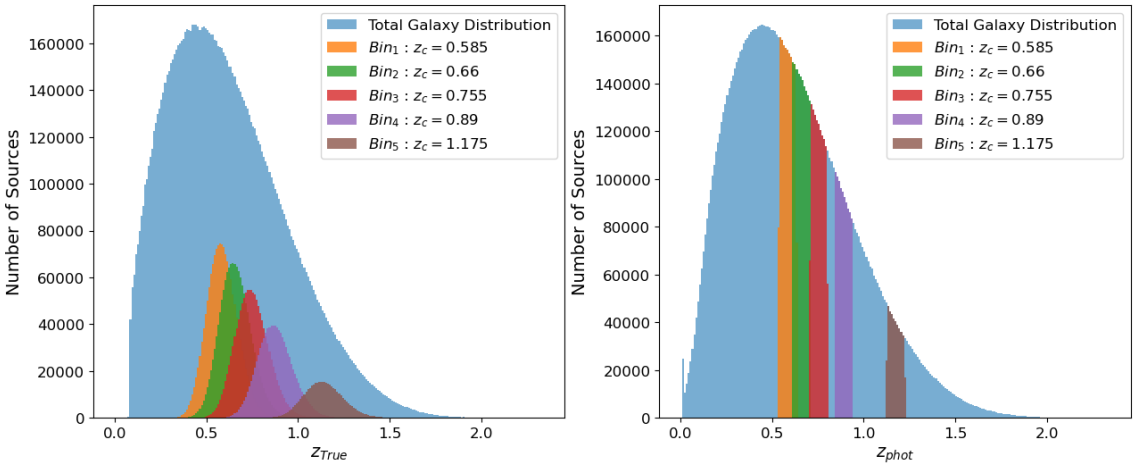


FIGURE 5.5: Redshift distribution of galaxies after selection into photometric redshift bins of width. The bins are defined by slicing the observed z_{phot} values (Right panel). The resultant "True" z-bins have gaussian distributions due to the photometric errors (Left panel).

In order to study the effect of magnification lensing we will consider the two more separated binned catalogues, $z_{c,\text{phot}} = \{0.585, 1.175\}$. These are the final product used in the subsequent correlation function analysis. Also, it's important to mention that, the corresponding angular position data (RA, Dec) for the galaxies in each bin have been saved.

5.3. Measuring Simulated LSS with Clustering Estimators

Once the observed galaxy catalogue has been prepared, we can now quantify the LSS of the simulation by measuring its two-point clustering statistics. The goal is to determine how well we can recover the intrinsic clustering signal from the observationally-biased data. To achieve this we measure clustering through two LSS estimators: C_ℓ, and $w(\theta)$.

5.3.1. Measuring the Angular Power Spectrum

We measure the angular power spectrum using the software package `NaMaster` [4]. `NaMaster` is a framework for estimating pseudo-C_ℓ angular power spectra, which is a fast and nearly optimal method for realistic surveys [4].

To use `NaMaster`, we first project the galaxy catalogue onto a celestial map. Using the `HEALPix` pixelisation scheme [42, 43] via the `healpy` package, we discretise the sphere into $N_{\text{pix}} = 12 \times n_{\text{side}}^2$ pixels (i.e. we transform our angular data, (RA,Dec), into pixels). Likewise, we set the resolution parameter to $n_{\text{side}} = 2^9$, matching the spatial resolution of

the initial simulation (N_{grid}). For each pixel in direction $\hat{\mathbf{n}}$, we calculate the galaxy over-density field, $\delta(\hat{\mathbf{n}})$, defined as

$$\delta(\hat{\mathbf{n}}) = \frac{n(\hat{\mathbf{n}}) - \bar{n}}{\bar{n}}\,, \tag{5.12}$$

where $n(\hat{\mathbf{n}})$ is the number of galaxies in the pixel, and $\bar{n} = N_{\text{gal}}/N_{\text{pix}}$ is the mean number density across the sky. This over-density map, which inherently includes the effects of RSD and lensing applied earlier, is the primary input for the C_ℓ estimation. Furthermore, Figure 5.6 shows the resulting maps for the two redshift bins centred at $z_c = 0.585$ and $z_c = 1.175$.

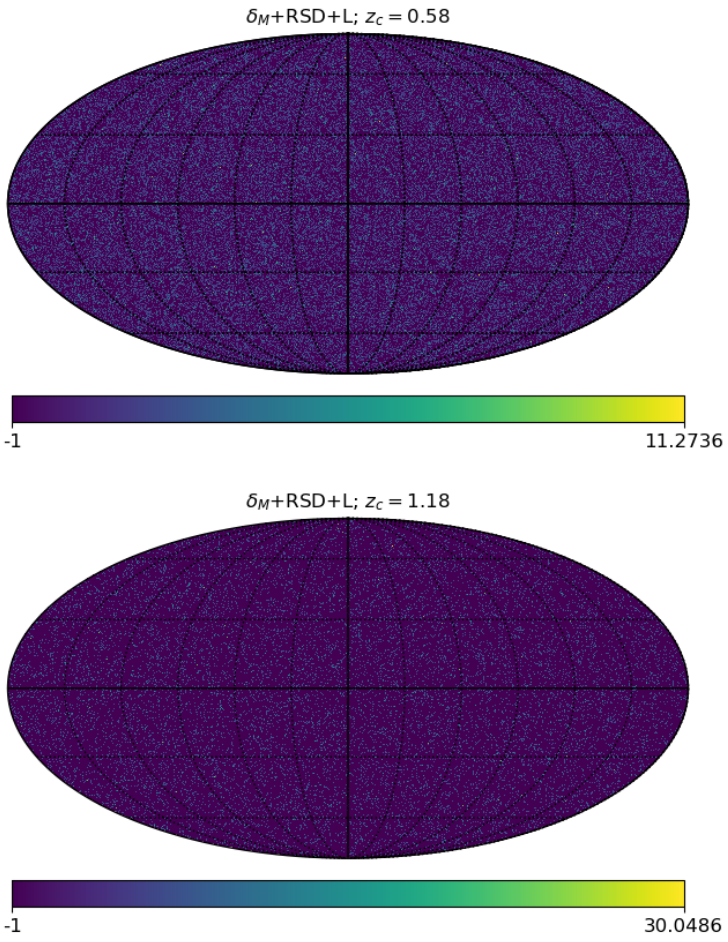


FIGURE 5.6: `HEALPix` maps of the galaxy over-density, $\delta(\hat{\mathbf{n}})$, for two different redshift bins, in which each pixel's colour represents the excess density of galaxies (see eq.(5.12)). The maps include the effects of photometric errors, RSD, and lensing as applied to the observed catalogue.

Once we have the sky maps in a format that healpy can read them, we can give them to `NaMaster`. For that we create a mask for each map, which will be 1 in every pixel since

we are considering the entire sky, and then we create a `NaMaster` field with our mask and our sky map.

For a full-sky observation the angular power spectrum between two maps, a_i and a_j, can be estimated from their spherical harmonic coefficients [44]

$$\tilde{C}_{\ell,ij} = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} a_{\ell m,i} a^*_{\ell' m',j}\,, \quad \text{with} \quad a_{\ell m,i} = \int \delta_i(\hat{n}) Y^*_{\ell m}(\hat{n}) d\Omega\,. \tag{5.13}$$

However, real-world complications like partial sky coverage and masks couple different multipole modes. The pseudo-C_ℓ method implemented in `NaMaster` accounts for this. It first computes the power spectrum of the masked (or "pixel-weighted") maps, known as the pseudo-C_ℓ, $\tilde{C}_\ell$. The relationship between the pseudo-spectrum and the underlying true spectrum is given by a coupling matrix, M, [4, 44]

$$\vec{\tilde{C}}_\ell(a_{\ell m,i}, a_{\ell' m',j}) = \sum_{\ell'} M^{s_a s_b}_{\ell\ell'} \vec{C}_{\ell,ij}\,, \tag{5.14}$$

where $M^{s_a s_b}_{\ell\ell'}$ is a matrix that encodes how the mask couples the power from a true multipole ℓ into the observed multipole ℓ'. Additionally, the s_a, s_b represent the spin of the fields/sky maps considered; and since we are only interested in pure spin-0 fields $M^{s_a s_b}_{\ell\ell'}$ becomes $M^{00}_{\ell\ell'}$.

In our case, we analyse the full sky, meaning our mask has a value of 1 for every pixel. For a full-sky, pure spin-0 field like the galaxy over-density, the coupling matrix becomes the identity matrix, $M_{\ell\ell'} = \mathbb{1}$. Therefore, in this instance, the pseudo-C_ℓ estimator is an unbiased estimate of the true C_ℓ, and equation (5.14) simplifies significantly.

Then, we provide `NaMaster` with our over-density maps and a corresponding full-sky mask. We then bin the multipoles into bands of width $\Delta\ell = 10$ to reduce noise. The resulting measurements of the cross-power spectra, for our redshift bins, are shown in Figure 5.7. These are compared against the theoretical prediction, described in Chapter 3, derived from the "True" galaxy catalogue.

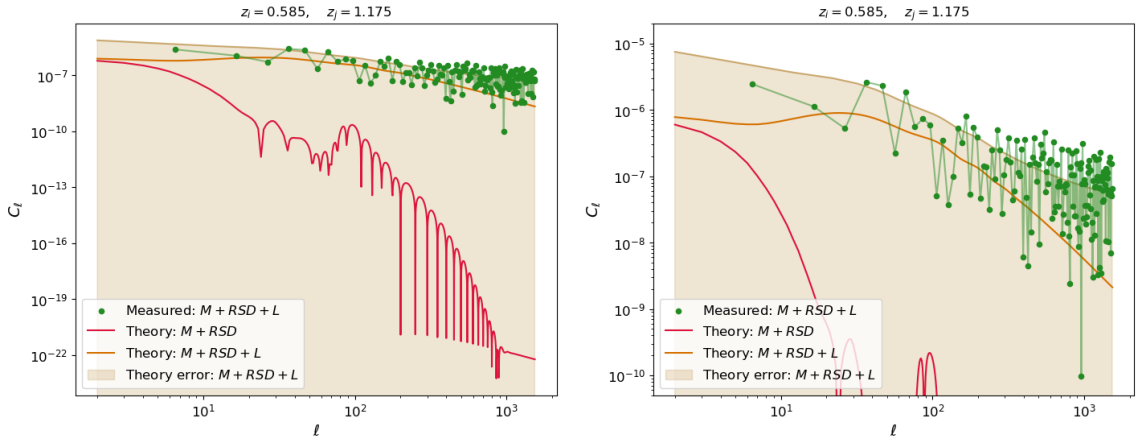


FIGURE 5.7: Measured angular power spectra from the simulated galaxy catalogue (green points) compared to the theoretical prediction with lensing (orange line) and without it (red line). The measurement is performed using the `NaMaster` pseudo-C_ℓ estimator.

When comparing the theoretical predictions of Figure 3.6 with those in Figure 5.7, the larger error bars in the latter are immediately noticeable. This is a direct consequence of our computational constraints, which forced us to reduce the simulated galaxy number density by a factor of 100. Therefore, as equation (3.19) indicates, this reduction amplifies the shot noise, $P_{N,ij}$, by a factor of 100.

5.3.2. Measure Angular Correlation function

Having measured the Angular Power Spectrum, the subsequent focus of study of our simulation will be the Angular Correlation function, $w(\theta)$. For this we use the `TreeCorr` package [5], which efficiently computes two-point statistics by using a tree data structure to count pairs of galaxies.

The simplest estimator of $w(\theta)$ is given by the fractional amount of pairs above random [25],

$$w_0(\theta) = \frac{DD(\theta)}{RR(\theta)} - 1\,, \tag{5.15}$$

where $DD(\theta)$ is the number of galaxy pairs separated by an angle θ, and $RR(\theta)$ is the analogous count for a uniform random distribution. However, this estimator is sensitive to edge effects and surface density uncertainties. Fortunately, multiple estimators have been constructed in order to improve from this one. The Landy-Szalay estimator [45], $w_{LS}(\theta)$, minimises the bias and variance of the angular range under investigation, therefore mitigating these issues. For this reason, $w_{LS}(\theta)$ is considered to be the best estimator of

$w(\theta)$ [25]:

$$w_{LS}(\theta) = \frac{r(r-1)}{n(n-1)}\frac{DD}{RR} - \frac{r-1}{n}\frac{DR}{RR} + \frac{RR}{RR}. \tag{5.16}$$

Here, n is the number of data points and r the number of random points.

Furthermore, for a cross-correlation between two data catalogues, D_i and D_j, `TreeCorr` uses this estimator as

$$w_{LS}(\theta) = \frac{D_iD_j - D_iR - RD_j + RR}{RR}, \tag{5.17}$$

where R is a catalogue of random points uniformly distributed within the survey footprint. The terms D_iR and RD_j are the cross-pair counts between the data and random catalogues.

We can now create `TreeCorr` catalogues from the RA and Dec of galaxies in each redshift bin. After that, we generate a random catalogue, R, containing a number of points twice as big as the largest of the two bins to minimise Poisson noise. We then use `TreeCorr` to compute the D_iD_j, D_iR, RD_j, and RR pair counts and apply equation (5.17) to estimate $w(\theta)$. So as to reduce the variance in the comparison of the data with random points, we repeat this process 10 times with different random catalogue realisations each instance. Afterwards we average the results and obtain the final measurement, which is shown in Figure 5.8, where it's compared to the theoretical model.

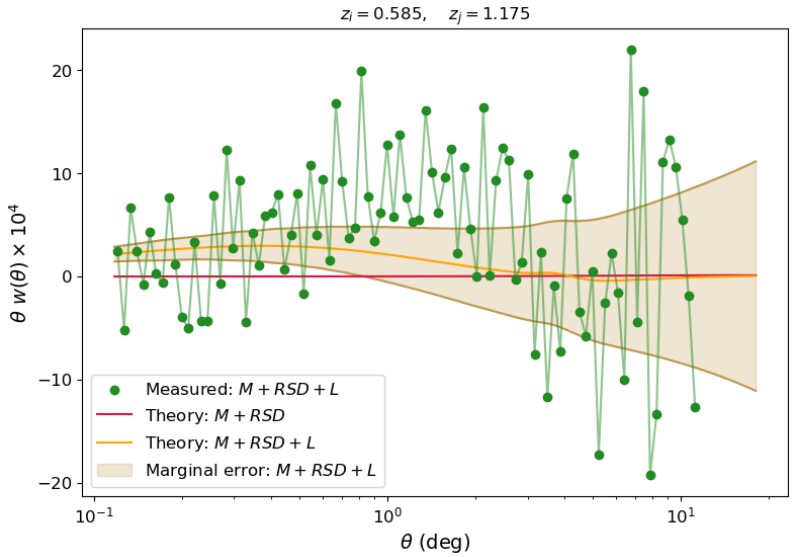


FIGURE 5.8: Measured angular correlation function, $w(\theta)$, from the simulated catalogue (points) using the `TreeCorr` implementation of the Landy-Szalay estimator. The solid line shows the theoretical prediction. Each point is the average of 10 measurements with different random point realisations.

5.4. Discussion

With this work we successfully generated a realistic, observable galaxy catalogue, that includes the effects of photometric redshift errors, RSD, and magnification lensing. The primary objective was to test whether standard LSS estimators could recover the underlying clustering signal and detect the imprinted lensing signature from this synthetic data. This section discusses the practical challenges encountered, interprets the results obtained, and outlines the main conclusions.

5.4.1. Challenges and Limitations encountered

The generation of high-fidelity cosmological simulations, and their analysis, can be computationally demanding. So, some limitations were encountered that required compromises between physical realism and computational feasibility.

The most significant constraint was Random-Access Memory (RAM) usage. The initial configuration, designed for high resolution and a large number of galaxies, proved to be too much. This led to a few adjustments:

1. **Reduced Galaxy Count:** The total number of galaxies in the simulation was lowered by a factor of 100, making it around 3.5×10^6 sources, in order to stay within available RAM limits. However, there is a critical trade-off: a sparser galaxy field increases Poisson noise, which can bias clustering measurements and cause deviations from theoretical predictions.

2. **Grid Resolution:** The spatial resolution was set by $N_{\text{grid}} = 512$. A lower resolution would have compromised the ability to resolve the finer structures of the cosmic web, directly impacting the accuracy of the generated LSS.

These compromises highlight a fundamental challenge in synthetic universe creation: balancing statistical significance and resolution against computational resources. Future work with access to better computing resources could improve our results.

5.4.2. Interpretation of the Clustering Measurements

The C_ℓ and $w(\theta)$ measured from the final observed catalogue are shown in Figures 5.7 and 5.8, respectively. While the measurements do not perfectly align with the theoretical predictions derived from the true galaxy distribution, they successfully capture the characteristic shape and scaling of the clustering signal across angular scales.

The observed discrepancies can be attributed to the following factors:

- **Sample Variance and Noise:** The smaller than initially intended volume and galaxy count of our simulation might have lead to inherent sample variance and shot noise, which deviate the measurements from the underlying theory curve.

- **Estimator Performance:** The measurement of $w(\theta)$ via the Landy-Szalay estimator can be further refined. Increasing the number of random points in the random catalogue (R) and averaging over a larger number of iterations would reduce Poisson noise and yield a smoother, more precise estimate.

Most importantly, the overall agreement in the profile of both C_ℓ and $w(\theta)$ with predictions demonstrates that the pipeline, from simulation generation to map-making and statistical estimation, is functionally correct. The discrepancies are consistent with expected statistical fluctuations rather than systematic errors in the methodology.

5.4.3. Conclusions on Measuring Magnification Lensing

The central goal of this chapter was to incorporate magnification lensing into a simulated galaxy catalogue and determine if its impact was detectable with standard LSS estimators. The results confirm that this is achievable!

Moreover, the measured angular clustering signals show a consistent relation with the theoretical predictions that include the lensing effect. The close correspondence between the measured and predicted profiles for both C_ℓ and $w(\theta)$ indicates that the lensing signal has been successfully encoded in the catalogue through the perturbation of galaxy positions and magnitudes, and subsequently recovered through the analysis.

For this reason, we conclude that it is possible to measure the impact of magnification lensing on galaxy clustering using standard angular clustering statistics. This successful proof-of-concept validates the entire workflow, from generating the initial density field with `CoLoRe` to applying observational effects and measuring the final statistics with `NaMaster` and `TreeCorr`. This pipeline provides a robust foundation for future studies aiming to isolate and constrain lensing effects in photometric galaxy surveys. The next step is, therefore, to use real data instead of the simulations and see if the magnification lensing is still measurable.

Furthermore, future surveys with high redshift range, such as LSST (Large Synoptic Survey Telescope) [46] or Euclid [47], can be used to constrain GR as already forecasted in some papers [21, 41, 48].

6. Conclusions

This dissertation aimed to detect the weak lensing magnification effect in photometric galaxy surveys. We have not only characterised the expected impact of magnification lensing in such surveys but also measured it from simulated catalogues using two different LSS estimators. Furthermore, we conducted a comparative analysis of those estimators to determine their relative effectiveness under different conditions.

In Chapter 1, we outlined this dissertation's main topic and structure. Next, Chapter 2 described the essential theoretical background, for cosmology and weak gravitational lensing, used throughout this work. Afterwards, Chapter 3 introduced the LSS estimators and demonstrated how magnification lensing affects them, showing that it becomes the dominant term for redshift bins with sufficient separation. Moreover, in Chapter 4 we compared the performance of C_ℓ and $w(\theta)$ by using them, in conjunction with Fisher Matrices, to constrain the values of specified parameters. Our analysis revealed that full-correlations between redshift bins are essential for measuring the magnification bias parameter, $s(z)$, and parameter β; as these parameters are lensing-dependent. Neither estimator demonstrated universally superior performance, due to the fact that their effectiveness depends on specific survey configurations, but the results obtained with full-correlations were much superior than those done with only autocorrelations. Furthermore, we found that the access to observations for a large fraction of the sky, large f_{sky}, is extremely important in order to obtain accurate measurements of cosmological parameters, or parameters $s(z)$ and β. Finally, in Chapter 5 we generated a simulation of a "perfect" galaxy catalogue and transformed it into an observed catalogue by introducing realistic observational effects. We then applied estimators of C_ℓ and $w(\theta)$ to measure the magnification lensing signal. Although our measurements exhibited some deviations from theoretical predictions, we successfully confirmed that the magnification lensing convergence effect is detectable in simulations that mimic observational data.

In conclusion, this work has advanced our understanding of how to detect magnification lensing in photometric galaxy surveys. We have identified the conditions under which

this effect is most significant and demonstrated its measurability using simulated catalogues.

Ergo, some possible extensions of this work are:

- **Application to Real Data:** The natural next step is to apply this methodology to real observational data, rather than only simulations, thereby validating our analysis with actual galaxy surveys.
- **Parameter Estimation with MCMC:** While this work used Fisher matrix forecasts to predict minimum uncertainties, future work should employ Markov Chain Monte Carlo (MCMC) methods to directly measure the values of $s(z)$ and β.
- **Constraining Lensing Convergence:** The measured values of $s(z)$ from MCMC, combined with estimates of Δ^L from C_ℓ or $w(\theta)$, can be used to constrain the magnification lensing convergence κ.
- **Testing General Relativity:** With constraints on κ and β, equation (2.57) can be used to constrain the sum of the gravitational potentials $\Phi + \Psi$. This provides a pathway to test theories of gravity, including General Relativity.

The comprehensive strategy outlined above, for the constrain of the value of $\Phi + \Psi$, is summarised in Figure 6.1.

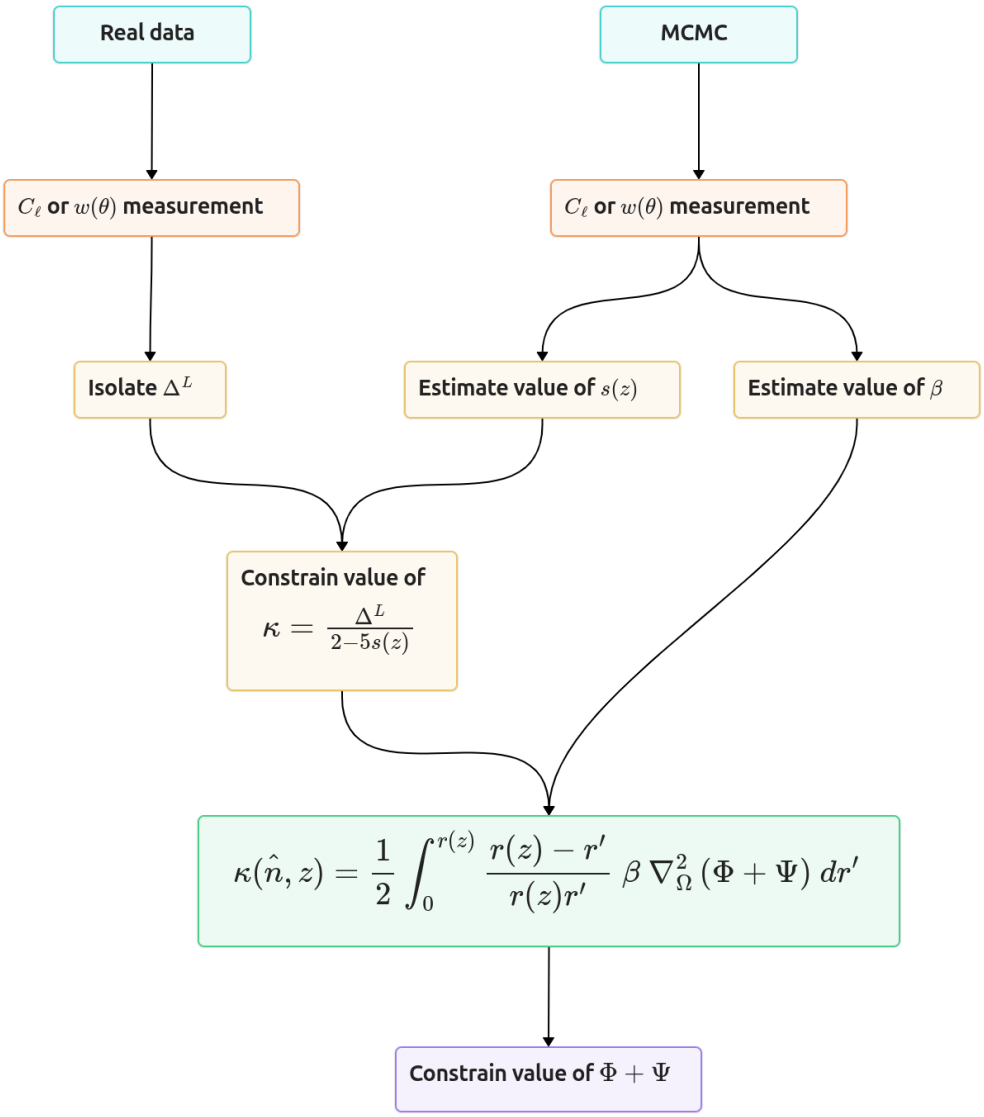


FIGURE 6.1: Strategy to constrain gravitational potentials Φ and Ψ.

Bibliography

[1] A. Challinor and A. Lewis, "Linear power spectrum of observed source number counts," *Physical Review D*, vol. 84, no. 4, Aug. 2011. [Online]. Available: http://dx.doi.org/10.1103/PhysRevD.84.043516 [Cited on pages iii, iv, 18, and 31.]

[2] N. E. Chisari et al., "Core cosmology library: Precision cosmological predictions for lsst," *The Astrophysical Journal Supplement Series*, vol. 242, no. 1, p. 2, May 2019. [Online]. Available: http://dx.doi.org/10.3847/1538-4365/ab1658 [Cited on pages iii, iv, and 31.]

[3] C. Ramírez-Pérez, J. Sanchez, D. Alonso, and A. Font-Ribera, "Colore: fast cosmological realisations over large volumes with multiple tracers," *Journal of Cosmology and Astroparticle Physics*, vol. 2022, no. 05, p. 002, May 2022. [Online]. Available: http://dx.doi.org/10.1088/1475-7516/2022/05/002 [Cited on pages iii, iv, 58, 61, and 62.]

[4] D. Alonso, J. Sanchez, and A. Slosar, "A unified pseudo-c-ell framework," *Monthly Notices of the Royal Astronomical Society*, vol. 484, no. 3, p. 4127–4151, Jan. 2019. [Online]. Available: http://dx.doi.org/10.1093/mnras/stz093 [Cited on pages iii, iv, 65, and 67.]

[5] M. Jarvis, G. Bernstein, and B. Jain, "The skewness of the aperture mass statistic," *mnras*, vol. 352, no. 1, pp. 338–352, Jul. 2004. [Cited on pages iii, iv, and 68.]

[6] W. C. Keel, *The road to galaxy formation*, 2007. [Cited on page 1.]

[7] J. K. Yadav, J. S. Bagla, and N. Khandai, "Fractal dimension as a measure of the scale of homogeneity: The scale of homogeneity," *Monthly Notices of the Royal Astronomical Society*, pp. no–no, Apr. 2010. [Online]. Available: http://dx.doi.org/10.1111/j.1365-2966.2010.16612.x [Cited on pages 1 and 3.]

[8] P. J. E. Peebles, *The Large-Scale Structure of the Universe*. Princeton University Press, 1980. [Online]. Available: http://www.jstor.org/stable/j.ctvxrpz4n [Cited on pages 1, 29, and 32.]

[9] P. J. E. Peebles, *Principles of Physical Cosmology*. Princeton University Press, 09 2020. [Online]. Available: https://doi.org/10.23943/princeton/9780691209814.001.0001 [Cited on page 1.]

[10] M. Bartelmann and P. Schneider, "Weak gravitational lensing," *Physics Reports*, vol. 340, no. 4, pp. 291–472, 2001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S037015730000082X [Cited on pages ix, 1, 8, 9, 10, 11, and 12.]

[11] D. Baumann, *Cosmology*. Cambridge University Press, 2022. [Cited on pages 3, 4, 5, 6, 7, 13, 14, and 15.]

[12] B. W. Carroll and D. A. Ostlie, *An introduction to modern astrophysics, Second Edition*, 2017. [Cited on page 5.]

[13] P. Coles and P. Lucchin, *Cosmology: The Origin and Evolution of Cosmic Structure*. Wiley, 2003. [Online]. Available: https://books.google.pt/books?id=BGYcivB1EtMC [Cited on pages 7, 8, and 15.]

[14] K. Umetsu, "Cluster–galaxy weak lensing," *The Astronomy and Astrophysics Review*, vol. 28, no. 1, Nov. 2020. [Online]. Available: http://dx.doi.org/10.1007/s00159-020-00129-w [Cited on pages 11 and 12.]

[15] J. M. Bardeen, "Gauge-invariant cosmological perturbations," *Phys. Rev. D*, vol. 22, pp. 1882–1905, Oct 1980. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.22.1882 [Cited on page 14.]

[16] H. e. a. Aihara, "The Eighth Data Release of the Sloan Digital Sky Survey: First Data from SDSS-III," *apjs*, vol. 193, no. 2, p. 29, Apr. 2011. [Cited on page 15.]

[17] A. Kravtsov and A. Klypin, "Simulation of the formation of the large-scale structure in the universe," siulations perfromed at the National Center for Supercomputer Applications. [Online]. Available: https://cosmicweb.uchicago.edu/sims.html [Cited on pages ix, 15, and 16.]

[18] M. Blanton and Sloan Digital Sky Survey, "SDSS galaxy map." [Online]. Available: https://www.sdss3.org/science/gallery_sdss_pie2.php [Cited on pages ix and 16.]

[19] C. Bonvin and R. Durrer, "What galaxy surveys really measure," *Physical Review D*, vol. 84, no. 6, Sep. 2011. [Online]. Available: http://dx.doi.org/10.1103/PhysRevD.84.063505 [Cited on pages 17, 33, 34, and 35.]

[20] E. D. Dio, F. Montanari, J. Lesgourgues, and R. Durrer, "The classgal code for relativistic cosmological large scale structure," *Journal of Cosmology and Astroparticle Physics*, vol. 2013, no. 11, p. 044–044, Nov. 2013. [Online]. Available: http://dx.doi.org/10.1088/1475-7516/2013/11/044 [Cited on pages 17, 18, 25, 26, 27, 33, and 83.]

[21] F. Montanari and R. Durrer, "Measuring the lensing potential with tomographic galaxy number counts," *Journal of Cosmology and Astroparticle Physics*, vol. 2015, no. 10, p. 070–070, Oct. 2015. [Online]. Available: http://dx.doi.org/10.1088/1475-7516/2015/10/070 [Cited on pages 18, 19, 21, 22, 26, 43, 55, 59, and 71.]

[22] M. LoVerde, L. Hui, and E. Gaztañaga, "Lensing corrections to features in the angular two-point correlation function and power spectrum," *Phys. Rev. D*, vol. 77, p. 023512, Jan 2008. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.77.023512 [Cited on page 19.]

[23] A. Lewis and A. Challinor, "Weak gravitational lensing of the cmb," *Physics Reports*, vol. 429, no. 1, p. 1–65, Jun. 2006. [Online]. Available: http://dx.doi.org/10.1016/j.physrep.2006.03.002 [Cited on page 21.]

[24] I. D. Saltas, I. Sawicki, L. Amendola, and M. Kunz, "Anisotropic stress as a signature of nonstandard propagation of gravitational waves," *Physical Review Letters*, vol. 113, no. 19, Nov. 2014. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.113.191101 [Cited on page 21.]

[25] J. Wall, "Practical statistics for astronomers, second edition," *Practical Statistics for Astronomers: Second Edition*, pp. 1–353, 06 2012. [Cited on pages ix, 24, 26, 27, 28, 31, 68, and 69.]

[26] G. B. Arfken, H. J. Weber, and F. E. Harris, *Mathematical Methods for Physicists: A Comprehensive Guide*, 7th ed. Boston: Academic Press, 2013. [Cited on pages 24 and 25.]

[27] C. Pryke, "Logbook entry for 2000-09-22," http://find.spa.umn.edu/~pryke/logbook/20000922/, 09 2000, accessed: 2024-11-21. [Online]. Available: http://find.spa.umn.edu/~pryke/logbook/20000922/ [Cited on pages ix and 25.]

[28] J. Fonseca and M. Liguori, "Measuring ultralarge scale effects in the presence of 21 cm intensity mapping foregrounds," *Monthly Notices of the Royal Astronomical Society*, vol. 504, no. 1, p. 267–279, Apr. 2021. [Online]. Available: http://dx.doi.org/10.1093/mnras/stab903 [Cited on pages 26 and 39.]

[29] P. Peebles, "Statistical Analysis of Catalogs of Extragalactic Objects. I. Theory," *apj*, vol. 185, pp. 413–440, Oct. 1973. [Cited on page 28.]

[30] J. Fonseca, R. Maartens, and M. G. Santos, "Probing the primordial Universe with MeerKAT and DES," *Monthly Notices of the Royal Astronomical Society*, vol. 466, no. 3, p. 2780–2786, Dec. 2016. [Online]. Available: http://dx.doi.org/10.1093/mnras/stw3248 [Cited on pages 29, 30, 31, 45, 59, and 63.]

[31] Z. Ma, W. Hu, and D. Huterer, "Effects of photometric redshift uncertainties on weak-lensing tomography," *The Astrophysical Journal*, vol. 636, no. 1, p. 21–29, Jan. 2006. [Online]. Available: http://dx.doi.org/10.1086/497068 [Cited on page 30.]

[32] M. Crocce, A. Cabré, and E. Gaztañaga, "Modelling the angular correlation function and its full covariance in photometric galaxy surveys: Angular clustering in photometric surveys," *Monthly Notices of the Royal Astronomical Society*, vol. 414, no. 1, p. 329–349, Mar. 2011. [Online]. Available: http://dx.doi.org/10.1111/j.1365-2966.2011.18393.x [Cited on pages 30, 40, 83, and 84.]

[33] Planck Collaboration, "Planck 2018 results: VI. cosmological parameters," *Astronomy & Astrophysics*, vol. 641, p. A6, Sep. 2020. [Online]. Available: https://doi.org/10.1051/0004-6361/201833910 [Cited on pages 31 and 42.]

[34] J. Asorey Barreiro, "Measuring large scale structure using angular cross-correlations," Ph.D. dissertation, 2013. [Cited on page 34.]

[35] J. Yoo, A. L. Fitzpatrick, and M. Zaldarriaga, "New perspective on galaxy clustering as a cosmological probe: General relativistic effects," *Phys. Rev. D*, vol. 80, p. 083514, Oct 2009. [Online]. Available: https://arxiv.org/abs/0907.0707 [Cited on page 34.]

[36] D. Jeong, F. Schmidt, and C. M. Hirata, "Large-scale clustering of galaxies in general relativity," *Phys. Rev. D*, vol. 85, p. 023504, Jan 2012. [Online]. Available: https://arxiv.org/abs/1205.1512 [Cited on page 34.]

[37] D. Huterer, *A Course in Cosmology: From Theory to Practice*. Cambridge University Press, 2023. [Cited on pages 37 and 39.]

[38] M. Tegmark, A. N. Taylor, and A. F. Heavens, "Karhunen-loeve eigenvalue problems in cosmology: How should we tackle large data sets?" *The Astrophysical Journal*, vol. 480, no. 1, p. 22–35, May 1997. [Online]. Available: http://dx.doi.org/10.1086/303939 [Cited on page 38.]

[39] A. Heavens, "Statistical techniques in cosmology," 2010. [Online]. Available: https://arxiv.org/abs/0906.0664 [Cited on page 39.]

[40] A. Gabasch, U. Hopp, G. Feulner, R. Bender, S. Seitz, R. P. Saglia, J. Snigula, N. Drory, I. Appenzeller, J. Heidt, D. Mehlert, S. Noll, A. Böhm, K. Jäger, and B. Ziegler, "The evolution of the luminosity functions in the FORS deep field from low to high redshift. II. The red bands," *aap*, vol. 448, no. 1, pp. 101–121, Mar. 2006. [Cited on page 59.]

[41] D. Alonso, P. Bull, P. G. Ferreira, R. Maartens, and M. G. Santos, "Ultra-large-scale Cosmology in Next-generation Experiments with Single Tracers," *apj*, vol. 814, no. 2, p. 145, dec 2015. [Cited on pages 59 and 71.]

[42] A. Zonca, L. Singer, D. Lenz, M. Reinecke, C. Rosset, E. Hivon, and K. Gorski, "healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in python," *Journal of Open Source Software*, vol. 4, no. 35, p. 1298, Mar. 2019. [Online]. Available: https://doi.org/10.21105/joss.01298 [Cited on page 65.]

[43] K. Górski, E. Hivon, A. Banday, B. Wandelt, F. Hansen, M. Reinecke, and M. Bartelmann, "HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere," *The Astrophysical Journal*, vol. 622, pp. 759–771, Apr. 2005. [Cited on page 65.]

[44] E. Hivon, K. M. Górski, C. B. Netterfield, B. P. Crill, S. Prunet, and F. Hansen, "Master of the cosmic microwave background anisotropy power spectrum: A fast method for statistical analysis of large and complex cosmic microwave background

data sets," *The Astrophysical Journal*, vol. 567, no. 1, p. 2, mar 2002. [Online]. Available: https://dx.doi.org/10.1086/338126 [Cited on page 67.]

[45] S. D. Landy and A. S. Szalay, "Bias and Variance of Angular Correlation Functions," *The Astrophysical Journal*, vol. 412, p. 64, Jul. 1993. [Cited on page 68.]

[46] LSST Science Collaboration, P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel, L. Armus, D. Arnett, S. J. Asztalos, T. S. Axelrod, S. Bailey, D. R. Ballantyne, J. R. Bankert, W. A. Barkhouse, J. D. Barr, L. F. Barrientos, A. J. Barth, J. G. Bartlett, A. C. Becker, J. Becla, T. C. Beers, J. P. Bernstein, R. Biswas, M. R. Blanton, J. S. Bloom, J. J. Bochanski, P. Boeshaar, K. D. Borne, M. Bradac, W. N. Brandt, C. R. Bridge, M. E. Brown, R. J. Brunner, J. S. Bullock, A. J. Burgasser, J. H. Burge, D. L. Burke, P. A. Cargile, S. Chandrasekharan, G. Chartas, S. R. Chesley, Y.-H. Chu, D. Cinabro, M. W. Claire, C. F. Claver, D. Clowe, A. J. Connolly, K. H. Cook, J. Cooke, A. Cooray, K. R. Covey, C. S. Culliton, R. de Jong, W. H. de Vries, V. P. Debattista, F. Delgado, I. P. Dell'Antonio, S. Dhital, R. Di Stefano, M. Dickinson, B. Dilday, S. G. Djorgovski, G. Dobler, C. Donalek, G. Dubois-Felsmann, J. Durech, A. Eliasdottir, M. Eracleous, L. Eyer, E. E. Falco, X. Fan, C. D. Fassnacht, H. C. Ferguson, Y. R. Fernandez, B. D. Fields, D. Finkbeiner, E. E. Figueroa, D. B. Fox, H. Francke, J. S. Frank, J. Frieman, S. Fromenteau, M. Furqan, G. Galaz, A. Gal-Yam, P. Garnavich, E. Gawiser, J. Geary, P. Gee, R. R. Gibson, K. Gilmore, E. A. Grace, R. F. Green, W. J. Gressler, C. J. Grillmair, S. Habib, J. S. Haggerty, M. Hamuy, A. W. Harris, S. L. Hawley, A. F. Heavens, L. Hebb, T. J. Henry, E. Hileman, E. J. Hilton, K. Hoadley, J. B. Holberg, M. J. Holman, S. B. Howell, L. Infante, Z. Ivezic, S. H. Jacoby, B. Jain, R, Jedicke, M. J. Jee, J. Garrett Jernigan, S. W. Jha, K. V. Johnston, R. L. Jones, M. Juric, M. Kaasalainen, Styliani, Kafka, S. M. Kahn, N. A. Kaib, J. Kalirai, J. Kantor, M. M. Kasliwal, C. R. Keeton, R. Kessler, Z. Knezevic, A. Kowalski, V. L. Krabbendam, K. S. Krughoff, S. Kulkarni, S. Kuhlman, M. Lacy, S. Lepine, M. Liang, A. Lien, P. Lira, K. S. Long, S. Lorenz, J. M. Lotz, R. H. Lupton, J. Lutz, L. M. Macri, A. A. Mahabal, R. Mandelbaum, P. Marshall, M. May, P. M. McGehee, B. T. Meadows, A. Meert, A. Milani, C. J. Miller, M. Miller, D. Mills, D. Minniti, D. Monet, A. S. Mukadam, E. Nakar, D. R. Neill, J. A. Newman, S. Nikolaev, M. Nordby, P. O'Connor, M. Oguri, J. Oliver, S. S. Olivier, J. K. Olsen, K. Olsen, E. W. Olszewski, H. Oluseyi, N. D. Padilla, A. Parker, J. Pepper, J. R. Peterson, C. Petry, P. A. Pinto, J. L. Pizagno, B. Popescu, A. Prsa, V. Radcka, M. J. Raddick, A. Rasmussen, A. Rau, J. Rho,

J. E. Rhoads, G. T. Richards, S. T. Ridgway, B. E. Robertson, R. Roskar, A. Saha, A. Sarajedini, E. Scannapieco, T. Schalk, R. Schindler, and S. Schmidt, "LSST Science Book, Version 2.0," *arXiv e-prints*, p. arXiv:0912.0201, Dec. 2009. [Cited on page 71.]

[47] F. Lepori, I. Tutusaus, C. Viglione, C. Bonvin, S. Camera, F. J. Castander, R. Durrer, P. Fosalba, G. Jelic-Cizmek, M. Kunz, J. Adamek, S. Casas, M. Martinelli, Z. Sakr, D. Sapone, A. Amara, N. Auricchio, C. Bodendorf, D. Bonino, E. Branchini, M. Brescia, J. Brinchmann, V. Capobianco, C. Carbone, J. Carretero, M. Castellano, S. Cavuoti, A. Cimatti, R. Cledassou, G. Congedo, C. J. Conselice, L. Conversi, Y. Copin, L. Corcione, F. Courbin, A. Da Silva, H. Degaudenzi, M. Douspis, F. Dubath, X. Dupac, S. Dusini, A. Ealet, S. Farrens, S. Ferriol, E. Franceschi, M. Fumana, B. Garilli, W. Gillard, B. Gillis, C. Giocoli, A. Grazian, F. Grupp, L. Guzzo, S. V. H. Haugan, W. Holmes, F. Hormuth, P. Hudelot, K. Jahnke, S. Kermiche, A. Kiessling, M. Kilbinger, T. Kitching, M. Kümmel, H. Kurki-Suonio, S. Ligori, P. B. Lilje, I. Lloro, O. Mansutti, O. Marggraf, K. Markovic, F. Marulli, R. Massey, S. Maurogordato, M. Melchior, M. Meneghetti, E. Merlin, G. Meylan, M. Moresco, L. Moscardini, E. Munari, R. Nakajima, S. M. Niemi, C. Padilla, S. Paltani, F. Pasian, K. Pedersen, W. J. Percival, V. Pettorino, S. Pires, M. Poncet, L. Popa, L. Pozzetti, F. Raison, J. Rhodes, M. Roncarelli, E. Rossetti, R. Saglia, P. Schneider, A. Secroun, G. Seidel, S. Serrano, C. Sirignano, G. Sirri, L. Stanco, J.-L. Starck, P. Tallada-Crespí, A. N. Taylor, I. Tereno, R. Toledo-Moreo, F. Torradeflot, E. A. Valentijn, L. Valenziano, Y. Wang, J. Weller, G. Zamorani, J. Zoubian, S. Andreon, S. Bardelli, G. Fabbian, J. Graciá-Carpio, D. Maino, E. Medinaceli, S. Mei, A. Renzi, E. Romelli, F. Sureau, T. Vassallo, A. Zacchei, E. Zucca, C. Baccigalupi, A. Balaguera-Antolínez, F. Bernardeau, A. Biviano, A. Blanchard, M. Bolzonella, S. Borgani, E. Bozzo, C. Burigana, R. Cabanac, A. Cappi, C. S. Carvalho, G. Castignani, C. Colodro-Conde, J. Coupon, H. M. Courtois, J.-G. Cuby, S. Davini, S. de la Torre, D. Di Ferdinando, M. Farina, P. G. Ferreira, F. Finelli, S. Galeotta, K. Ganga, J. Garcia-Bellido, E. Gaztanaga, G. Gozaliasl, I. M. Hook, S. Ilić, B. Joachimi, V. Kansal, E. Keihanen, C. C. Kirkpatrick, V. Lindholm, G. Mainetti, R. Maoli, N. Martinet, M. Maturi, R. B. Metcalf, P. Monaco, G. Morgante, J. Nightingale, A. Nucita, L. Patrizii, V. Popa, D. Potter, G. Riccio, A. G. Sánchez, M. Schirmer, M. Schultheis, V. Scottez, E. Sefusatti, A. Tramacere, J. Valiviita,

M. Viel, and H. Hildebrandt, “Euclid preparation: Xix. impact of magnification on photometric galaxy clustering,” *Astronomy & Astrophysics*, vol. 662, p. A93, Jun. 2022. [Online]. Available: http://dx.doi.org/10.1051/0004-6361/202142419 [Cited on page 71.]

[48] D. J. Bacon, R. A. Battye, P. Bull, S. Camera, P. G. Ferreira, I. Harrison, D. Parkinson, A. Pourtsidou, M. G. Santos, L. Wolz, F. Abdalla, Y. Akrami, D. Alonso, S. Andrianomena, M. Ballardini, J. L. Bernal, D. Bertacca, C. A. P. Bengaly, A. Bonaldi, C. Bonvin, M. L. Brown, E. Chapman, S. Chen, X. Chen, S. Cunnington, T. M. Davis, C. Dickinson, J. Fonseca, K. Grainge, S. Harper, M. J. Jarvis, R. Maartens, N. Maddox, H. Padmanabhan, J. R. Pritchard, A. Raccanelli, M. Rivi, S. Roychowdhury, M. Sahlén, D. J. Schwarz, T. M. Siewert, M. Viel, F. Villaescusa-Navarro, Y. Xu, D. Yamauchi, and J. Zuntz, “Cosmology with phase 1 of the square kilometre array red book 2018: Technical specifications and performance forecasts,” *Publications of the Astronomical Society of Australia*, vol. 37, 2020. [Online]. Available: http://dx.doi.org/10.1017/pasa.2019.51 [Cited on page 71.]

A. Theoretical error bars

A.1. Error bars of the Angular Power Spectrum

Here, we will determine the equations of $\sigma\left(C_\ell\right)$, which represent the theoretical error associated with our calculations of C_ℓ.

For the case in which the redshift bins that we are evaluating are the same, i.e. they are autocorrelated ($z_i = z_j$) we use the parameter $\sigma\left(C_{\ell,ii}\right)$ to obtain the error associated with each C_ℓ value determined previously. So, from Dio et al. [20] we have that

$$\sigma\left(C_{\ell,ii}\right) = \sqrt{\frac{2}{2\ell+1}\frac{1}{f_{sky}\Delta\ell}}(C_{\ell,ii} + P_{N,ii})\,, \tag{A.1}$$

where $\Delta\ell$ represents the size of the bin in ℓ, i.e. if $\Delta\ell = 10$ then there are 10 values of ℓ inside each bin. The Shot Noise, $P_{N,ij}$ was explained in equation (3.19) and the fraction of the observed sky, f_{sky}, was explained in section 4.1.2.

On the other hand, in the cross-correlation case, $z_i \neq z_j$, the expression is a little more complex. From Crocce et al. [32] we see that

$$\sigma\left(C_{\ell,ij}\right) = \sqrt{\frac{1}{2\ell+1}\frac{1}{f_{sky}\Delta\ell}\left[C_\ell^2(z_i,z_j) + (C_{\ell i} + P_{Ni})(C_{\ell j} + P_{Nj})\right]}\,. \tag{A.2}$$

A.2. Error bars of the Angular Correlation Function

As before, the theoretical error associated with our calculations of $w(\theta)$ is $\sigma\big(w(\theta)\big)$, which is given by [32]

$$\sigma\big(w_{marg}(\theta)\big) \geq \sqrt{Cov\big(w(\theta), w(\theta')\big)}\,, \tag{A.3}$$

where $Cov\big(w(\theta), w(\theta')\big)$ is the value of one element of the Covariance matrix of $w(\theta)$.

We start by obtaining the Covariance matrix of $w(\theta)$. If we are considering only the autocorrelation case, then we can simply use the expression of $Cov\big(w_{ii}(\theta), w_{ii}(\theta')\big)$, otherwise we have to use the more general cross-correlation case, $Cov\big(w_{ij}(\theta), w_{ij}(\theta')\big)$. Then,

in order to do this calculations we will use the following equations: [32]

$$Cov\big(w_{ii}(\theta), w_{ii}(\theta')\big) = \langle\, w_{ii}(\theta) w_{ii}(\theta')\,\rangle \tag{A.4}$$
$$= \frac{2}{f_{sky}} \sum_{\ell \geq 0} \frac{2\ell+1}{(4\pi)^2} P_\ell(\cos\theta) P_\ell(\cos\theta') \left[(C_{\ell,ii} + P_{N,ii})\right]^2 ,$$

$$Cov\big(w_{ij}(\theta), w_{ij}(\theta')\big) = \langle\, w_{ij}(\theta) w_{ij}(\theta')\,\rangle \tag{A.5}$$
$$= \frac{1}{f_{sky}} \sum_{\ell \geq 0} \frac{2\ell+1}{(4\pi)^2} P_\ell(\cos\theta) P_\ell(\cos\theta') \left[(C_{\ell,ii} + P_{N,ii})(C_{\ell,jj} + P_{N,jj}) + C_{\ell,ij}^2\right] .$$

Here, equations (A.4) and (A.5) represent each element of the Covariance matrices of the auto-correlations, and the cross-correlations, of $w(\theta)$, respectively.

After having the expressions of the Covariance matrices, we can easily determine the marginal errors, $\sigma\big(w_{marg}(\theta)\big)$, by square rooting the values on its diagonal.

B. Analytic derivatives

B.1. Analytic derivative of $\partial C_\ell/\partial\beta$

From equation (3.10), it is known that the angular power spectrum is given by

$$C_{\ell,ij} = \int \frac{1}{k}\mathcal{P}(k)\Delta_{\ell,i}(k)\Delta_{\ell,j}(k)\, dk\,. \tag{B.1}$$

Let's also recall that, in our analysis, we are not considering every component of the number counts, we're just using $\Delta_i^G = \Delta_i^M(k) + \Delta_i^{RSD}(k) + \Delta_i^L(k)$, as in equation (3.20). Therefore, the derivative of C_ℓ in order to β can be expressed as

$$\begin{aligned}
\frac{\partial C_{\ell,ij}}{\partial\beta} &= \frac{\partial}{\partial\beta}\int \frac{1}{k}\mathcal{P}(k)\Delta_i^G\Delta_j^G\, dk \\
&= \frac{\partial}{\partial\beta}\int \frac{1}{k}\mathcal{P}(k)\left(\Delta_i^M + \Delta_i^{RSD} + \Delta_i^L\right)\left(\Delta_j^M + \Delta_j^{RSD} + \Delta_j^L\right)\, dk \\
&= \int \frac{1}{k}\mathcal{P}(k)\frac{\partial}{\partial\beta}\Big[(\Delta^M)^2 + (\Delta^{RSD})^2 + \Delta_i^M\Delta_j^{RSD} + \Delta_i^{RSD}\Delta_j^M \\
&\qquad\qquad +(\Delta_i^M + \Delta_i^{RSD})\Delta_j^L + \Delta_i^L(\Delta_j^M\Delta_j^{RSD}) + (\Delta^L)^2\Big]\, dk\,.
\end{aligned}$$

This expression can be simplified by remembering that, of the three components of Δ, only the Lensing component has the term β. This can be seen by joining equations (2.53) and (2.56):

$$\begin{aligned}
\Delta_i^L &= -(2-5s)\kappa_i \\
&= -(2-5s)\left(-\frac{1}{2}\beta\nabla^2\psi_i\right) \\
&= \left(1 - \frac{5}{2}s\right)\beta\nabla^2\psi_i\,.
\end{aligned} \tag{B.2}$$

So, when we derivate them in order of β, the terms with no lensing dependency will be null. The expression then simplifies to

$$\frac{\partial C_{\ell,ij}}{\partial\beta} = \int \frac{1}{k}\mathcal{P}(k)\frac{\partial}{\partial\beta}\left[(\Delta_i^M + \Delta_i^{RSD})\Delta_j^L + \Delta_i^L(\Delta_j^M + \Delta_j^{RSD}) + (\Delta^L)^2\right]\, dk\,.$$

If we now substitute equation (B.2) in the expression, we obtain

$$\begin{aligned}\frac{\partial C_{\ell,ij}}{\partial\beta} &= \int \frac{1}{k}\mathcal{P}(k)\frac{\partial}{\partial\beta}\left[(\Delta_i^M+\Delta_i^{RSD})\Delta_j^L+\Delta_i^L(\Delta_j^M+\Delta_j^{RSD})+(\Delta^L)^2\right]\,dk \\ &= \int \frac{1}{k}\mathcal{P}(k)\frac{\partial}{\partial\beta}\left[\left(1-\frac{5}{2}s\right)\beta\left((\Delta_i^M+\Delta_i^{RSD})\nabla^2\psi_j+\nabla^2\psi_i(\Delta_j^M+\Delta_j^{RSD})\right)\right. \\ &\qquad\left.+\left(1-\frac{5}{2}s\right)^2\beta^2(\nabla^2\psi_i\cdot\nabla^2\psi_j)\right]\,dk \\ &= \int \frac{1}{k}\mathcal{P}(k)\left[\left(1-\frac{5}{2}s\right)\left((\Delta_i^M+\Delta_i^{RSD})\nabla^2\psi_j+\nabla^2\psi_i(\Delta_j^M+\Delta_j^{RSD})\right)\right. \\ &\qquad\left.+\left(1-\frac{5}{2}s\right)^2 2\beta(\nabla^2\psi_i\cdot\nabla^2\psi_j)\right]\,dk\,.\end{aligned}$$

Considering that we are using the ΛCDM model, the fiducial value of β is 1, which implies that $\Delta_i^L=\left(1-\frac{5}{2}s\right)\nabla^2\psi_i$. So,

$$\begin{aligned}\frac{\partial C_{\ell,ij}}{\partial\beta} &= \int \frac{1}{k}\mathcal{P}(k)\left[\left(1-\frac{5}{2}s\right)\left((\Delta_i^M+\Delta_i^{RSD})\nabla^2\psi_j+\nabla^2\psi_i(\Delta_j^M+\Delta_j^{RSD})\right)\right. \\ &\qquad\left.+\left(1-\frac{5}{2}s\right)^2 2(\nabla^2\psi_i\cdot\nabla^2\psi_j)\right]\,dk \\ &= \int \frac{1}{k}\mathcal{P}(k)\left[(\Delta_i^M+\Delta_i^{RSD})\Delta_j^L+\Delta_i^L(\Delta_j^M+\Delta_j^{RSD})+2\Delta_i^L\Delta_j^L\right]\,dk\,.\end{aligned}$$

It might look like we're going nowhere, but if we now use the simple trick of adding and subtracting the terms: $(\Delta^M)^2$, $(\Delta^{RSD})^2$, $\Delta_i^M\Delta_j^{RSD}$ and $\Delta_i^{RSD}\Delta_j^M$; then the expression simplifies to

$$\begin{aligned}\frac{\partial C_{\ell,ij}}{\partial\beta} &= \int \frac{1}{k}\mathcal{P}(k)\left[\Delta_i^G\Delta_j^G+\Delta_i^L\Delta_j^L-(\Delta^M)^2-(\Delta^{RSD})^2-\Delta_i^M\Delta_j^{RSD}-\Delta_i^{RSD}\Delta_j^M\right]\,dk \\ &= \int \frac{1}{k}\mathcal{P}(k)\left[\Delta_i^G\Delta_j^G+\Delta_i^L\Delta_j^L-(\Delta_i^M+\Delta_i^{RSD})(\Delta_j^M+\Delta_j^{RSD})\right]\,dk \\ &= \int \frac{1}{k}\mathcal{P}(k)\Delta_i^G\Delta_j^G\,dk+\int \frac{1}{k}\mathcal{P}(k)\Delta_i^L\Delta_j^L\,dk-\int \frac{1}{k}\mathcal{P}(k)\left(\Delta_i^M+\Delta_i^{RSD}\right)\left(\Delta_j^M+\Delta_j^{RSD}\right)\,dk\end{aligned}$$

$$\boxed{\frac{\partial C_{\ell,ij}}{\partial\beta}=C_{\ell,ij}^G+C_{\ell,ij}^L-C_{\ell,ij}^{M+RSD}}\,. \tag{B.3}$$

So, from equation (B.3) we can tell that, by calculating the values of: $C_{\ell,ij}^G$ (which includes Matter density, RSD and Lensing); $C_{\ell,ij}^L$ (which only has the Lensing term); and of $C_{\ell,ij}^{M+RSD}$ (which includes both matter density and RSD terms), we have enough information to determine $\partial C_{\ell,ij}/\partial\beta$ analytically.

B.2. Analytic derivative of $\partial C_\ell / \partial A_s$

From equations (3.10) and (3.8), one can write the angular power spectrum as

$$\begin{aligned} C_{\ell,ij} &= \int \frac{1}{k} \mathcal{P}(k) \Delta_{\ell,i}(k) \Delta_{\ell,j}(k)\, dk \\ &= \int \frac{1}{k} A_s \left(\frac{k}{k_0}\right)^{n_s - 1} \Delta_{\ell,i}(k) \Delta_{\ell,j}(k)\, dk\,. \end{aligned} \tag{B.4}$$

Therefore, the derivation of $\partial C_\ell / \partial A_s$ can be simplified to

$$\begin{aligned} \frac{\partial C_{\ell,ij}}{\partial A_s} &= \frac{\partial}{\partial A_s} \int \frac{1}{k} A_s \left(\frac{k}{k_0}\right)^{n_s - 1} \Delta_{\ell,i}(k) \Delta_{\ell,j}(k)\, dk \\ &= \int \frac{1}{k} \left(\frac{k}{k_0}\right)^{n_s - 1} \Delta_{\ell,i}(k) \Delta_{\ell,j}(k) \frac{\partial A_s}{\partial A_s}\, dk \\ &= \int \frac{1}{k} \left(\frac{k}{k_0}\right)^{n_s - 1} \Delta_{\ell,i}(k) \Delta_{\ell,j}(k)\, dk \end{aligned}$$

$$\boxed{\frac{\partial C_{\ell,ij}}{\partial A_s} = \frac{C_{\ell,ij}}{A_s}}\,. \tag{B.5}$$