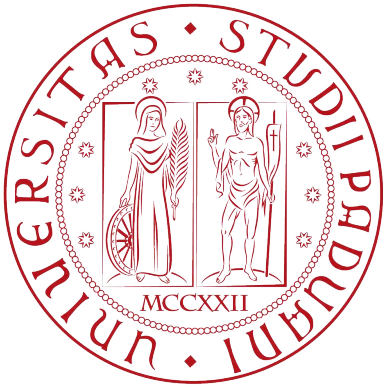


Department of Physics and Astronomy "Galileo Galilei"
University of Padua

Data analysis of Euclid's NISP instrument: from the first data to the study of cosmological parameters

PhD Course in Physics

Bianca De Caro

XXXVI Cycle

Advisors: Prof. Alessandro Renzi

Co-Advisors: Prof. Luca Stanco

Thesis written with the financial contribution of Istituto Nazionale di Fisica Nucleare

Abstract

Over the past two decades, the standard cosmological model, the so-called ΛCDM model, has achieved remarkable success. We have been able to accurately measure the power spectrum of the cosmic microwave background (CMB) as observed by the Planck experiment. Additionally, the galaxy clustering measurements from Large Scale Structure (LSS) surveys are compatible with the standard model. Despite ΛCDM's fit to our observations, tensions arise in measurements between the different datasets. These tensions could be attributed to either new physics or to unknown systematical effects, opening new scientific objectives for future experiments.
Extensions to ΛCDM model could provide more insights into the true nature dark energy. Unraveling the mysteries of this dark sector stands as the central question for cosmology in the coming decades. One of the prominent analyses in this direction is to correlate the data from CMB and LSS experiments, such as *Euclid*. The idea, as it will be discussed in this thesis, is to exploit the additional information provided by this cross-correlation in order to reach more stringent constraints on the nature of the dark sector. In particular, the work discussed in this thesis is based on the analysis of the Integrated Sachs-Wolfe (iSW) effect. This effect arises at late time from the interaction between the CMB photons and the evolving gravitational potential, which is traced by the LSS, e.g. the distribution of galaxies. The iSW, which is imprinted in the CMB spectrum at very large scales, encodes information on the nature of dark energy. Unfortunately, the iSW signal is difficult to detected from the CMB aloe, as it is subdominant with respect to the other temperature anisotropies. We can exploit the correlation with the distribution of galaxies. In particular, this thesis aims to forecast the detection significance of the iSW effect from the cross-correlation of the *Euclid* galaxy distribution, detected by the Near Infrared Spectrometer and Photometer (NISP) instrument, with the CMB temperature perturbations. In this work, we first develop a promising estimator of the temperature-galaxies cross-correlation observable, based on a type of spherical wavelets, and we compare it to the widely used harmonics estimator. Second, we validate the likelihood approximation that will be implemented on the official Euclid analysis for the estimation of cosmological parameters by means of the cross-correlation signal.
This thesis aims to prove the robustness of the cross-correlation analysis between the large-scale CMB temperature anisotropies and the *Euclid* galaxy distribution, in preparation for the first data arriving in the next years.

Contents

Introduction

In the last twenty years, the large amount of data gathered by the cosmological experiments has allowed us to consolidate our understanding of the fundamental properties of the Universe. This collective knowledge is summarized in the concordance cosmological model, the so-called ΛCDM. Despite the remarkable success of this model in fitting the data, there are several puzzles to be solved. One of these is about the nature of dark matter and the dark energy.
In the coming years, galaxy surveys such as *Euclid* will extensively map the distribution of matter across the extragalactic sky. These surveys will employ both the distribution of galaxies and the subtle effects of weak gravitational lensing as tools for cosmological investigation. We will test gravity on cosmological scales and shed light on the underlying causes of the Universe's accelerated expansion.
Both the Large Scale Structures (LSS) and the Cosmic Microwave Background (CMB) surveys will yield a substantial amount of data, and understanding how to better exploit it to extract cosmological information is a formidable challenge. One promising analysis for studying the evolution of the late Universe is the cross-correlation between LSS and CMB data. Indeed, at low redshift, CMB photons interact with LSS structures, leaving an imprint on the CMB temperature and polarization fields. One of these effects is the so-called "*integrated Sachs-Wolfe*" (iSW) effect, generated when CMB photons interact along their path with evolving gravitational potential perturbations. The photons, climbing out the time-varying gravitational potential, gain energy overall, leading to temperature anisotropies on the CMB temperature field at large scales. This effect encodes information about dark energy. Detecting the iSW effect from the CMB temperature field alone is challenging due to its weak and suppressed signal, particularly at large scales where the CMB signal is dominated by cosmic variance, an intrinsic source of uncertainty in any CMB measurements. However, cross-correlating the CMB temperature field and LSS tracer enhances our detection capability of the iSW.

The primary focus of this thesis is the analysis of the iSW effect through the cross-correlation of CMB temperature anisotropies at large scales and the distribution of galaxies observed by the Euclid Near Infrared Spectrometer and Photometer (NISP) instrument. The aim is to provide realistic results regarding the detection of the iSW effect based on the specifications of the Euclid photometric survey. This thesis summarizes the work I have undertaken during my Ph.D. as a member of the Euclid Consortium collaboration, particularly within the CMBX Science Working Group.
The first major topic of this thesis involves validating the likelihood approximation of CMB temperature-galaxies cross-correlation analysis. The objective is to infer the probability distribution of a set of Dark Energy cosmological parameters and assess biases induced by the photometric settings of the Euclid survey. We perform all the analyses within the framework of the ΛCDM model and minimal extensions to it, using the prescribed values for the cosmological parameters for the *Euclid* forecast. The likelihood approximation we aim to validate is meant to be implemented in the official Euclid analysis.
The second main topic is the estimates of the cross-correlation observable from a set of two-dimensional simulated maps of the CMB temperature and galaxy distribution. We propose a promising estimator based on a type of spherical wavelets known as "*needlets*". Like spherical harmonics, needlet functions are defined over the two-dimensional sphere and have been studied for CMB data analysis. The peculiar properties of needlets can be exploited to improve the study of the systematic effects that could affect the signal.
They have also been explored as a tool for building a new estimator for the cross-correlation between CMB (both temperature and lensing) and galaxy distribution (see [1] and [2]). In this thesis, we validate the needlet estimator across a set of simulated maps, both in the case of full sky and masked observations. We test the significance of the iSW signal detection and recover the probability distribution of the dark energy density parameter. We apply the same analysis to simulated maps of the Euclid survey.

Outline of the thesis This thesis is structured as follows. Chapter 1 provides an overview of the standard cosmological model and the physics underlying the cosmological probes relevant to this work, namely the CMB and Galaxy Clustering. Chapter 2 offers an in-depth look at the Euclid mission, including its scientific objectives and the precision of measurements the mission aims to achieve during its survey lifetime. The NISP instrument is introduced in chapter 3, detailing its spectroscopy and photometric survey characteristics. Chapter 4 delves into the cross-correlation analysis between CMB temperature and the distribution of galaxies for detecting the iSW effect. It explains how the cross-correlation observable is estimated from sky

maps and includes projections for the expected iSW detection in the Euclid survey. Chapter 5 presents the methodology and results for validating the likelihood approximation, focusing on Dark Energy cosmological parameter estimations. The last chapter 6 introduces needlet functions, describes the estimator for the cross-correlation observable and its associated variance, and validates this estimator across various sets of simulations, including those simulating the Euclid survey.

Chapter 1

The Standard Model of Cosmology

This Chapter aims to review the basics of the Standard Model in Cosmology and the physics of the key cosmological observables, i.e. the Cosmic Microwave Background (CMB) and the distribution of galaxies as probes of the Large Scale Structures (LSS) of the Universe.

Unless otherwise specified, in this chapter I am following Weinberg [3] and Durrer [4], adopting units such that

$$\hbar = c = k_B = 1\,.$$

1.1 Introduction to the Standard Model of Cosmology

The Standard Model of Cosmology rests on two assumptions:

- the Cosmological Principle (CP) which states that the Universe is isotropic and homogeneous at sufficiently very large scales;
- the theory of General Relativity (GR) which describes the gravitational interaction at cosmological scales;

From these assumptions, the model pictures a Universe that evolved from an initial singularity and it has been expanding ever since. During its history, the Universe went through transitions from hotter and denser to cooler and dispersed states, where all the particles that composed the fluid were in thermal equilibrium. This model, known also as the Standard Big Bang model, stands on three fundamental observations: the recession of the galaxies due to the expansion of the Universe, known as *Hubble law*, the abundance of light elements, in concordance with the prediction from the model, and the Cosmic Microwave Background (CMB), a uniform and isotropic radiation that is the snapshot of the Universe at the early stage of its history.

from these observations, the standard model fits a Universe that is composed today of approximately 5% of baryonic matter (neutrons, protons, electrons) and 25% of dark matter, whose properties are not known yet. The remaining component is an unknown source of energy that cosmologists address as "dark energy". According to the ΛCDM, is related to the cosmological constant Λ and it drives the late-time acceleration of the Universe.
The small fluctuations observed in the CMB temperature field and in the distribution of matter are explained, according to the current knowledge, by the inflationary model. According to this simplest inflationary model, the early Universe passed through a stage of exponential expansion, driven by a scalar field called inflation. The quantum fluctuations of the inflaton generated Gaussian and adiabatic perturbations in the matter and radiation field, after the end of inflation. The CMB temperature anisotropies were observed by the Planck experiment with high precision, confirming the picture proposed by the concordance model [5].
Despite these achievements, some tensions between different datasets, e.g. CMB and LSS, suggest the need to explore more theories to explain the cosmic acceleration and to dig into the nature of dark energy. Some of these theories involve a modification of General Relativity - the so-called Modified Gravity (MG) - and a dark energy fluid with a dynamical equation of state.

1.2 The background geometry of the Universe

The Cosmological Principle states that on average and on sufficiently large scales[1], the Universe is homogeneous (invariant under spatial translations, i.e. the cosmological properties are the same throughout) and isotropic (invariant under spatial rotations, i.e. there are no preferred directions). It follows that a family of comoving observers[2] must observe the same points of the Universe at the same time (with a suitable choice of time). This universal time variable is called proper time. Together with the evidence from Hubble's observations that the Universe is not static, the Friedmann-Lemaître-Robertson-Walker (FLRW) metric is the most general metric describing an expanding spacetime. Using the comoving spherical coordinates $x^i = (r, \theta, \phi)$, the line element in this metric is:

$$\begin{aligned} ds^2 &= -dt^2 + a^2(t)\left[\frac{d^2r}{1-Kr^2} + (d\theta^2 + sin^2\theta d\phi^2)\right] \\ &= -dt^2 + a^2(t)\left[\frac{d^2r}{1-Kr^2} + r^2 d^2\Omega\right] \end{aligned} \tag{1.1}$$

where t is the proper time, $a(t)$ is the scale factor, that parameterizes the expansion of the Universe. The scale factor conventionally is defined in

[1] Scales bigger than a hundred Mpc

[2] A comoving observer is the one that moves with the average velocity of typical galaxies in their respective neighborhoods.

such a way that $a(t_0) = 1$, with t_0 being the present cosmic time (today). The adimensional constant K can conventionally assume the following values:

$$K = \begin{cases} -1 & \text{Open Universe} \\ 0 & \text{Flat Universe (Euclidean)} \\ +1 & \text{Closed Universe} \end{cases} \tag{1.2}$$

each one representing one of the three classes of equivalence of the possible geometries. Eq. (1.1) can be expressed in terms of the *conformal time* η, such that $d\eta = dt/a(t)$, where $a(\eta) \stackrel{\text{def}}{=} d\eta$:

$$ds^2 = a^2(\eta)\left(-d^2\eta + \left[\frac{d^2r}{1-Kr^2} + r^2d^2\Omega\right]\right). \tag{1.3}$$

An important consequence of the Universe's expansion is that light gets stretched. A photon emitted by a source at a time t_{em} with wavelength λ_{em} will be observed at a time t_{obs} with a wavelength λ_{obs} given by

$$\frac{a(t_{obs})}{a(t_{em})} = \frac{\lambda_{obs}}{\lambda_{em}} \equiv 1 + z\,, \tag{1.4}$$

where z is the *cosmological redshift*. If $a(t_{em})$ is increasing, then the wavelength λ_{obs} is redshifted and z pushed toward higher values. Redshift, as we will discuss in the following section, is considered as the third coordinate to obtain cosmological distance.

1.2.1 Distances

In cosmology, distances are not uniquely defined.
From eq. (1.3) we can obtain the distance a photon travels to us once emitted at a comoving radius r:

$$d(t) = a(t) \times \underbrace{\int_0^r \frac{d\tilde{r}}{\sqrt{1-K\tilde{r}^2}}}_{\text{f(r)}}, \tag{1.5}$$

called *proper distance*. Here we are considering photons moving in the radial direction, i.e. fixing (θ, ϕ) in one direction. The form of *f(r)* depends on K:

$$f(r) = \begin{cases} \text{arcsin}r & \text{K = 1} \\ r & \text{K = 0} \\ \text{arcsinh}r & \text{K = -1} \end{cases}^{3} \tag{1.6}$$

For sources relatively close, $z \lesssim 0.1$, to the lowest order in redshift eq. (1.4) becomes $z \simeq \frac{\dot{a}}{a}\big|_{t_0} d = H_0 d$, where $H \equiv \frac{da/dt}{a}$ is the *Hubble parameter* and H_0

[3] Note that for $r \sim 0$ f(r)$\sim r$ for all values of K

is the Hubble parameter *today* [4], at t_0 [5]:

$$H_0 = 67.7 \pm 0.5 \text{ km/s/Mpc} \qquad \text{(Planck 2018)} \tag{1.7}$$

We can define also the *comoving* Hubble rate as $\mathcal{H} \equiv \frac{\dot{a}}{a} = aH$, where $\dot{} \equiv d/d\eta$ is the derivative with respect to the conformal time [5].
For nearby objects, we have $v \sim z$, where v is the recession speed, and so $v = H_0 d$. This is known as Hubble law and was proven by Edward Hubble in 1929 [6].
It is useful to relate the proper distance to the *comoving distance* factorizing out the scale factor as $\chi = d(t)/a(t) = f(r)$, giving a quantity that does not change in time due to the expansion of space:

$$\chi = \int_{t_0}^{0} \frac{d\tilde{t}}{a(\tilde{t})} = \int_a^1 \frac{d\tilde{a}}{\tilde{a}^2 H(\tilde{a})} = \int_0^z \frac{d\tilde{z}}{H(\tilde{z})}\,. \tag{1.8}$$

The greatest comoving distance from which light could have reached us by now (assuming the Big Bang at t_i = 0) is the *comoving horizon* η:

$$\eta(t) \equiv \int_0^t \frac{d\tilde{t}}{a(\tilde{t})} = \int_a^0 \frac{d\ln\tilde{a}}{\tilde{a}H(\tilde{a})}\,. \tag{1.9}$$

Regions separated by distances greater than η are not causally connected, So, if two regions are separated by $\lambda > \eta$, they have never communicated. The quantity $(aH)^{-1}$ is the *comoving Hubble radius*, that is the approximate distance over which light can travel in the course of one expansion time. Regions separated by $\lambda > (aH)^{-1}$ are not casually connected *now*. Note that the comoving distance and the Hubble radius are not the same, since η, as we will discuss later, takes the main contribution at early times.
The comoving distance is not observable, so it is common in cosmology to rely on the so-called "standard candles", i.e. objects with known luminosity L (like the Type-Ia SN), and "standard rulers", i.e. sources with a known physical size (such as the fluctuations in the CMB temperature field and in the distribution of galaxies). From the observation of a source with luminosity L and flux $F(a) = \frac{L}{4\pi D_L^2(a)}$ we can derive the *luminosity distance* as:

$$D_L \equiv \sqrt{\frac{L}{4\pi F}} = \frac{f(r)}{a} = (1+z)f(r)\,. \tag{1.10}$$

Measuring the angle θ subtended by an object with physical size ℓ allows us to measure the *angular distance* from us to that object:

$$D_A(a) \equiv \frac{\ell}{\theta} = a(t)f(r) = \frac{f(r)}{1+z}\,. \tag{1.11}$$

The two definitions eq. (1.11) and (1.10) are related as $D_L = (1+z)^2 D_A$.

[4] One of the bigger discrepancies in cosmology measurements is the tension between measurements of H_0 from different observables, such as CMB and Type Ia Supernovae. Here I report the value from Planck analysis of CMB temperature and polarization fluctuations.

[5] We refer to the derivative respect to the physical time t as $f' \equiv df/dt$

1.2.2 Friedmann equations

In the last sections, we have discussed the geometry of the spacetime defining the FLRW metric. The dynamics of the background Universe, ruled by the scale factor, is related to the metric through the Einstein Field Equations:

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} + \Lambda g_{\mu\nu} \,, \tag{1.12}$$

where $R_{\mu\nu}$ is the Ricci tensor and $R = R^{\mu}_{\nu}$ is the Ricci scalar, both terms depend on the choice of the metric $g_{\mu\nu}$. $T_{\mu\nu}$ is the *energy-momentum tensor*; due to the symmetry of spacetime, it has the form of a perfect fluid at rest in comoving coordinates:

$$T^{\mu}_{\nu} = (\rho + \mathcal{P}) u^{\mu} u_{\nu} - \mathcal{P} \delta^{\mu}_{\nu} \,, \tag{1.13}$$

where u^{μ} is the relative four-velocity between the fluid and the observer, $\rho = \rho(t) = -T^0_0$ and $\mathcal{P} = \mathcal{P}(t) = T^i_i$ are the *energy density* and *pressure* in the rest frame of the fluid. All the other components vanish because of the requirements of isotropy and homogeneity. The last term on eq. (1.12) is the one related to the *cosmological constant* Λ, discussed in detail in the paragraph 1.2.2.
With these ingredients, the time-time element of eq. (1.12) becomes, yielding the time evolution of Hubble rate H:

$$\mathcal{H}^2 = \frac{8\pi G}{3} \rho a^2 + \frac{\Lambda}{3} a^2 - K \tag{1.14}$$

and the space components:

$$\frac{d}{d\eta} \left(\frac{\dot{a}}{a} \right) = \dot{\mathcal{H}} = -\frac{4\pi G}{3} a^2 (\rho + 3\mathcal{P}) + \frac{1}{3} \Lambda a^2 \,. \tag{1.15}$$

Eq. (1.14) and (1.15) are the so-called *Friedmann equations*. Differentiating with respect to η eq. (1.14) and inserting into (1.15), we obtain the continuity equation:

$$\dot{\rho} = -3\mathcal{H}(\rho + \mathcal{P}) \,, \tag{1.16}$$

that is a consequence of the conservation of energy and momentum, derived also from $\nabla^{\mu} T_{\mu\nu} = 0$. The cosmological fluids are assumed to be *barotropic*, namely that the pressure is an explicit function of the energy density, $\mathcal{P} = \mathcal{P}(\rho)$. In cosmology, it is common to define the adimensional constant $w \equiv \mathcal{P}/\rho$ and the *sound speed* $c_s^2 \equiv \delta\mathcal{P}/\delta\rho$. The integration of eq. (1.16) leads to $\rho \propto a^{-3(1+w)}$. If the curvature K can be neglected and the energy density is dominated by one component at the time, inserting it into Friedmann's equations gives:

$$\begin{aligned} a &\propto \eta^{\frac{2}{1+3w}} \propto t^{\frac{2}{3(1+w)}} \\ \mathcal{H} &\propto \frac{2}{(1+3w)\eta} \,. \end{aligned} \tag{1.17}$$

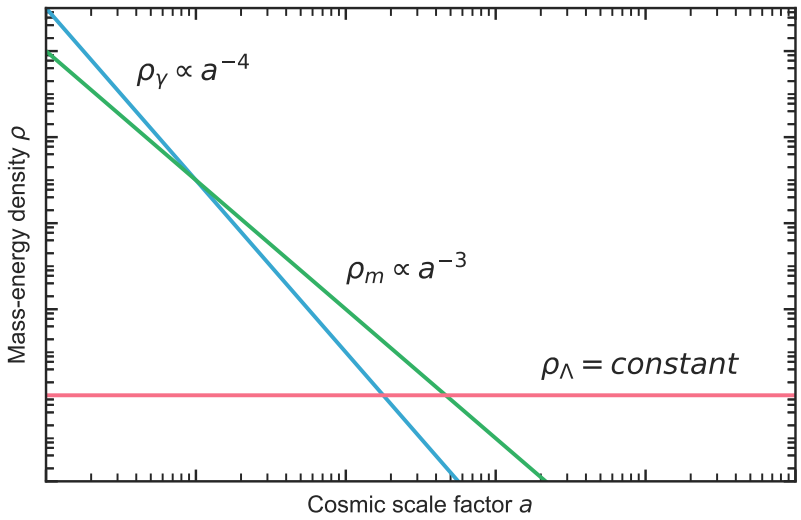


Figure 1.1: Evolution of the energy density ρ as a function of the scale factor a, for different species: matter, radiation and dark energy. The underlying physical model is ΛCDM.

Thus the scale factor and the Hubble scale evolve differently depending on the parameter of the equation of state of the fluids. In Fig. 1.1 it is shown how the energy density for different species evolves as a function of the scale factor.

It is useful to report the case of different species contributing to the cosmological fluids:

- **Matter**: Non-relativist particles (where the mass m is much bigger than the momentum p), like baryons (protons, neutrons, and electrons), dark matter, that have $P = 0$ or negligible with respect to the energy density. In this case $w = 0$ and $\rho \propto a^{-3}$; the scale factor evolves as $a \propto \eta^2 \propto t^{\frac{2}{3}}$, whereas $\mathcal{H} \propto 2/\eta$.
- **Radiation**: For this species $P = \frac{1}{3}\rho$ and thus $w = \frac{1}{3}$ and $\rho \propto a^{-4}$. This is the case of non-relativistic particles ($m \ll p^2$), like photons and light neutrinos. In this case, the scale factor scales as $a \propto \eta \propto t^{\frac{1}{2}}$ and $\mathcal{H} \propto 1/\eta$.
- **Dark Energy**: Any component with $P = -\rho$ and $w = -1$. From eq. (1.14) we see that $\rho > 0$ and so the pressure has to be negative; the solution is $a \propto \exp(Ht)$ and the Hubble scale is constant. In the concordance model, this component, called Dark Energy (DE), is sourced by the cosmological constant Λ added at the r.h.s. of Einstein's equations.

From eq. (1.14), with $K = 0$ and $\Lambda = 0$ we can define the *critical density* as:

$$\rho_{crit} \equiv \frac{3\left(\frac{\mathcal{H}}{a}\right)^2}{8\pi G} \tag{1.18}$$

and the dimensionless *density parameters* for different species as $\Omega_i(t) = \frac{\rho_i(t)}{\rho_{crit}(t)}$. The present value of the critical density ρ^0_{crit} is approximately 10^{-27} kg/m^3. With these new dimensionless quantities, we can rewrite eq. (1.14) as, including all the species we discussed so far:

$$H^2 = H_0^2 \left[\Omega_{r0} a^{-4} + \Omega_{m0} a^{-3} + \Omega_{K0} a^{-2} + \Omega_\Lambda\right] . \tag{1.19}$$

From Fig. 1.1 we can separate three epochs in cosmic history, depending on the most abundant species: *radiation domination*, *matter domination* and *Dark Energy domination*.
The curvature density can be calculated from eq. (1.19) at $a = a_0$. Thus $\Omega_{K0} = 1 - \Omega_{\rho 0} - \Omega_{m0} - \Omega_{\Lambda 0}$.

Cosmological constant and Dark Energy The cosmological constant Λ was first introduced by Einstein in order to obtain a static solution to the field equations. In fact, adding a term $\propto \Lambda g_{\mu\nu}$ in the l.h.s. of Einstein's equations does not spoil the symmetry properties of the spacetime and prevents the scale factor from growing in the case of a closed Universe ($K > 0$). But this model is unstable against small perturbations and so it was soon abandoned in favor of expanding Universe solutions, supported also by Hubble's observations of the recession of galaxies.
At the end of the '90s, the observations of the redshift-distance relation for Type-Ia supernovae [7] indicated that the expansion of the Universe is accelerating. To explain the acceleration, Einstein's equation where rewritten as in eq. (1.12), adding the cosmological constant in the r.h.s and interpreting it as a particular kind of fluid. Isolating its contribution to the total energy-momentum tensor, the energy density and the pressure related to Λ are $\rho_\Lambda = -\mathcal{P}_\Lambda = \Lambda/8\pi G$, so the equation of state is parameterized by $w = -1$. In a flat Universe dominated by this component, eq. (1.14) reads:

$$\begin{aligned} a &\propto e^{Ht} = e^{\sqrt{\frac{\Lambda}{3}}t} \\ H &= \sqrt{\frac{\Lambda}{3}}, \end{aligned} \tag{1.20}$$

and thus $\ddot{a} = \frac{\Lambda}{3} a$. So, a fluid with negative pressure and sourced by a cosmological constant, proportional to the metric $g_{\mu\nu}$, acts as a repulsive force and, according to the sign of Λ, produces an acceleration or deceleration. This energy is commonly referred to as *Dark Energy*.
In the past years were proposed different scenarios that could explain

the cosmic acceleration, in extension of the standard cosmological model. One of these involves a dark energy with a time-evolving EoS parameter, $w = w(a)$, parametrizes as [8] [9] [6]:

$$w^{CPL}(a) = w_0 + w_a(1-a)\,, \tag{1.21}$$

such as $w(a_0) = -1$ at late time. The concordance model is recovered for $w_0 = -1$ and $w_a = 0$.

1.3 The Perturbed Universe

The assumption of homogeneity and isotropy is a pillar in building the Standard Cosmological Model. However, observations of the CMB and matter fields show small perturbations of the order of 10^{-5} both at large scales and small scales. According to the concordance model, these fluctuations are produced by the *inflationary mechanism*, a stage at the very early time of the Universe characterized by an exponential cosmic expansion.

1.3.1 Perturbative General Relativity

Let's think of the *physical* (*perturbed*) spacetime as close to a symmetric *background* spacetime, the FRW Universe described in the previous section. This means that we can write the metric of the perturbed spacetime as:

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu}\,, \tag{1.22}$$

where $\bar{g}_{\mu\nu}$ is the metric of the background spacetime (the overbar will be dropped eventually) and $\delta g_{\mu\nu}$ is small. At first-order (or *linear*) perturbation we drop all the terms that contain products of such small quantities and their derivatives.
We are interested in deriving the evolution of perturbation in time, so we start from the perturbation to Einstein's equations at first order:

$$\delta G_{\mu\nu} = 8\pi G \delta T_{\mu\nu}\,. \tag{1.23}$$

Assuming the background metric to be spatially flat, the most general expression for the perturbed line element is:

$$ds^2 = a^2(\eta)\{-(1+2A)d\eta^2 - 2B_i d\eta dx^i + [(1-2D)\delta^k_{ij} + 2E_{ij}]dx^i dx^j\}\,. \tag{1.24}$$

where A and D are two scalars, B_i is a 3-D vector and E_{ij} is a tensor that is symmetric and traceless. The vector B_i and the tensor E_{ij} can be split into scalar, vector and tensor parts. In this chapter, we will not discuss vector and tenor perturbations, since we are only interested in the density

[6] CPL stands for Chevallier and Polarski and Linder, who first proposed this model

inhomogeneity originated by the scalar perturbations of the metric. It is worth mentioning that tensor perturbations are an important prediction of the inflationary model and describe gravitational waves.
The terms in eq. (1.24) are not uniquely defined but depend on the *gauge choice*. Strictly speaking, it means to fix the coordinate system and thus the coordinate transformations between such coordinate system and another in the perturbed spacetime. Another possibility is to work with variables that are invariant under any *gauge transformations* [10].
In this work, we use the gauge freedom to set the value of the scalar perturbations B and E equal to zero [7]. This choice is called *Conformal-Newtonian* gauge (or *longitudinal*), which is a commonly used gauge in cosmological perturbation theory[8]. Eq. (1.24) becomes:

$$ds^2 = a^2(\eta)\{-(1+2\phi)d\eta^2 + [(1-2\psi)\delta^k_{ij}]dx^i dx^j\}\,, \tag{1.25}$$

Where we have defined the two scalar $A = \phi$ and $D = \psi$. In the Newtonian limit, ϕ corresponds to the Newtonian gravitational potential.
The perturbations to the energy-momentum tensor $\delta T^{\mu\nu}$ are written in terms of the perturbations to the density, pressure and velocity [11]:

$$\rho = \bar{\rho} + \delta\rho, \quad p = \bar{p} + \delta p, \text{ and } u^i = \bar{u}^i + \delta u^i = \delta u^i \equiv \frac{1}{a} v_i\,. \tag{1.26}$$

The δu^0 component is not independent, because of the constraint $u_\mu u^\mu = -1$.
The metric and energy-momentum tensor perturbations yield the perturbed Einstein's equations (1.23), going in Fourier space:

$$3\mathcal{H}(\dot{\psi} + \mathcal{H}\phi) + k^2\psi = -4\pi G a^2 \delta\rho \tag{1.27a}$$

$$k^2(\dot{\psi} + \mathcal{H}\phi) = 4\pi G a^2(\bar{\rho} + \bar{\mathcal{P}})\theta \tag{1.27b}$$

$$\ddot{\psi} + \mathcal{H}(\dot{\phi} + 2\dot{\psi}) + (2\dot{\mathcal{H}} + \mathcal{H}^2)\phi + \frac{k^2}{3}(\psi - \phi) = 4\pi G a^2 \delta\mathcal{P} \tag{1.27c}$$

$$k^2(\psi - \phi) = 12\pi G a^2(\bar{\rho} + \bar{\mathcal{P}})\sigma\,, \tag{1.27d}$$

where $\theta = ik^j v_j$ is the divergence of the fluid velocity in Fourier space, and σ is related to the traceless component of T^i_j and is defined as $\sigma \equiv 2\Pi\bar{\mathcal{P}}/(\bar{\rho} + \bar{\mathcal{P}})$, where Π is called the anisotropic stress tensor. All the terms in the r.h.s of eq. (1.27) are a sum of all the species that contribute to the energy-momentum tensor. For a perfect fluid, Π is equal to zero and thus $\psi = \phi$. In reality, in the early Universe, neutrinos decoupled producing an anisotropic pressure[9]. This term, however, is negligible for non-relativistic

[7] Here B and E are the scalar components of E_{ij} and B_i

[8] Other popular gauges are the *synchronous gauge* (see [11]), often used in numerical work, and the *spatially-flat gauge*, a convenient gauge for computing inflationary perturbations.

[9] The anisotropic stress in eq.(1.27) is related to the quadrupole of the photons, which is negligible, and of neutrinos.

particles, such as cold dark matter, and for photons before decoupling. So, for matter perturbations, Π is negligible since the neutrino contribution is small. Also for CMB perturbations the anisotropic stress is negligible, since the photons are tight couplet to baryons, and thus the radiation is isotropic. To close the hierarchy of perturbed Friedmann's equations, the first-order continuity equations are, defining the *relative energy density perturbations* $\delta = \frac{\delta\rho}{\bar{\rho}}$:

$$\dot{\delta} = -(1+w)(\theta - 3\dot{\psi}) - 3\mathcal{H}\left(c_s^2 - w\right)\delta \tag{1.28a}$$

$$\dot{\theta} = -\mathcal{H}(1-3w)\theta - \frac{\dot{w}}{1+w}\theta + \frac{c_s^2}{1+w}k^2\delta - k^2\sigma + k^2\phi\,. \tag{1.28b}$$

The first of eq. (1.28) is referred to as *energy continuity equations* and the second one is the first *Euler equation*.

1.3.2 Initial conditions

Before solving eq. (1.28), we need to set the initial conditions.
The standard model we have discussed is not able to explain some theoretical and observational problems. One of these, for example, is known as *horizon problem*. From observations, the CMB appears to be isotropic (at the level of 10^{-5} precision) at any scale: how can be explained that points separated by an angular distance larger than the horizon today share the same temperature, being casually connected? Moreover, from the observations of the CMB, the Universe appears to be flat and with a density very close to the critical one ρ_{cri} [5]. We also know from Friedmann's equations that the density parameters grow with time: this means that at very early times the Universe was supposed to be even flatter, requiring unnatural initial conditions. This is known as the *flatness problem*.
These problems can be solved by postulating the existence of a period of accelerating expansion in the early Universe, the so-called *inflation* [12]. In fact, an exponential increase in the size of the Universe stretches the geometry and washes out any curvature, solving the flatness problem. Moreover, inflation leads to a decrease of the Hubble radius: so, regions that were smaller than the Hubble radius before inflation, after inflation become larger than the Hubble radius and thus disconnected, but still share the same temperature.
Accelerated expansion happens if the Universe is dominated by a fluid with negative pressure, such that, from the second Friedmann's equation (eq. (1.15)):

$$\frac{a''}{a} = -\frac{4\pi G}{3}(\rho + 3\mathcal{P}) > 0 \Rightarrow \mathcal{P} < -\frac{1}{3}\rho \Rightarrow w < -\frac{1}{3}\,. \tag{1.29}$$

Let's consider the case in which $w = -1$ and thus H = const[10]: these conditions imply that $a(t) \propto \exp(Ht)$[11]. So the Universe passed through a phase of exponential expansion, that lasted "enough" to generate the perturbations we see today and to solve the *horizon* and *flatness* problems.
The candidate to fulfill these requirements is a scalar field $\phi(t)$, called *inflaton*, subjected to a potential $V(\phi)$[12]. This scalar field behaves as a perfect fluid with:

$$\begin{aligned} T^0_0 &= -\rho_\phi = -\left(\frac{1}{2}\phi'^2 + V(\phi)\right) \\ T^i_j &= \mathcal{P}_\phi \delta^i_j = \left(\frac{1}{2}\phi'^2 - V(\phi)\right)\delta^i_j\,, \end{aligned} \tag{1.30}$$

and it moves like a damping oscillator with the following equation of motion:

$$\phi'' + 3H\phi' + \frac{dV}{d\phi} = 0\,, \tag{1.31}$$

where the damping term is due to the Hubble expansion. From eq. (1.30), in order to have $\mathcal{P}_\phi \simeq -\rho_\phi$, we need $V(\phi) \gg \frac{1}{2}\phi'^2$, that is that the energy density of the Universe must be dominated by the potential energy of the scalar field. This implies that also inflation is nearly constant during inflation. Thus Friedmann's equation during inflation becomes:

$$H^2 = \frac{8\pi G}{3}\left[V(\phi) + \frac{1}{2}\phi'^2\right] \simeq \frac{8\pi G}{3}V(\phi)\,, \tag{1.32}$$

which means that $V(\phi)$ changes slowly as H must be nearly constant. These models, according to which the potential and the scalar field are slowly changing, are referred to as *slow-roll inflation*.
It is customary to parameterize the inflationary potential by means of the so-called *slow-roll parameters*, which provide the conditions for the inflation to take place [13]:

$$\begin{aligned} \epsilon &\equiv -\frac{H'}{H^2} = \frac{3}{2}\frac{\phi'^2}{H^2} \simeq \frac{1}{16\pi G}\left(\frac{V_{,\phi}}{V}\right) \\ \eta_V &\equiv \frac{1}{8\pi G}\left(\frac{V_{,\phi\phi}}{V}\right)\,. \end{aligned} \tag{1.33}$$

[10] Actually, the Hubble parameter can not be exactly constant, but slowly changing, because at some point the inflation must end and $a(t)$ stops growing exponentially.

[11] This is known as *de Sitter phase*, but acceleration can be obtained with other conditions.

[12] Over the years, a lot of different models have been proposed, all of them including one or more scalar fields and an associated potential. This is known as *inflationary zoology* (see, for example, reference [13]). In this work, we will focus on single-field inflation and a nearly flat potential.

[13] Here $V_{,\phi}$ and $V_{,\phi\phi}$ are the first and the second derivative of V with respect to ϕ, respectively.

The condition $\epsilon \ll 1$ ensures that the Hubble parameter changes slowly, while $\eta_V \ll 1$ means that the velocity of $V(\phi)$ does not increase as long as the inflation lasts.
When the kinetic energy of the field is not negligible anymore and the field starts oscillating around the minimum of the potential, the conditions set by the slow-rolls parameters are not valid anymore and the inflation ends. Physically, the inflaton decays into lighter particles which acquire energy. This phase is called *reheating*.
These oscillations we have just mentioned are the ones responsible for the large-scale density perturbations in the matter and radiation fields after inflation. Let's consider some small perturbations of the scalar field $\delta\phi(\mathbf{x}, t)$ such that:

$$\phi(\mathbf{x}, t) = \phi(t) + \delta\phi(\mathbf{x}, t) = \phi(t - \delta t(\mathbf{x}))\,. \tag{1.34}$$

This means that inflation does not acquire the same value at a given time t in all the space, but acquires different values from one spatial point $\mathbf{x}$ to another, yielding fluctuations in the fields. Expanding at first order and solving eq. (1.31), we find that the super-horizon solution is, in Fourier space:

$$|\delta\phi| = \frac{H}{\sqrt{2k^3}} \left(\frac{k}{aH}\right)^{\frac{3}{2}-\nu} \quad \text{with } \nu \simeq \frac{3}{2} + 3\epsilon - \eta_V\,, \tag{1.35}$$

Thus super-horizon perturbations are constant during inflation. Eventually, these perturbations re-enter the horizon, after being stretched from quantum to cosmological scales as the Universe expands, and become the initial conditions for the density perturbations.
The power spectrum of eq. (1.35) is:

$$\mathcal{P}_{\delta\phi}(k) = \left(\frac{H}{2\pi}\right)^2 \left(\frac{k}{aH}\right)^{3-2\nu}\,. \tag{1.36}$$

To relate eq. (1.36) to the power spectrum of the perturbations that come out after inflation, we need to define the gauge-invariant quantity $\mathcal{R} = \psi + \mathcal{H}\frac{\delta\phi}{\dot{\phi}}$, called *comoving curvature perturbation*. For super-Hubble scales and adiabatic perturbations[14], we have:

$$\mathcal{R} \simeq \psi + \mathcal{H}\frac{\delta\phi}{\dot{\phi}} \simeq \psi + \mathcal{H}\frac{\delta\rho}{\dot{\rho}}\,. \tag{1.37}$$

[15] In this way, we can relate a perturbation with scale k of the inflation field to the density perturbation of the radiation/matter field after the scale

[14] Perturbations are adiabatic if, given density and pressure, their perturbations with respect to the background are the same: $\frac{\delta\rho}{\dot{\rho}} = \frac{\delta\mathcal{P}}{\dot{P}}$.

[15] To be precise, $\zeta = \psi + \mathcal{H}\frac{\delta\rho}{\dot{\rho}}$ is the curvature perturbation in the uniform density gauge, but in the super-Hubble limit and for adiabatic perturbation is equal to $\mathcal{R}$.

becomes sub-Hubble.
The power spectrum for the comoving curvature perturbation is:

$$\mathcal{P}_{\mathcal{R}}(k) = \frac{\mathcal{H}}{\dot{\phi}^2}\mathcal{P}_{\delta\phi}(k) = \left(\frac{\mathcal{H}^2}{2\pi\dot{\phi}}\right)^2 \left(\frac{k}{\mathcal{H}}\right)^{3-2\nu} . \tag{1.38}$$

The spectral index of the power spectrum is defined as $n_s - 1 = \dfrac{d\ln\mathcal{P}(k)}{d\ln k}$: if it is equal to one the power spectrum is scale-invariant and the amplitude of the initial fluctuations is the same on all cosmological scales. In the case of eq. (1.38), we have:

$$n_s - 1 = 3 - 2\nu = 2\eta_V - 6\epsilon\,, \tag{1.39}$$

and thus $\mathcal{P}_{\mathcal{R}}$ is nearly scale-invariant for slow-roll inflation. The power spectrum can be parameterized as

$$\mathcal{P}_{\mathcal{R}}(k) = A_s \left(\frac{k}{k_p}\right)^{n_s-1} , \tag{1.40}$$

where A_s is the scale-invariant amplitude and k_p is the pivot scale (commonly taken to be 0.05 h/Mpc or 0.02 h/Mpc).
The last step is to relate the quantum fluctuations to the initial conditions for the density perturbations. Using the continuity equation (1.16) and eq. (1.37), we can write:

$$\mathcal{R} = \phi - \frac{\delta\rho}{3(1+w)\rho} = \frac{5+3w}{3(1+w)}\phi\,. \tag{1.41}$$

Thus, for different values of w:

$$\phi_{init}^{RD} = \frac{2}{3}\mathcal{R} \text{ and } \phi_{init}^{MD} = \frac{3}{5}\mathcal{R}\,, \tag{1.42}$$

for scales that enter the horizon during radiation domination and matter domination, respectively.
Similarly to eq. (1.40), we can define the power spectrum of tensor fluctuations as:

$$\mathcal{P}_t(k) = A_t \left(\frac{k}{k_p}\right)^{n_t} . \tag{1.43}$$

Gravitational waves can be viewed as tensor perturbations of spacetime with two degrees of freedom, corresponding to the two states of polarization. Therefore, the power spectrum in eq. (1.43) is the power spectrum of the stochastic background of gravitational waves produced by inflation. The ratio of the tensor power spectrum over the curvature power spectrum $r \equiv \dfrac{A_t^2}{A_s^2}$ is called the *tensor-to-scalar ratio*. A direct measurement of this not

only will be a detection of the gravitational wave background, but also a probe of the inflation physics, since it is directly related to the energy scale of inflation, as [14]:

$$V^{1/4} \simeq \left(\frac{r}{0.07}\right)^{1/2} \times 1.8 \times 10^{16}\,\text{GeV} \tag{1.44}$$

and to the field variation during inflation in units of the Planck mass m_{Pl}[16], as [14]:

$$\frac{\Delta\phi}{m_{Pl}} \simeq 0.46 \left(\frac{r}{0.07}\right)^{1/2} . \tag{1.45}$$

1.3.3 Evolution of dark matter perturbations

Having set up the perturbed Friendmann's equations, we can calculate the evolution for the perturbations of dark matter. These are coupled to photon perturbations[17] through gravitational potential. Looking at the scale k of the perturbations, it is useful to separate two regimes: scales that are larger than the Hubble radius $k \ll \mathcal{H}$, called super-Hubble modes, and scales that are smaller than the Hubble radius $k \gg \mathcal{H}$, referred as sub-Hubble modes. The latter regime includes all the scales that enter the horizon during the radiation-dominated era or when the Universe is dominated by matter. During matter domination, the gravitational potential depends only on dark matter and is independent of radiation. The net effect is that ϕ and ψ remain constant and the density perturbations start growing due to the effect of gravitational attraction. In early times, during the radiation-dominated epoch, the radiation pressure opposes the gravitational growth, so scales that "enter the horizon"[18] during the radiation domination grows logarithmically. On the other hand, e will see that the radiation perturbations oscillate without growing.

Scales larger than the Hubble radius In this case, we can drop all the terms $\propto k$ and eq. (1.27) together with eq. (1.28) leads to:

$$\phi = \phi_{init} \quad \text{and} \quad \delta \approx -2\phi_{init}\,, \tag{1.46}$$

meaning that the metric and density perturbations outside the horizon are frozen and thus proportional to the primordial fluctuations set by inflation.

[16] $m_{pl} = (8\pi G)^{-1/2} = 2.4\times\ 10^{18}$ GeV

[17] Other perturbations, such as neutrinos, will not be discussed since it goes beyond the purpose of this thesis.

[18] With this expression we refer to a scale $\lambda_{phys}/2\pi = k_{phys}^{-1}$ that grows slower than the Hubble radius $1/\mathcal{H}$ here referred as the horizon (which, as we discussed in section 1.2.1 are not the same quantity).

Scales smaller than the Hubble radius We can identify two main regimes in which the growth rate of perturbations[19]: scales that enter the horizon before and after the epoch of *matter-radiation equality* $z_{eq} \sim 10^4$. Small scales enter the horizon before z_{eq} during the radiation-dominated Universe. During this epoch, the dark matter perturbations are determined by the gravitational potential, which is governed by radiation. In this case, solving the field equations eq. (1.27) and eq. (1.28) we obtain:

$$\phi \approx -9\phi_{init}^{\mathrm{RD}} \frac{\cos\left(k\eta/\sqrt{3}\right)}{(k\eta)^2} \quad \text{and} \quad \delta \approx -9\phi_{init}^{\mathrm{RD}} \ln\left(0.62 k\eta\right), \tag{1.47}$$

where we define ϕ^{RD} in (1.42). While ϕ oscillates and decays $\propto 1/(k\eta)^2$, the density perturbations δ grow logarithmically. For larger scales, i.e. scales that have entered the horizon well after z_{eq}, during the matter-domination, Einstein's equations reduce to:

$$\dot{\delta} + k\theta = -3\dot{\phi} \tag{1.48a}$$

$$\dot{\theta} + \mathcal{H}\theta = +k\phi \tag{1.48b}$$

$$k^2\phi = -4\pi G a^2 \rho \delta\,. \tag{1.48c}$$

The first two lines are the continuity and Euler equations for matter perturbations, the third one is the *Poisson* equation.
Combining these equations into a second-order differential equation in the density contrast, we obtain a general solution:

$$\delta = C_1 D_+(\eta) + C_2 D_-(\eta)\,. \tag{1.49}$$

where $D_+ \propto \eta^2$ is the *growing mode* and $D_- \propto \eta^{-3}$ is the *decaying mode*. The former mode, which is the relevant one for structure formation, persists while the latter mode rapidly vanishes. The gravitational potential, on the contrary, for scales that enter the horizon during the matter-domination epoch, stays constant. It is common to define the *growth factor* such that it is normalized today:

$$D(a) \equiv \frac{D_+(a)}{D_+(a=1)}\,. \tag{1.50}$$

The two regimes we have discussed allow for an analytical solution to Friedmann's equation for the dark matter perturbations and have been obtained after some assumptions. In general, what is commonly done is to separate the time dependence, carried by the growth factor $D_+(a)$, and the scale dependence, defining the transfer function $T(k)$, such that $\phi = \phi_{init} T(k) D_+(a)/a$. This function takes into account all the interactions between different components and the behavior at the horizon crossing of

[19] For the very large scales, there is also a third regime, the domination od dark energy today.

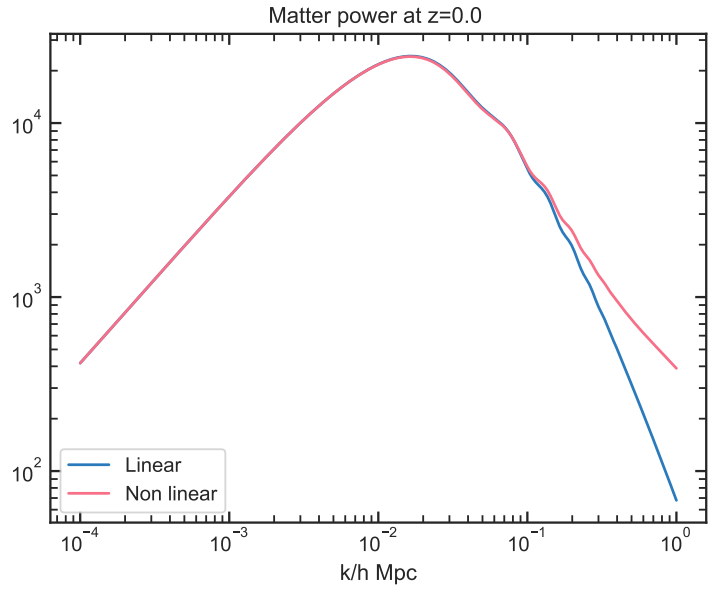


Figure 1.2: The matter power spectrum $P(k)$ at redshift $z = 0$. Here are shown both the linear and non-linear power spectrum. The underlying cosmological model is ΛCDM and the cosmological parameters are fixed at the Planck 2018 best-fit values.

different scales; it is defined such that $T(k) \to 1$ for $k \to 0$[20]. To compute it, one has to solve the complete system of Einstetin's and Boltzmann's equations. Cosmologists employ codes to solve numerically this system for each component, such as CAMB (Code for Anisotropies in the Microwave Background [16]) and CLASS (The Cosmic Linear Anisotropy Solving System [17]).
The power spectrum, i.e. the variance of the distribution of $\langle|\delta(\mathbf{k}, z)|\rangle$ is proportional to the primordial power spectrum and the square of the transfer function and the growth factor:

$$P_{\delta\delta}(k, a) \propto \mathcal{P}_{\mathcal{R}} T^2(k) D_+^2(a)\,. \tag{1.51}$$

As shown in Fig. 1.2, the large scales that enter the horizon well after k_{eq}, which is the scale of matter-radiation equality, grow proportional to the primordial power spectrum, i.e. $\propto k^{n_s}$ and $n_s = 1$. On the contrary, the growth of smaller scales, i.e. scales that enter the horizon during the radiation-dominated epoch, is suppressed and thus the transfer function is smaller than 1: the power spectrum, therefore, decreases as a function of k after the turnover scale k_{eq}.
Fig. 1.3 shows the behavior of the growth factor as a function of a: at late

[20]The slight decline of the largest scale at the time of Universe passes through the epoch of matter-radiation equality is conventionally removed [15].

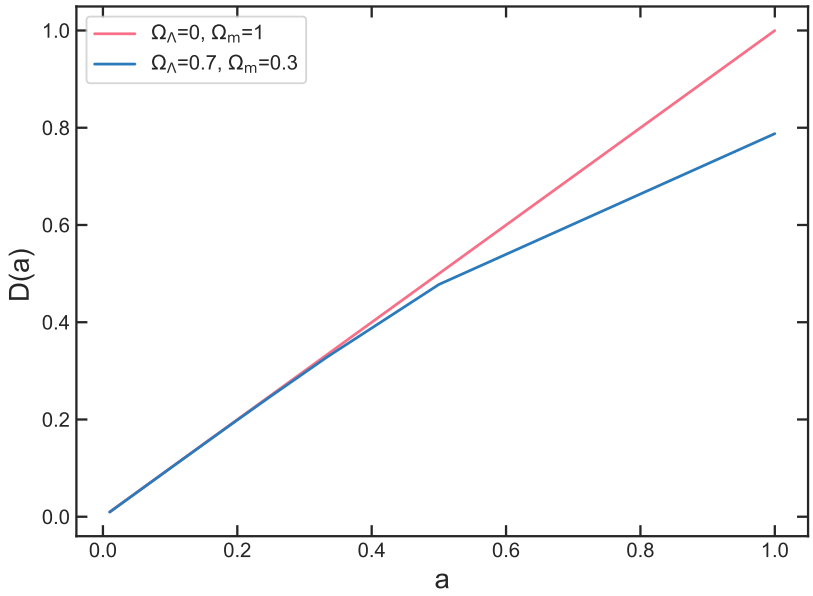


Figure 1.3: The growth factor $D_+(a)$ as a function of the scale factor a. At late time the growth factor is linear. The pink line is the growth factor in a Universe with only matter ($\Omega_m = 1$), and the blue line is in the case of a Universe with dark energy. When the dark energy starts dominating, the growth factor increases slower as a function of a. The underlying cosmological model is ΛCDM fixed at the Planck 2018 best-fit values.

time, is growing linearly, as expected since ϕ has to be constant. Furthermore, it shows the difference on the growth factor in a Universe with only matter, i.e. $\Omega_m = 1$, and in a Universe with dark energy: for $a \sim 1$ the dark energy starts dominating and $D_+(a)$ increase slower as a function of a.

As δ is no longer of the order of unity and at small scales ($k > 0.1h$ Mpc^{-1}), the perturbation regime is no longer valid. In this case, it is not possible to obtain a linear solution to the evolution of matter perturbations as the one in eq. (1.49). Different approaches have been developed in order to calculate the evolution of matter perturbations in the non-linear regime. For example, one of these approaches is to rely on numerical simulations of dark matter particles (also called N-body simulations). The matter power spectrum in the non-linear theory is shown in Fig. 1.2.

1.3.4 Evolution of photon perturbations

Calculation of the evolution of photon perturbations requires also the interaction with baryons (electrons, protons and neutrons), other than gravitational potential. At early times, photons and baryons are tightly coupled through Compton scattering. The distribution function of the photon is therefore modified by the interaction in such a way that the number of particles

is conserved. The equation that encodes this physics is the Boltzmann equation, written as:

$$\frac{df}{dt} = C[f]\,, \tag{1.52}$$

here the right-hand side contains the collision term, whereas the left-hand side is the derivative of the distribution function over time.
We are interested in the perturbation to the distribution function of photons due to the perturbed spacetime. Since the perturbations are small, we can expand eq. (1.52) up to the first order.
Starting from the perturbed photon distribution function, written as:

$$f(\mathbf{x}, p, \hat{p}, t) = \left[\exp\left\{\frac{p}{T(t)(1+\Theta(\mathbf{x}, \hat{p}, t))}\right\} - 1\right]^{-1}\,, \tag{1.53}$$

where Θ is the temperature perturbation, which could also be called $\delta T/T$, and $\hat{p}$ is the direction of the momentum of the scattered photon. Switching to Fourier space, other ingredients required to obtain the Boltzmann equation are:

- The expansion in multiples of the temperature perturbations as:

$$\Theta_l \equiv \frac{1}{(-1)^l}\int_{-1}^{1}\frac{d\mu}{2}\mathcal{P}_l(\mu)\Theta(\mu, k, \eta)\,, \tag{1.54}$$

  where $\mu \equiv \dfrac{\mathbf{k}\cdot\hat{p}}{k}$ is the cosine of the angle between the wavenumber $\mathbf{k}$ and the photon direction $\hat{p}$, describing the direction of photon propagation[21].

- The velocity of the baryons is defined as $\mathbf{u}_b(\mathbf{k}, \eta) = \dfrac{\mathbf{k}}{k}u_b(\mathbf{k}, \eta)$, which means that the bulk velocities of electrons are longitudinal and point in the same direction of $\mathbf{k}$.

- The *optical depth* encodes all the physics about the Compton scattering and is defined such that $\dot{\tau} \equiv \dfrac{d\tau}{d\eta} = -n_e\sigma_T a$. Here n_e is the number density of electrons and σ_T is the Thompson cross-section[22]. At late time, n_e is small and $\tau \ll 1$, while at early time is large.

With these definitions, the Boltzmann equation for photons is:

$$\dot{\Theta} + ik\mu\Theta + \dot{\phi} + ik\mu\psi = -\dot{\tau}\left[\Theta_0 - \Theta + \mu u_b\right]\,. \tag{1.55}$$

[21] Since $\mathbf{k}$ is parallel to the gradient of the temperature, $\mu = 1$ means that the photon is traveling in the direction along which the temperature is changing.

[22] Here we are considering the non-relativist Compton scattering since $k \ll m_e$, where m_e is the electron mass, known as Thompson scattering, with $\sigma_T \approx 6.6510^{-29}$ m^{-2}.

First of all, note from eq. (1.55) that different Fourier mode **k** are decoupled. Then we can make say that, in the absence of electron velocity, the scattering tends to uniformize the photon distribution bringing Θ toward the *monopole* Θ_0, whereas a value of u_b different from zero produces a non-vanishing dipole, while high-order terms vanish. So, in the regime of *tight coupling*, photons behave as a fluid, with a non-vanishing monopole and dipole. When Compton scattering ceases to be efficient, photons and baryons decouple and photons start *free-streaming* toward us.

Solution for large-scale anisotropies In the limit of $k \ll \mathcal{H}$, we consider scales that enter the horizon during MD. Moreover, the Compton scattering at late time becomes inefficient and thus $\tau \approx 0$. The solution to the Boltzmann equation is related to the initial conditions set by inflation:

$$\Theta_0(\mathbf{k}, \eta) = \frac{2}{5}\mathcal{R}(\mathbf{k}) = \frac{2}{3}\phi(\mathbf{k}, \eta) = -\frac{1}{3}\delta\,, \tag{1.56}$$

where the last equality comes from (1.46). The observed anisotropy is the monopole of perturbation and the gravitational redshift ϕ, so: $[\Theta_0(\mathbf{k}, \eta) + \phi(\mathbf{k}, \eta)] = -\frac{1}{6}\delta$.

Solution for small-scale anisotropies During radiation domination, baryons and photons are tightly coupled, that is $\tau \gg 1$, and, as said previously, all the multipoles higher than the dipole are washed out. In this limit eq. (1.55) becomes:

$$\left\{\frac{d^2}{d\eta^2} + \frac{R'}{1+R}\frac{d}{d\eta} + k^2 c_s^2\right\}[\Theta_0 + \phi](\mathbf{k}, \eta) = \frac{k^2}{3}\left[\frac{1}{1+R}\phi - \psi\right](\mathbf{k}, \eta)\,, \tag{1.57}$$

where $R \equiv \dfrac{3\rho_b}{4\rho_\gamma}$ is the baryon-to-photon ratio and c_s is the sound speed of the fluid, defined also as:

$$c_s(\eta) \equiv \sqrt{\frac{1}{3(1+R(\eta))}}\,. \tag{1.58}$$

Baryons make the fluid heavier and lower the sound speed. If baryon density is negligible with respect to the photon density, c_s is the sound speed for a relativistic fluid $c_s \sim \sqrt{1/3}$. Eq. (1.57) is the equation of motion of a forced, damped oscillator, where the drag term depends on the baryons and the force term is $F(k, \eta) \equiv \frac{k^2}{3}\left[\frac{1}{1+R}\phi - \psi\right](\mathbf{k}, \eta)$, encoding the expansion due to the gravitation potential. The solution to this equation in

the tight-coupled limit is:

$$\Theta_0(\mathbf{k},\eta)+\phi(\mathbf{k},\eta) = [\Theta_0(\mathbf{k},0)+\phi(\mathbf{k},0)](\cos(kr_s)) + \frac{k}{\sqrt{3}}\int_0^\eta d\tilde{\eta}[\Theta_0(\mathbf{k},\tilde{\eta})+\psi(\mathbf{k},\tilde{\eta})]\sin[k(r_s(\eta)-r_s(\tilde{\eta}))]\,. \tag{1.59}$$

Here, first of all, we have defined the *sound horizon* as $r_s \equiv \int_0^\eta d\tilde{\eta} c_s(\tilde{\eta})$, that is the comoving distance traveled by a sound wave by time η. To obtain eq. (1.59) we assumed that R is negligible [18]. This is true if we consider that the baryons density, as set from nucleosynthesis, is low: $\Omega_b h^2 \sim 0.025$ [5] and therefore negligible with respect to photon density during the radiation-dominated era. If R is larger, the effect is to enhance the amplitude of the oscillations, since baryons load the fluid that feels more pressure.
Similarly, we can solve eq. (1.57) for the dipole and find $\Theta_1(\mathbf{k},\eta) \propto \sin(kr_s)$, so it is out of phase with respect to the monopole. These solutions are usually referred to as acoustic oscillations of the photon-baryon fluid.
The tight-coupling limit we exploit to solve for the acoustic oscillations is not always valid. In reality, photons travel a mean free path λ_{MFP}, related to the optical depth, while scattering off baryons. As photons random walk, they generate anisotropic stress in the fluid and cause heat conduction across the wavelength of perturbation [19]. This leads to a damping of the perturbation that is smaller than the *diffusion length* $\lambda_D = \frac{1}{\sqrt{n_e \sigma_T H}}\frac{1}{a}$, that is the mean comoving distance a photon travels in one Hubble time. The damping[23]of the perturbations is:

$$\mathcal{D}(k) \sim \exp\left\{-\frac{k^2}{k_D^2}\right\}, \tag{1.60}$$

where the damping length is defined via:

$$k_D^{-2} \equiv \int_0^\eta \frac{d\tilde{\eta}}{6(1+R)\tau'}\left[\frac{R^2}{1+R}+\frac{8}{9}\right]. \tag{1.61}$$

1.4 Cosmological probes

Now that we have discussed the physics behind cosmological perturbations and their evolution, we can relate this knowledge to the observation and sketch the cosmological information that we can extract from those. Here we will focus on the Cosmic Microwave Background and Galaxy Clustering.

[23]This is known as Silk damping, from who first described this phenomenon in the paper at the reference [20] .

1.4.1 Cosmic Microwave Background

The Cosmic Microwave Background (CMB) is a cornerstone in modern cosmology, providing an abundance of information about the early Universe. CMB is the relic radiation left by the primordial plasma - when the Universe was a hot and dense state - and it permeates uniformly the entire cosmos. It is characterized by a blackbody spectrum with a temperature of about 3 K. In early times, as we have discussed in the previous sections, photons and baryons were the components of a tightly-coupled fluid, brought together by Compton scattering, which coupled electrons and photons, and Compton scattering, which coupled protons and electrons. Together these species were in thermal equilibrium. As the Universe expanded and the temperature dropped, electrons could begin to recombine into hydrogen atoms, lowering the rate of scattering with radiation. This is known as *recombination era*. Recombination happens at redshift $z_{rec} \sim 1400$ and $T_{rec} \sim 4000$ K, when half of the free electrons present in the fluids are ionized and captured by protons into hydrogen atoms. At some point, the electron-photon interaction rate becomes smaller than the Hubble expansion rat, since the remaining free electrons are too sparse to interact. At redshift $z_{dec} \sim 1100$ and $T_{dec} \sim 3000$ K, the interaction rate Γ_C is similar to H, i.e. $\Gamma_C(z_{dec}) \simeq H(z_{dec})$, and thus photon decouple from matter, starting to free-streaming without scattering until today. From z_{dec}, the Universe becomes transparent to radiation[24]. This radiation, which constitutes what is observed today as Cosmic Microwave Background radiation, appears to come from a spherical surface around the observer with a radius equal to the distance each photon has traveled since it was last scattered at the epoch of recombination. This surface is called *last scattering surface*. The value of z_{dec} set by the last measurements is $z_{dec} = 1089.92 \pm 0.25$ [5].
The CMB radiation shows a perfect blackbody Spectral Energy Distribution (SED), described by Planck's law:

$$B_\nu(T) = 4\pi\nu^3 \frac{1}{\exp\left(\frac{2\pi\nu}{T}\right) - 1}, \tag{1.62}$$

where h is the Planck constant, ν is the frequency and k_B is the Boltzmann constant. The FIRAS instrument on COBE satellite first measured the temperature T of the CMB in the early nineties. They showed that the CMB spectrum follows the space described in eq. (1.62), with $T = 2.72548 \pm 0.00057$ K [21].
Although very spatially uniform, CMB contains small temperature fluctuations that depend on the angle of observation of the sky. Some of these, called *primary* anisotropies, are related to the cosmological perturbations that come out from inflation and are imprinted in the photon-baryons plasma, as we discussed in 1.3.4. Other inhomogeneities in the average

[24] In other words, the optical depth τ is smaller than one.

temperature field are sourced by the interactions of the photons and matter after the decoupling. These are called *secondary* anisotropies of the CMB. The CMB temperature field that we observe is decomposed as:

$$T^{obs}(\hat{\mathbf{n}}) = T_0 \left[1 + (\boldsymbol{\beta} \cdot \hat{\mathbf{n}}) + \Theta\right] , \tag{1.63}$$

where T_0 is the average temperature [25], $\boldsymbol{\beta}$ is the velocity of the Earth with respect to the CMB rest frame, and $\Theta \equiv \frac{\Delta T}{T_0}$ are the intrinsic temperature fluctuations[26]. The dipole term $(\boldsymbol{\beta} \cdot \hat{\mathbf{n}})$, which is of the order of $\sim \mathcal{O}(10^{-3})$ in units of T_0, is usually removed by the calibration of the experiment that is observing the sky. Instead, the intrinsic fluctuations are much smaller, $\Theta \sim \mathcal{O}(10^{-5})$.
CMB is also linearly polarized, due to Thomson scattering of photons off free electrons in the *last scattering surface*. CMB polarization carries independent information to the temperature anisotropies, in particular about the theory of inflation. The dominant contribution to CMB polarization is from the scalar perturbations arising from inflation, as in the case of temperature anisotropies, and create a particular pattern of polarization called *E*-mode. Moreover, detecting *B*-mode polarization is compelling for proving the tensor perturbations and the gravitational wave background sourced by those.
The Planck satellite [5], launched by the European Space Agency (ESA) in 2009, is responsible for the most notable advancements in CMB temperature and polarization measurements. The data from the survey unveiled with unprecedented precision the pattern of CMB anisotropies and polarization. Fig. 1.4 shows the CMB temperature anisotropies map measured by Planck 2018.

CMB angular power spectrum Before discussing the nature of anisotropies in the CMB field, it is essential to establish the observable from which we can extract information.
The CMB temperature field is a function on a sphere, and thus it is natural to decompose it in spherical harmonics $Y_{\ell m}$, such that:

$$\Theta(\hat{\mathbf{n}}) = \sum_{\ell=0}^{\infty} \sum_{=-\ell}^{\ell} a^T_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}) \quad \text{where} \quad a^T_{\ell m} = \int d\Omega \Theta(\hat{\mathbf{n}}) Y^*_{\ell m} , \tag{1.64}$$

where the multipole moment ℓ and the angular scale θ between two points of the sky are related by $\theta \sim 1/\ell$, thus low-ℓs correspond to larger scales and high-ℓs to the smaller scales. We will not consider the first two multiples, $\ell = 0$ and $\ell = 1$, since are related to the monopole T_0 and the dipole $(\boldsymbol{\beta} \cdot \hat{\mathbf{n}})$,

[25] There are different definitions of the mean temperature, one of these is obtained averaging over the blackbody spectrum in eq. (1.62)

[26] Here ΔT is the difference of the observed temperature $T^{obs}(\hat{\mathbf{n}})$ and the mean temperature T_0

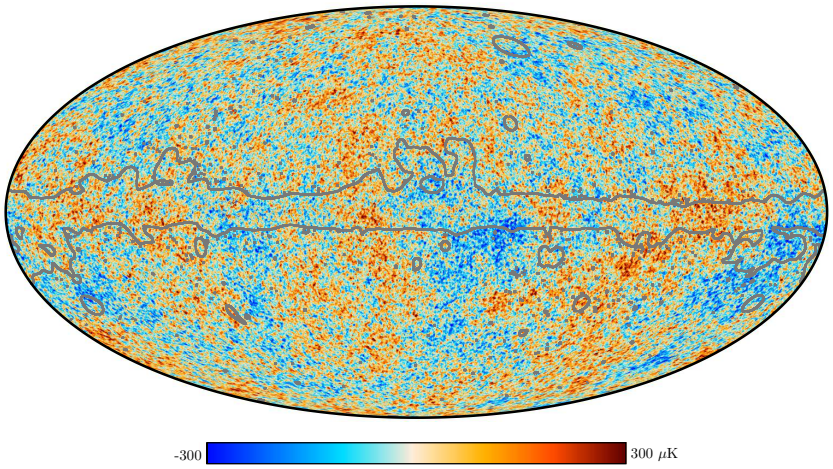


Figure 1.4: The map of the temperature anisotropies of the CMB measured by Planck 2018. The grey line shade the mask. Image credit: ESA/Planck.

as in eq. (1.63).
Since the primordial perturbations are imprinted in the CMB field as they evolve in time, also the $a^T_{\ell m}$ are Gaussian random variables. Therefore we expect the $a^T_{\ell m}$ to be drawn from a Gaussian distribution with zero expectation value[27],

$$\langle a^T_{\ell m} \rangle = 0\,, \tag{1.65}$$

and variance

$$\langle a^T_{\ell m} a^{*T}_{\ell' m'} \rangle = \delta_{\ell\ell'}\delta_{mm'} C^{\mathrm{TT}}_\ell\,. \tag{1.66}$$

C^{TT}_ℓ is the *angular power spectrum*, defined as:

$$C^{\mathrm{TT}}_\ell = \frac{1}{2\ell+1}\sum_m \langle |a^T_{\ell m}|^2 \rangle\,. \tag{1.67}$$

In eq. (1.66) we use the condition that $\Theta(\hat{\mathbf{n}})$ has real values and thus $a^T_{\ell -m} = (-1)^m a^{*T}_{\ell m}$. In addition, under the assumption of isotropy (Cosmological principle), C^{TT}_ℓ depends only on ℓ.
Under the assumptions of a Gaussian field, the angular power spectrum C^{TT}_ℓ contains all the statistical information about the CMB anisotropies. Thus from the observations we calculate the angular power spectrum and then compare it to the one predicted by theory. The C^{TT}_ℓ are the coefficients of the multipoles expansion of the two-point correlation function $C^{\mathrm{TT}}(\theta) \equiv \langle \Theta(\hat{\mathbf{n}})\Theta(\hat{\mathbf{n}}') \rangle$ at angular scale θ:

$$C^{\mathrm{TT}}(\theta) = \sum_\ell \frac{2\ell+1}{4\pi} C^{\mathrm{TT}}_\ell P_\ell(\hat{\mathbf{n}}\cdot\hat{\mathbf{n}}')\,, \tag{1.68}$$

[27] $\langle\cdot\rangle$ denotes the ensemble average over all the possible sky realizations.

where $\hat{\mathbf{n}}$ and $\hat{\mathbf{n}}'$ are two different directions such that $\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}' = \cos\theta$.
We are unable to calculate the angular power spectrum from the data as in eq. (1.67), since we observe only one realization of the stochastic CMB anisotropy field. We can replace the ensemble average with a spatial average over all the $2\ell+1$ value of m[28] and estimate the C_ℓ^{TT} as:

$$\hat{C}_\ell^{\mathrm{TT}} = \frac{1}{2\ell+1}\sum_m |a_{\ell m}^T|^2 \,, \tag{1.69}$$

$\hat{C}_\ell^{\mathrm{TT}}$ is called the *estimator* of C_ℓ^{TT} and has the property $\langle \hat{C}_\ell^{\mathrm{TT}} \rangle = C_\ell^{\mathrm{TT}}$ [29]. The variance of the estimator is given by:

$$\begin{aligned} (\Delta \hat{C}_\ell^{\mathrm{TT}})^2 &\equiv \langle (\hat{C}_\ell^{\mathrm{TT}} - C_\ell^{\mathrm{TT}})^2 \rangle \\ &= \frac{2}{2\ell+1} C_\ell^{\mathrm{TT}} \,. \end{aligned} \tag{1.70}$$

This is known as *cosmic variance* and it is an intrinsic source of uncertainties. It is larger at the largest scales because at low-ℓ we have only a limited number of samples to average over with respect to the small scales.

From cosmological perturbations to CMB anisotropies The next step is to calculate the theoretical prediction for the CMB angular power spectrum and connect it with the evolution of primordial perturbations.
An approach is to evaluate the source terms of perturbations along the photon line of sight[30]. The C_ℓ^{TT} is the solution to the following integral:

$$C_\ell^{\mathrm{TT}} = \frac{2}{\pi}\int dk k^2 \mathcal{P}_R(k) |\Theta_\ell(k, \eta = \eta_0)|^2 \tag{1.71}$$

where $\mathcal{P}_R$ is the primordial power spectrum (1.40) and $\Theta_\ell(k, \eta_0)$ is the transfer function for the photons.
The perturbations that we observe in the CMB have been generated at the time of recombination, $\eta = \eta_*$, and after photon-matter decoupling have been propagating freely until today, $\eta = \eta_0$. Thus $\Theta_\ell(k, \eta_0)$ is the projection in spherical waves of the source function $S(k, \eta)$, which is eq. (1.55) calculated very close to recombination:

$$\Theta_\ell(k, \eta_0) = \int_0^{\eta_0} d\tilde{\eta} S(k, \tilde{\eta}) j_\ell[k(\eta_0 - \tilde{\eta})] \,, \tag{1.72}$$

[28] This is the *ergodic hypothesis*: since the CMB is homogeneous and isotropic, different points widely separated in the sky can be considered as different realizations of the same stochastic process [22].

[29] When this property is valid, the estimator is *unbiased*.

[30] This is known as Line of Sight approach [23] .

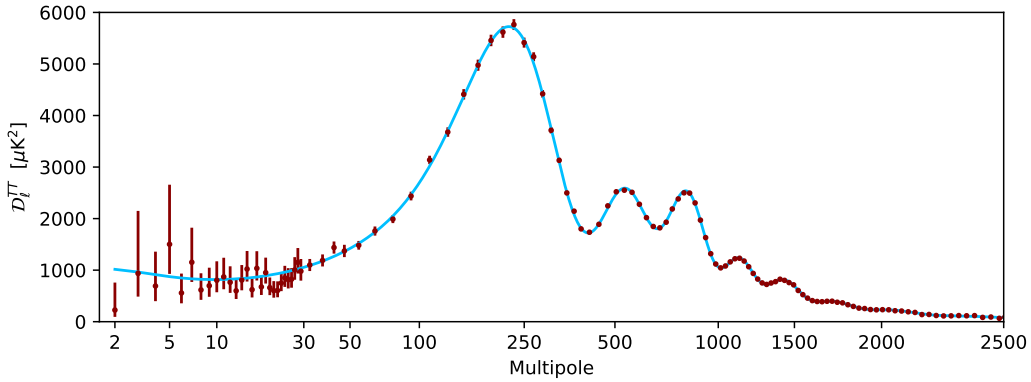


Figure 1.5: The temperature angular power spectrum of the CMB ($C_\ell^{\rm TT}$), normalized by a factor $\ell(\ell+1)/2\pi$, measured by Planck 2018.

where $j_\ell(k\eta)$ are the Bessel functions. The source function is written as:

$$\begin{aligned} S(k,\eta) \equiv & g(\eta)\left[\Theta_0(k,\eta)+\psi(k,\eta)\right] \\ & +\frac{i}{k}\frac{d}{d\eta}\left[u_b(k,\eta)g(\eta)\right] \\ & +e^{-\tau}\left[\dot{\psi}(k,\eta)+\dot{\phi}(k,\eta)\right], \end{aligned} \tag{1.73}$$

where we have defined the *visibility function* as $g(\eta) \equiv -\dot{\tau}(\eta)e^{-\tau(\eta)}$. The visibility function is the probability that a photon last scattered at η and it peaks sharply at recombination. This allows us to evaluate eq. (1.72) at η_* and then obtaining[31]:

$$\begin{aligned} \Theta_\ell(k,\eta_0) \simeq & \left[\Theta_0(k,\eta_*)+\phi(k,\eta_*)\right] j_\ell[k(\eta_0-\eta_*)] \\ & +3\Theta_1(k,\eta_*)\left[j_{\ell-1}[k(\eta_0-\eta_*)]-\ell(\ell+1)\frac{j_\ell[k(\eta_0-\eta_*)]}{k(\eta_0-\eta_*)}\right] \\ & +\int_0^{\eta_0} d\tilde{\eta}\, e^{-\tau(\tilde{\eta})}\left[\dot{\psi}(k,\tilde{\eta})+\dot{\phi}(k,\tilde{\eta})\right] j_\ell[k(\eta_0-\eta)]\,. \end{aligned} \tag{1.74}$$

Finally, we can insert the solutions for Θ_0 and Θ_1 that we have found in section 1.3.4 and calculate the $C_\ell^{\rm TT}$. The monopole and the dipole in eq. (1.74) are dominant at the time of recombination. The last term, called the *Integrated Sachs-Wolfe* (iSW) term is not negligible when the gravitational potential evolves, as during matter-radiation equality and at dark energy domination. Diffusion damping must be also included: the correct way to do it is to multiply $S(k,\eta)$ by the damping factor $\mathcal{D}$ (eq. (1.60)).
Fig. 1.5 is the $\mathcal{D}_\ell = \frac{\ell(\ell+1)}{2\pi}C_\ell^{\rm TT}$, normalized by a factor $\ell(\ell+1)/2\pi$, measured by Planck [5]. There are three main features:

- At low ℓ ($\ell \lesssim 100$) the $\mathcal{D}_\ell$ shows the so-called *Sachs-Wolfe plateau*: as expected, very large scales are frozen at the initial conditions (see eq.

[31] In the tight coupling limit $u_b \simeq -3i\Theta_1$.

(1.56)) ; thus $\frac{\ell(\ell+1)}{2\pi}C_\ell^{\rm TT} \propto A_s = \text{const}$, where A_s is the scalar amplitude of $\mathcal{P}_R$ (eq. (1.40)).

- The acoustic oscillations in eq. (1.59) , driven by the combined effects of gravity and baryonic pressure, produced the *acoustic peaks*. These peaks are located approximately at the extrema of $\cos kr_s$, i.e. $k_{pk} = n\pi/r_s$ ($n = 1, 2, ...$). Thus, the first peak corresponds to the sound horizon at the time of recombination, at $\ell \sim 100$. Changing the baryon density Ω_b changes the sound horizon and shifts the peak locations as well as the heights of the peaks.
- Small scales ($\ell \gtrsim 1000$) are damped by the *Silk damping*, due to the fact that the finite thickness of the last scattering is comparable to the mean free path of photons just before recombination. In this case, the damping factor is not negligible and washes out all the smaller scales. For higher values of $\Omega_b h^2$ the photon mean free path becomes shorter and the damping moves to smaller scales, so the angular power spectrum on scales $\ell = 1000$ is larger.

CMB secondary anisotropies Along their path, CMB photons interact with the cosmic structure and the intergalactic medium. These effects produce temperature anisotropies that are called *secondary* since are generated after the recombination. A comprehensive discussion about the CMB secondary anisotropies is detailed in Ref. [24].
The main contributions are:

- **The Sunyaev-Zel'dovich effect** is caused by the inverse Compton interaction between the CMB photons and the free electrons of a hot ionized gas along the line of sight.
- **Gravitational lensing of the CMB** by the large-scale structures along the line of sight does not produce anisotropies but distributes the photon's energy. This induced a shift in the CMB angular power spectrum along smaller scales.
- **Polarization** is primarily produced at the last scattering surface. The secondary polarization is generated during *reionization*. Between $15 \lesssim z \lesssim 6$ (it is not fully understood), the gas present in the Universe is ionized again. Thus CMB photons that are propagating interact with the free electrons of the ionized medium. The Compton scattering between photons and free electrons at reionization produces a quadrupole moment in the photon distribution different from zero, giving rise to second-order polarizations. This signal takes place at small scales and has a small amplitude;
- **The integrated Sachs-Wolfe effect** is sourced by the interaction of CMB photons with time-varying gravitational potential. Photons passing through a potential well are blue-shifted and red-shifted as

they fall into and then climb out. If the potential well is evolving meanwhile, the net energy of photons is different from zero, giving rise to anisotropies. This can happen after z_{eq}, when the Universe is matter-dominated but the contribution of radiation is still strong, and at the onset of Dark Energy domination. Thus we can distinguish the *early iSW*, which happens around recombination, and the *late iSW*, which contributes at late time. The early iSW effect affects scales that enter the horizon during the radiation-matter transition. In the CMB spectrum, this is reflected in the enhancement of the height of the first peaks. The late-iSW affects very large scales, enhancing the fluctuations at low-ℓ. Since the effect is small and the cosmic variance is dominating at those scales, the basic idea is to measure the late-iSW cross-correlation CMB temperature map with external LSS tracers, such as galaxies survey. The detection of this effect would hint at the presence of dark energy. In the next chapter, we will discuss it extensively.

1.4.2 Galaxy Clustering

Galaxy clustering is related to the three-dimensional distribution of galaxies, determined by observing the angular positions of galaxies in the sky and analyzing the redshifts of these galaxies. Studying the distribution of galaxies allows the direct measurements of the cosmic expansion history, by studying baryon acoustic oscillations (BAO). Moreover, it allows us to explore the growth of cosmic large-scale structures.
As discussed in the last section, the photons and baryons are coupled together, behaving like a fluid. The oscillations are imprinted on the CMB angular power spectrum at the typical scale of the sound horizon at the time of recombination. These oscillations leave a signature in the distribution of matter, the BAO in the matter power spectrum, approximately at the same scale r_s. The BAO can be used as a standard ruler.
Measuring the matter power spectrum from the observed galaxy distribution is not an easy task. First, it requires modeling a relation between the distribution of galaxies δ_g and the underlying distribution of matter perturbation δ_m. Usually, it is assumed the *linear bias relation*

$$\delta_g(\mathbf{x}) = b\delta_m(\mathbf{x})\,, \tag{1.75}$$

where $b(\eta)$ is a parameter that depends on the redshift and sensitively on the galaxy sample considered and $\mathbf{x}$ is the 3D position of the galaxy $\mathbf{x} = (z, \theta, \phi)$. This relation holds at large scales and at liner order in perturbations. Another source of uncertainties is the distortions in the redshift distribution of galaxies due to their peculiar velocities [25]. At large scales, we see galaxies moving towards an overdensity region due to gravity: the apparent effect is that galaxies closest to us appear to move

away, whereas we see the galaxies on the other side of the overdense region become closer. The net effect is that the density field becomes squashed along the line of sight and thus in redshift space the galaxy clustering is stronger. On the other hand, on smaller scales the effect is opposite and the clustering amplitude is smaller. These distortions are the so-called *redshift-space distortions*.
Since the number of galaxies must be preserved between real and redshift space, the correction in the density field is

$$\delta_{g,\mathrm{RSD}}(\mathbf{x}) = b\delta_m(\mathbf{x}) - \frac{\partial}{\partial \mathbf{x}}\left[\frac{\mathbf{u}_m(\mathbf{x})\cdot\hat{\mathbf{z}}}{aH}\right], \tag{1.76}$$

where $\mathbf{u}_m$ is the peculiar velocity of the field (see eq (1.26)). For large scales entered during the matter domination, it is related to the matter density perturbations through the Euler equation (first of eq. (1.48)) in Fourier space. Taking into account that the time dependence of large-scale perturbations δ_m is encoded into the growth factor D_+ and putting equal to zero the term on the left side $-3\dot{\phi}$ since the gravitational potential is constant, we can write the Euler equation as, in Fourier space:

$$\mathbf{u}_m(\mathbf{k},\eta) = \frac{i\mathbf{k}}{k^2}\frac{\dot{D_+}}{D_+}\delta_m(\mathbf{k},\eta) = aHf\frac{i\mathbf{k}}{k^2}\delta_m(\mathbf{k},\eta)\,. \tag{1.77}$$

The term f is the linear growth rate defined as:

$$f \equiv \frac{d\ln D_+}{d\ln a}, \tag{1.78}$$

that it is approximately predicted from the ΛCDM model to be $f \simeq [\Omega_m(s)]^{0.55}$, for $\Omega_m \sim 0.3$.
Transforming eq. (1.76) in Fourier space and inserting eq. (1.77), one obtains:

$$\delta_{g,\mathrm{RSD}}(\mathbf{k}) = \left[b + f\mu_k^2\right]\delta(\mathbf{k})\,. \tag{1.79}$$

Here μ_k is defined as the angle between the line of sight and the wavevector, $\mu_k = \hat{z}\cdot\hat{k}$, where $\hat{z}$ is the direction of the line of sight (in distance-observer approximation RSD affects only the $\hat{z}$-direction, not the transverse one). The first comment we can make about eq. (1.79) is that $f\mu_k^2 > 0$ and thus the apparent overdensity is larger in redshift space than in real space, whose contribution is $b\delta_m$, as anticipated. The second comment is that this apparent overdensity is only in the direction parallel to the line of sight. It follows that the observed matter power density is dependent on the direction:

$$P_{g,\mathrm{RSD}}(k,\mu_k,z) = P_L(k,z)\left[b + f\mu_k^2\right]^2, \tag{1.80}$$

where here P_L is the linear matter power spectrum defined in eq. (1.51). The last measurements of $P_{g,\mathrm{RSD}}$ are shown in Fig. 1.6, from different

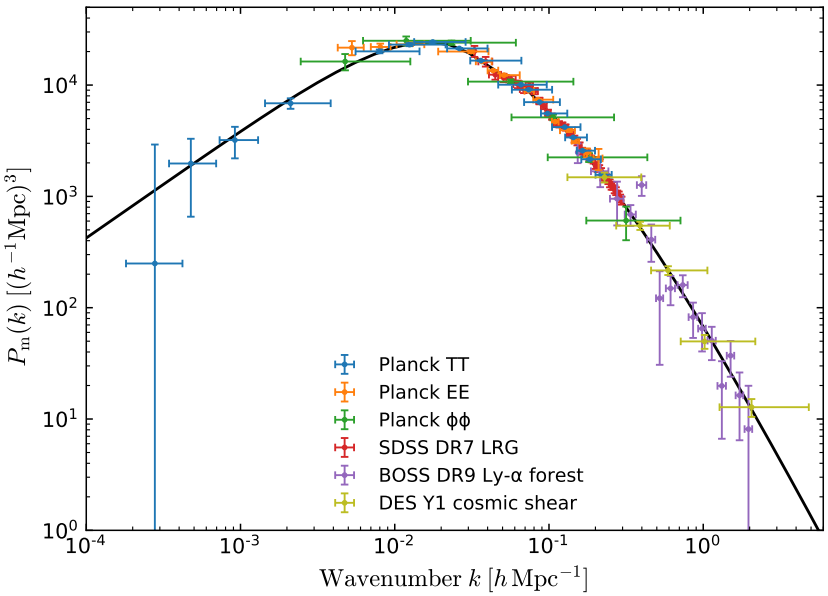


Figure 1.6: The (linear) matter power spectrum at $z = 0$ obtained from different cosmological probes: from CMB temperature, polarization and lensing measured by Planck 2018, from the BAO measured by the Sloan Digital Sky Surveys (SDSS) and Baryon Oscillation Spectroscopic Survey (BOSS) [26], from cosmic shear measured by the Dark Energy Survey (DES) [27]. Image courtesy of ESA/Planck.

cosmological probes. To disentangle b and f from the GC measurements what is usually done is to decompose the $P_{g,\rm RSD}$ in multipoles.As described in [28] and [29], the growth rate is a probe to modified gravity models, since an increase in gravity leads to an increase in the growth of structures.

1.5 The concordance model

The inflationary Λ Cold Dark Matter (ΛCDM) model is, arguably, the model that best fits both the observations of the CMB and of the distribution of galaxies, as well as other cosmological probes that we have not discussed. For now, it is thus considered by the scientific community as the concordance model of Cosmology.
This model sketches a Universe dominated by non-interacting cold Dark Matter and the cosmological constant Λ. The Universe is expanding according to the Hubble rate $H(z)$, which depends on the energy content at different times. At very early times, the Universe passed through an *inflationary* stage which provoked an exponential expansion. This inflationary mechanism is predicted to be the cause of the small density perturbations, at the level of 10^{-5}, which are observed in the CMB temperature field and in the matter distribution.

Base parameters		
$\Omega_b h^2$	$\Omega_c h^2$	$\ln(10^{10} A_s)$
0.02236 ± 0.00015	0.1202 ± 0.0014	3.045 ± 0.016
n_s	$\tau_{\rm reio}$	$100\theta^*$
0.9649 ± 0.0044	$0.0544^{+0.0070}_{-0.0081}$	1.04090 ± 0.00031
Derived parameters		
H_0 [km s^{-1} Mpc^{-1}	Ω_Λ	Ω_m
67.27 ± 0.60	0.6834 ± 0.0084	0.3166 ± 0.0084

Table 1.1: The best-fit and 1σ value for the ΛCDM cosmological parameter measured by Planck 2018. The base parameters are the 6 parameters of the concordance mode. The other parameters were derived from the base parameters. The data includes CMB temperature and polarization. The parameter τ_{reio} is the optical depth at reionization.

The Planck CMB observation, in combination with other probes (i.e. BAO, Type Ia supernovae), put the tightest constraints on the cosmological parameters that describe the concordance model. In table 1.1 are listed the ΛCDM cosmological parameters measured by Planck [5]. With these values of the density of the different components, the transitions between the matter-radiation era and matter-dark energy era happen at $a_{eq} = \frac{\Omega_{m,0}}{\Omega_\Lambda} = 2.9 \times 10^{-4}$ ($z_{eq} \sim 3408$) and $a_\Lambda = \frac{4.15 \times 10^{-5}}{\Omega_m h^2} = 0.46$ ($z_\Lambda = 1.18$), respectively, and the curvature density Ω_K is compatible with zero.
Planck measurements show that the primordial power spectrum $\mathcal{P}_R$ is nearly scale-independent, with $n_s = 1$ excluded at 8σ.

Chapter 2

The Euclid Mission

Euclid is a space mission conducted by the European Space Agency (ESA) with the primary objective of investigating Dark Energy and Dark Matter by mapping the large-scale universe across cosmic time, as detailed in the Euclid Red Book [cite]. Launched in July 2023, Euclid is positioned at the L2 Sun-Earth Lagrangian point, where it will observe the extragalactic sky for approximately six years.
The mission leverages two primary cosmological probes: Weak (Gravitational) Lensing (WL) and Galaxy Clustering (GC), covering a substantial area of 15,000 square degrees, equivalent to about 36% of the celestial sphere, up to a redshift of approximately z=2.
Euclid aims to utilize WL to constrain cosmological parameters. To achieve this, the mission will collect data on the shapes and distances of approximately 30 million galaxies. GC requires measurements of the 3-D distribution of galaxies based on their spectroscopic redshifts. To fulfill these tasks, Euclid is equipped with two onboard instruments: a visible imager (VIS) and a near-infrared photometer and spectrometer (NISP).
The VIS instrument employs a 6x6 array of 4k CCD detectors sensitive to the visible range, allowing it to measure the shape distortions of 1.5 billion galaxies. On the other hand, NISP is equipped with a matrix of 4x4 detectors sensitive to the near-infrared range (900-2000 nm) and operates in two modes: photometric and spectroscopic. In the photometric mode, NISP uses three broadband filters to estimate the photometric redshifts of galaxies imaged by VIS, enhancing the information obtaines from the WL survey through tomography. In spectroscopic mode, NISP uses grisms to observe the spectra of about 50 million galaxies. A grism is an optical element that combines a prism and a grating, enabling light of a specific central wavelength to pass directly through. By combining these two primary methods, Euclid aims to achieve unprecedented precision in measuring fundamental cosmological parameters.
This chapter provides a description of the Euclid mission, structured as follows: Section 2.1 outlines the scientific goals of the mission; Section 2.2

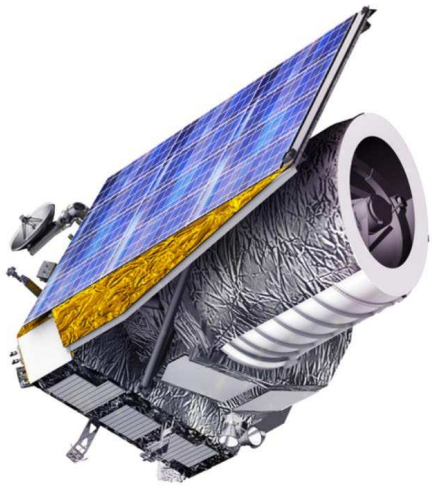


Figure 2.1: Overview of the *Euclid* satellite. Image credit: ESA/Euclid.

provides details about the cosmological probes; Section 2.3 offers insights into the satellite's structure, including the VIS and NISP instruments. Finally, Section 2.4 details the Euclid survey, outlining the observational strategy for scanning the sky and the scientific requirements for the primary probes. Most of the information contained in this chapter can be found in the Euclid Red Book [30].

2.1 Scientific objectives

Euclid aims to provide precise and accurate measurements of the expansion and evolution of the Universe. Here, we summarize the primary scientific objectives of the mission:

Dark Energy The *Euclid* project aims to precisely determine the values of w_0 and w_a as defined in equation (1.21). The targeted precision for these parameters is approximately 0.01 for w_0 and 0.1 for w_a. Achieving this level of precision will offer unparalleled constraints, allowing for the discerning between a cosmological constant and more complex dynamical dark energy models.

Modified Gravity Another possible explanation for cosmic acceleration arises from Modified Gravity (MG) theories, which introduce modifications to General Relativity (GR) on cosmological scales. In certain MG models, these deviations from GR are imprinted in the growth of cosmic structure, altering the parameter γ in the growth rate of dark matter fluctuations, denoted as $f(a)$:

$$f(a) \equiv \frac{d\ln D(a)}{d\ln a} \simeq [\Omega_{m,0}(a)]^{\gamma} \,, \tag{2.1}$$

where the ΛCDM model predicts $\gamma \simeq 0.55$. The primary objective of the *Euclid* project is to measure this parameter with a precision of 1σ uncertainty of 0.02, providing a rigorous test of the validity of ΛCDM. If any deviation from the expected value is detected, it would signify the presence of new physics beyond the currently accepted concordance model.

Massive neutrinos and dark matter Oscillation experiments have confirmed the existence of mass for certain neutrino species, although they contribute only a small fraction to the total dark matter content in the Universe, as indicated by observations of LSS and the CMB. Determining the absolute masses of neutrinos poses a challenging task, but it is essential for understaing their mass hierarchy (whether normal or inverted) and the mechanism responsible for their mass generation. In cosmology, neutrinos tend to free-stream out of dense regions since they are influenced only by gravity and are relatively light. This behavior leads to a suppression of the matter power spectrum at small scales. The extent of this suppression is directly linked to the sum of neutrino masses, denoted as Σm_ν. By achieving a precision of $\Delta m_\nu < 0.03$ eV, *Euclid* can measure this effect on the matter power spectrum and discern the neutrino mass hierarchy, provided that the total mass falls within the range $\Sigma m_\nu < 0.1$ eV. Moreover, *Euclid* is set to conduct highly precise measurements of dark matter halo density profiles and the matter power spectrum. These measurements will significantly advance our comprehension of dark matter by substantially enhancing constraints on its self-interaction cross-section, improving them by three orders of magnitude compared to current limits. These measurements will significantly enhance our understanding of dark matter by improving constraints on its self-interaction cross-section by three orders of magnitude compared to current limits. This will help determine if dark matter is warm and also establish a lower limit for its mass.

Inflation and Non-Gaussianity The concordance model predicts that the primordial power spectrum of density fluctuations, originating from inflation, is both Gaussian and nearly scale-invariant. *Euclid* is set to measure the scalar spectral index n_s, as defined in (1.40), with a level of precision comparable to that achieved by Planck. Additionally, the mission aims to determine the parameter f_{nl}, which quantifies the amplitude of non-Gaussianity. In other words, it measures how much the primordial fields deviate from a perfectly Gaussian distribution. *Euclid* aims to achieve this with a precision of approximately 2%, representing an improvement over the accuracy attained by Planck.

2.2 Cosmological probes

Euclid's ambitious scientific objectives will be accomplished through an unprecedented galaxy survey, characterized by its remarkable scale. This survey will employ the VIS instrument to capture images of approximately 30 galaxies per square arcminute, a total of 1.5 billion galaxies. This large number will enable the mapping of the distribution of dark matter through weak gravitational lensing. Additionally, near-infrared photometry will provide precise estimates of the galaxies' photometric redshifts, with a precision of $\sigma_z/(1+z) < 0.05$. Knowledge of these photometric redshifts allows for the partial recovery of the three-dimensional mass distribution through a technique known as Weak-Lensing tomography. The precision of cosmological parameter measurements from a WL survey can be significantly enhanced by tomographically determining the evolution of the statistical properties of large-scale structure across the finite redshift width of the source distribution [31]. In the case of *Euclid*, the measurements of photometric redshifts are complemented by ground-based imaging data, such as data from DES ([32]) and LSST ([33]), to achieve an accuracy of $\sigma_{\langle z \rangle}/(1+z) = 0.02$ in the measurements of the mean redshift of each tomographic bin [34].
Concurrently, *Euclid* will conduct a spectroscopic redshift survey utilizing the NISP instrument. This survey's goal is to acquire spectroscopic data for 50 million galaxies within the redshift range of $0.9 < z < 1.8$. The required accuracy for spectroscopic redshift measurements is $\sigma_z^{sp}/(1+z) < 0.01$. Spectroscopic redshifts will be determined by detecting the H-alpha (Hα) emission line, which has a wavelength of approximately 656.3 nm. To cover the desired redshift range of 0.9 to 1.8, a spectroscopic instrument sensitive to the near-infrared (NIR) range (1100-2000 nm) is essential. This choice is justified by the fact that near-infrared light within this range is fully absorbed by the Earth's atmosphere, making it impossible to explore this part of the electromagnetic spectrum with ground-based experiments.
The cross-correlation analysis between the two primary probes can enhance the information available for addressing the mission's key cosmological questions. The main advantage arises from conducting both a comprehensive deep imaging survey and an extensive redshift survey across the exact same region of the sky, increasing statistical power and control of the systematics [30].
In addition to the primary probes, WL measurements will be complemented by additional cosmological probes obtained from external data sources, such as CMB temperature anisotropies and CMB lensing provided by Planck [35]. The cross-correlation analysis of these complementary datasets will be discussed in Chapter 4.

Figure 2.2: The*Euclid*'s payload, containing the telescope and the scientific instruments. Credits: ESA/Euclid.

2.3 The spacecraft

The spacecraft can be divided into three primary components: a Service Module (SVM), a Payload Module (PLM) housing the telescope, and the Scientific Instruments [36]. This configuration fits within a diameter of 3.74 meters and a height of 4.8 meters.

The SVM includes the spacecraft subsystems that support payload operations. It hosts the warm electronics of the payload and establishes structural connections with both the PLM and the launch vehicle. Within the SVM is the Sunshield, which shields the PLM from direct sunlight and serves as a foundation for the photovoltaic assembly, generating electrical power for the spacecraft.

The PLM (see Fig. 2.2) contains the main instruments, the optical components of the telescope, and the mechanical and thermal interfaces to the instruments (radiating areas and heating lines). The material that constitutes the telescope ensures excellent thermo-elasticity and stiffness. Throughout the survey, the main mirror will be actively cooled and maintained at a low temperature. The two instruments onboard are the VIS and the NISP, both of which capture images of the same portion of the sky. This arrangement enables the acquisition of multiple sets of data with a single telescope alignment.

2.3.1 The Visible Imager

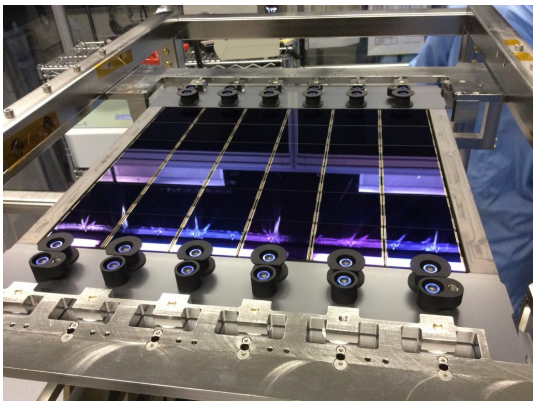


Figure 2.3: The VIS instrument's focal plane. Image credits: Euclid Consortium.

The VIS instrument is primarily designed for precise WL measurements. It achieves this by accurately measuring the shapes of galaxies within the visible wavelength range of 550-900 nm [37].
Beyond WL shape measurements, VIS data also play a crucial role in improving photometric redshift (photo-z) estimation. This is accomplished by enhancing the photometric information from ground-based observations, particularly for less resolved sources, thanks to the exceptional image quality achieved through diffraction-limited imaging in the VIS band [36].
Following data acquisition, the VIS central processing unit will construct images from pixel data onboard the spacecraft. Subsequently, the data will be compressed and transmitted to the ground for further analysis [36].
The focal plane of the VIS instrument features a 6x6 CCD matrix, consisting of approximately 4000x4000 pixels. Each pixel measures 12 μm on each side, corresponding to an angular scale of 0.1 arcsec. This scale translates to a total Field of View[1] (FoV) of 0.5 deg^2. With an expected image quality of around 0.23 arcsec, the instrument enables highly precise measurements of galaxy shapes with high accuracy.
Additional components of VIS include the calibration unit, the shutter, and two electronic units. One electronic unit processes data, while the other controls the instrument's mechanisms and manages the allocated power.

[1]The FoV is the solid angle hrough which a detector is sensitive to electromagnetic radiation.

2.3.2 The Near Infrared Spectrometer and Photometer

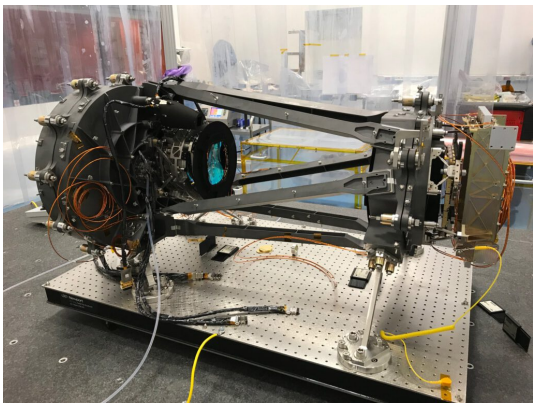


Figure 2.4: The NISP instrument. Image credits: Euclid Consortium.

NISP is designed to observe the sky in the near-infrared spectrum, capturing both photometric and spectroscopic data from extragalactic sources within the same FoV as VIS. The NISP focal plane consists of a grid of 4x4 detectors, each featuring 2048x2048 pixels with a size of 18 μm, providing an angular resolution of 0.3 arcsec.
Near-infrared photometry will be conducted on the same sources imaged by VIS. The photometric measurements employ three filters, namely Y_E, J_E, and H_E, covering the wavelength range of 950-2020 nm. These filters are placed in the Filter Wheel (FWA), alongside two additional slots: OPEN, allowing unfiltered light for spectroscopy to get through the Grism Wheel (GWA), and CLOSED, used for capturing dark frames to estimate detector thermal response.
NISP will acquire spectra of all sources in the FoV using a technique called *slitless spectroscopy*. The incoming light from the telescope will be dispersed using one of the grisms mounted on the GWA, namely RGS000, RGS180, and RGS270, three "red" grisms, covering the wavelength range between 1200 and 1850 nm. Each grism provides spectra along a different direction, which will be combined offline to refine the spectrum of each source by removing contaminants. The grism RGS270 will not be used, due to a non-conformity found in 2020. Observations with RGS000 and RGS180 will be supplemented by two additional observations, tilting RGS000 by -4° and RGS180 by +4°. A fourth "blue" grism (BGS000) will cover the range of 920-1400 nm and will only be employed during deep surveys [35]. In addition to the focal plane, NISP is composed of the warm electronics unit, located in the Service Module, which synchronizes the acquisition of the detectors and processes data in-flight. There is also a calibration unit

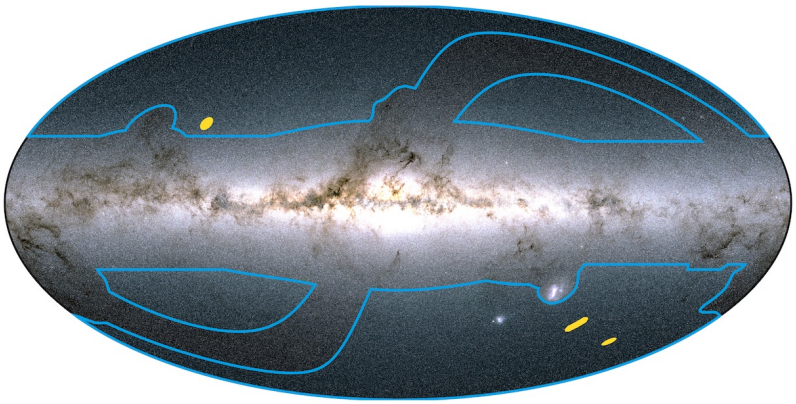


Figure 2.5: Two-dimensional map of the Euclid Wide Survey, in blue. The additional Deep Surveys are marked in yellow. Credit: ESA/Euclid.

2.4 The Euclid Survey

As pointed out in the previous sections, in order to accomplish its primary scientific objective, *Euclid* needs to observe a substantial portion of the extragalactic sky using both multiband imaging and slitless spectroscopy. This extensive survey is referred to as the *Euclid* Wide Survey (EWS) [36], which will encompass approximately 15,000 deg^2 for both WL and GC measurements. The survey is planned to be completed within the nominal six-year mission lifespan. Fig. 2.5 shows a visual representation of the survey area. In conjunction with the EWS, *Euclid* will conduct an additional deep survey (DS), spanning 40 deg^2 over the northern and southern Galactic hemispheres. The DS will supply scientific calibration datasets to the EWS and contribute data for legacy science purposes. This includes the observation of faint high-redshift galaxies, quasars, and active galactic nuclei (AGN) [35].

2.4.1 The survey strategy

On a daily basis, *Euclid* observes a circular region in ecliptic coordinates, with the size of the circle varying depending on the latitude. The axis of the Sun-spacecraft undergoes a movement of approximately 1 degree per day, ensuring that the satellite remains as perpendicular to the Sun as possible. This alignment ensures that the spacecraft's Sunshield consistently faces the Sun, thereby maintaining the thermal stability of the mirrors and instruments.
Each survey tile is observed using a "step-and-stare" approach, where the telescope is initially pointed at a fixed position in the sky before moving to the next position to capture both imaging and spectroscopic measurements. *Euclid* employs a dithering technique, which involves a series of four observations with slight changes in the telescope's pointing direction between them. During each frame, both the VIS and NISP instruments carry out sky exposures simultaneously. First VIS and NISP acquire the data, and then the NISP photometric imaging follows. During this phase, VIS closes its shutter in order to avoid disturbance with the measurements. Subsequently, a dither step is executed. This entire procedure is repeated for each of the observations. The purpose of this strategy is to partially correct instrumental effects in the images, such as those caused by the gaps between detectors in the focal planes. Additionally, this strategy mitigates the impact of cosmic rays on the data [36] .

2.4.2 Photometric survey

As previously mentioned, to meet the scientific objectives of the mission, the photometric survey must encompass a galaxy count of at least 30 per square arcminute.
The information gleaned from WL analysis is further enhanced when considering the redshift distribution of the sources. WL tomography relies on robust estimates of galaxy redshifts. The process of deriving photometric redshifts relies on identifying the optimal match between the broad-band colors and the spectral template of galaxies. The precision of photometric redshift measurements depends on the number of filters employed and the signal-to-noise ratio of the observations. Consequently, there can be a dispersion between the photometric redshifts and the true (spectroscopic) redshifts. The requirement for the dispersion is $\sigma_z^{\rm ph}/(1+z) \leq 0.05$.
AThis level of precision can only be achieved by the combination of the three broad-band NIR bands with additional ground-based data. The survey area must be imaged from the ground with a minimum of four additional filters, spanning the full wavelength range of 420 - 930 nm.
In order to achieve the required accuracy on the estimate of the mean redshift of each tomographic bin, it is crucial that the set of templates employed for calculating photometric redshifts accurately represent true

galaxy spectra. These templates can be acquired through ground-based spectroscopic observations, numbering approximately 10^5 galaxies, with insignificant fraction (less than 10^{-4}) exhibiting incorrect redshift estimates.

2.4.3 Spectroscopic survey

The galaxy-clustering signal becomes stronger with a larger number of observed galaxies. *Euclid* is required to obtain correct redshifts for a minimum of 3,500 galaxies per square degree, resulting in a total sample of 52.5×10^6 galaxies over the entire sky area (15,000 deg^2). This will be achieved through slitless spectroscopy, specifically targeting Hα emitting galaxies. As previously mentioned, the required accuracy of the spectroscopic redshift measurement is $\sigma_z^{\rm sp}/(1+z) \leq 0.001$.
The target of the survey is the Hα emitting galaxies, which emit spectral radiation due to the transition between the n=3 and n=2 energy levels of hydrogen, with wavelength $\lambda \simeq 656.3$ nm. Due to the large number of detectable spectra within the FoV and the broad wavelength coverage, there is a significant possibility of spectral contamination or confusion arising from neighboring galaxies. This contamination can lead to failures in the redshift measurement, meaning that not all identified spectra can provide a reliable measurement of the redshift. Completeness and purity are the parameters that quantify the quality of the resulting galaxy catalog. The completeness C is defined as the ratio between the number of galaxies with measured redshifts $N_{\rm means}$ and the total number that can be detected at this flux limit $N_{\rm tot}$, considering the spatial density and luminosity function[2] of Hα emitting galaxies. For *Euclid*, C should exceed 45%. On the other hand, purity is the ratio between the number of galaxies with correctly measured redshift ($N_{\rm corr}$) among the total number of measured spectra ($N_{\rm means}$). To correctly detect the BAO signal, the purity has to be $> 80\%$. Achieving these levels of completeness and purity requires observing the same FoV multiple times with varying orientations. For this reason, NISP utilizes multiple grisms with different orientations.
To reduce ambiguities, photometric data provide reference images for extracting the spectra during the spectroscopic survey.

[2]The luminosity function of galaxies is defined as the distribution of the galaxies given their luminosity

Chapter 3

The Near Infrared Spectrometer and Photometer

This chapter provides a comprehensive overview of the NISP instruments. Section 3.1 describes the acquisition system responsible for collecting signals in the detectors. In Section 3.2 we discuss the advantages of *slitless spectroscopic* for precise spectral measurements.Finally, Section 3.3 describes into the NISP photometric system and summarizes how photometric redshift distribution of the galaxy catalog is derived from the data.

3.1 The NISP detection system

As detailed in Section 2.3.2, the NISP is equipped with highly sensitive semiconductor detectors paired with dedicated readout electronics designed specifically for infrared detection applications. These detectors feature arrays with dimensions of 2048 x 2048 pixels. Their primary function is to convert incoming near-infrared photons from astrophysical sources into electron-hole pairs. Each photosensitive pixel measures 2040x2040 in size and is accompanied by 4 rows and 4 columns of reference pixels, which play a crucial role in monitoring potential biases and temperature fluctuations that may arise during extended exposures, as discussed in Ref. [38].
The *quantum efficiency* (QE), which denotes the conversion rate, may exhibit variation across the pixel matrix. Additionally, due to the operational temperature, *dark current* - the charge generated within the detector in the absence of external radiation - affects the detection. The detectors selected for integration into the NISP for the *Euclid* mission are characterized by their superior quantum efficiency ($> 80\%$) and minimal dark current (< 0.01 e-/pix/sec) compared to alternative infrared detector materials, as referenced in Teledyne's documentation on detectors [39].
However, an additional source of signal contamination, known as *readout noise*, is caused by the read-out method utilized for signal digitization. The

QE, the dark current, and the readout noise are significant considerations for the performance of the NISP and must be accurately quantified through instrument measurements and subsequently incorporated into the data simulation models. Ensuring that these parameters meet specified values is a crucial objective. Furthermore, it is essential that at least 95% of pixels meet these requirements to guarantee efficient coverage throughout the entire survey. [40]
In the next section, we will discuss how to reduce the readout noise during frame acquisition on-flight.

3.1.1 Readout modes for the detectors

During an acquisition process, continuous measurements of the integrated signal are taken from the beginning of the exposure. One approach to determine the integrated signal value over an exposure duration of t_{exp} involves measurements immediately following the "reset" images. These reset images are acquired by the detectors when the observation is paused, as well as measurements taken at the end of the exposure. For a pixel's single value, the acquired data can be expressed as $\Delta S = S(t) - S(t_0)$, where $S(t)$ represents the signal measurements at time t in Analog-to-Digital Units (ADU) [1].
The uncertainty associated with ΔS is calculated as follows (where CDS denotes correlated double sampling):

$$\sigma_{CDS} = \sqrt{N_{photons} + 2 \times \sigma^2_{readout}}\,, \tag{3.1}$$

Here, $N_{photons}$ accounts for the Poisson noise related to the distribution of incident photons, while $\sigma_{readout}$ represents the readout noise for the pixel.
To mitigate the effects of readout noise, a commonly employed technique, which will also be implemented in NISP detectors, is the use of the multiple accumulated sampling (MACC) mode, as discussed in Ref. [38]. The MACC readout mode combines Flower-M and Up-the-Ramp (UTR) acquisition methods.
In the Fowler-M mode, the integrated signal is measured by averaging a set of n_f images, with the images grouped together before the averaging process. For a measurement consisting of n_g groups, assuming a null signal from the observation ($N_{photons}$=0), the accuracy in this acquisition mode can be expressed as follows:

$$\sigma_{Fowler-M} = \sqrt{2 \times \frac{\sigma^2_{readout}}{n_f}} = \frac{\sigma_{CDS}}{\sqrt{n_f}}\,. \tag{3.2}$$

Hence, as the number of frames per group increases, the readout noise associated with each group decreases, enhancing the accuracy of the measurement.

[1] The units in which the output signal is expressed

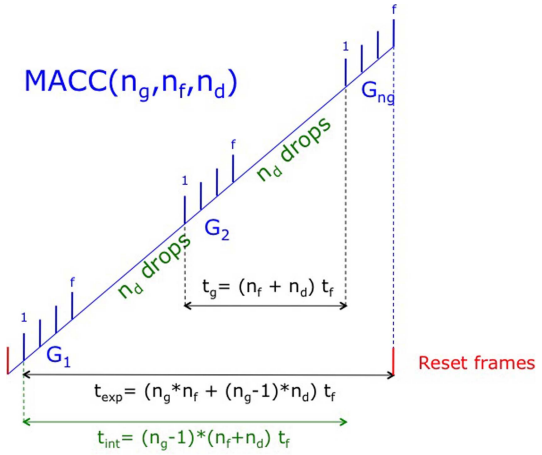


Figure 3.1: Multiple accumulated sampling MACC[n_g, n_f, n_d] with n_g equally spaced groups sampled Up the Ramp, n_f frames per group and n_d dropped frames between two successive groups [38].

When all the images are acquired throughout the exposure duration, the signal is sampled "Up the Ramp". This means that the signal accumulates over time and can be estimated, for instance, by linearly fitting the data points within the ramp. In this scenario, the noise level is higher because two subsequent measurements are correlated by the same incoming photon flux. Additionally, this approach necessitates the storage and processing of a substantial amount of data. However, it offers the advantage of detecting potential anomalies that might occur during a ramp, such as the effects of cosmic rays.
Combining these two methods, the acquisition mode involves reading n_f images within each of the n_g groups obtained during the exposure period, with intervals of n_d non-transferred images in between. To simplify the notation, this readout mode is denoted as MACC[n_g, n_f, n_d], as illustrated Fig. 3.1. Similar to the Fowler-M mode, the slope of the ramp is proportional to the photon flux.
The data initially transmitted to the ground undergo in-flight processing due to telemetry constraints, as only a limited amount of information can be transmitted. Subsequently, the flux is estimated based on the ramp's slope.

3.2 Spectroscopy with the NISP

The spectra of galaxies play a crucial role in astrophysics as they unveil distinct features at specific wavelengths, and thus redshift, known as absorption and emission lines. Grism spectroscopy offers a valuable method to simultaneously investigate the properties of galaxies. This technique employs specialized optical components called "grisms" to examine the electromagnetic spectrum of observed objects. By combining a diffraction

grating and a prism, the grism disperses the incident light and then redirects it back into the detector, enabling the observation and analysis of the dispersed light.
In *Euclid*'s mission, the instrument for spectroscopy measurements is NISP, described in 2.3.2, which will record two-dimensional images of the spectra. From these images will be extracted the calibrated one-dimensional spectra of the galaxies.

3.2.1 Slitless spectroscopy

Spectra obtained from astrophysical objects serve multiple purposes, offering intricate insights into the source's characteristics, including its chemical composition, pressure, density, and temperature. Galactic spectra additionally provide data concerning the rotation speed and radial velocity of galaxies. In the context of galaxy surveys, spectral data plays a pivotal role in determining the redshift of galactic sources. By identifying specific spectral features, such as an emission line with a known wavelength, it becomes possible to estimate the source's redshift accurately. In the *Euclid* survey, as previously mentioned, spectroscopic redshifts will be determined by identifying the Hα emission line.
Spectra of celestial sources are typically acquired using specialized instruments known as spectrometers. In specific surveys, the light emanating from celestial objects is captured through a narrow aperture referred to as a "slit," resulting in what is known as a slit spectrometer. The primary advantage of employing a slit spectrometer lies in its targeted approach to observation, enabling the selection of specific sources and minimizing contamination from other objects. However, in cases where recording spectra for multiple sources is required, using a slit may not be practical. In such circumstances, an alternative technique known as *slitless spectroscopy* is employed.
Slitless spectroscopy functions by defining "slits" based on the position, size, and shapes of sources in the sky. In this method, an object disperses its light in a slitless image, effectively creating its own virtual slit. Consequently, each source generates a linear track on the detector, with an emission line manifesting as a bright point along that track, as shown in Fig. 3.2. For non-circular objects like galaxies, it is possible to define a tilted virtual slit using a specific axis, like the major axis of the object. The positions of these virtual slits are determined by the spatial arrangement of objects in the sky, which can lead to instances of overlapping spectra. This cross-contamination primarily occurs in the dispersion direction since the spectra are significantly longer in that direction compared to the target. This source of noise can be resolved by capturing multiple exposures of the same field of view with different dispersion directions. This approach ensures that two sources with overlapping spectra in one dispersion direction will not overlap in the orthogonal direction. By combining the data obtained from images dispersed orthogonally, it becomes possible to mitigate confusion

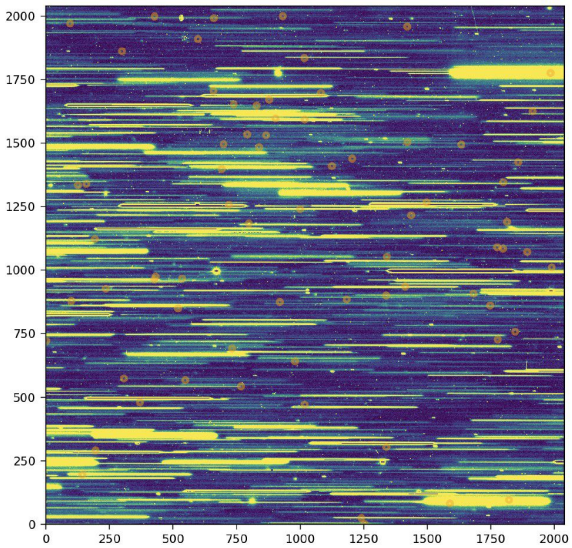


Figure 3.2: Simulation of the NISP exposure, each source produces a straight line in the detector, and each point is a Hα [41].

arising from cross-contamination.
Another characteristic of slitless spectroscopy is the higher sky background level compared to slit spectroscopy. In the context of the *Euclid* mission, the primary contribution to this background light is the so-called *zodiacal light* [41], arising from sunlight scattering off interplanetary dust. Zodiacal light is most pronounced along the ecliptic plane, coinciding with the Earth's orbit. To mitigate the impact of zodiacal light, the *Euclid* survey prioritizes regions at higher ecliptic latitudes, where the background light is less intense [36]. Accurate modeling of the zodiacal light and its subsequent removal is crucial, as it, along with cross-contamination, becomes one of the principal sources of noise.
Furthermore, slitless spectroscopy is subject to image distortions, which occur when the optical system's imaging scale varies across the field of view, particularly for off-axis points. These spectral distortions cause deviations in the spectra traces from straight lines aligned with the dispersion direction and may lead to stretching or elongation of the spectra.

3.2.2 Calibration of the spectra

The *Euclid* spectroscopic pipeline produces 1D spectra for each object by processing the raw images captured by the NISP. These spectra are then stored within the catalog provided by the photometric survey. To ensure their accuracy, each spectrum must undergo meticulous calibration in terms of wavelength and flux.

Given the central wavelength of the observed range, which is $\lambda = 15 \times 10^3$ Å, the required accuracy for wavelength calibration should adhere to the

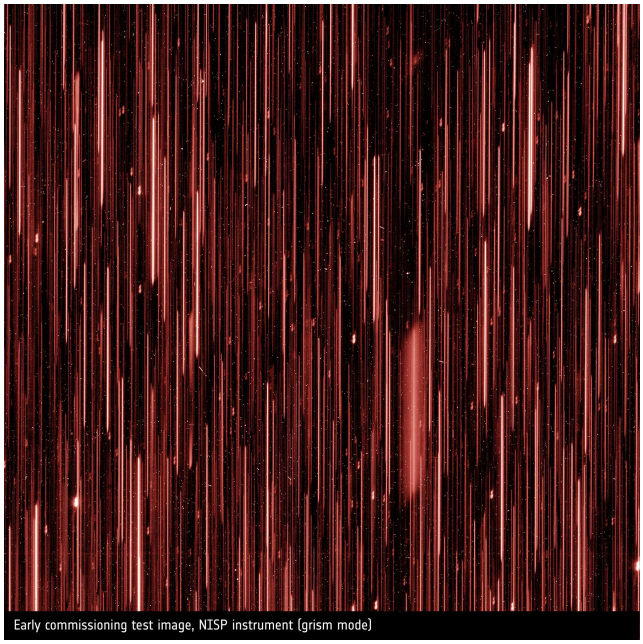


Figure 3.3: First image after *Euclid* launch of NISP in spectroscopic mode. This was an image to test if the instrument is on focus and works as expected. Credits: Euclid Consortium.

following criterion:

$$\sigma_{\Delta\lambda} < 15 \times 10^3 \,\mathring{\mathrm{A}} \times 0.001 = 15 \,\mathring{\mathrm{A}}\,. \tag{3.3}$$

Since a pixel of the NISP detector corresponds to 13.4 $\mathring{\mathrm{A}}$, eq. 3.3 is $\sigma_{\Delta\lambda} \simeq 1.1$ pixels.
Additionally, the flux must be accurately determined through calibration procedures.

3.3 Photometry with the NISP

As mentioned earlier, *Euclid* will capture images in the 0.95–2.02 μm range (NIR) using three distinct passbands (Y_E, J_E, and H_E) as illustrated in Fig 3.4. These passbands are rectangular in shape, share the same relative spectral width $\Delta\lambda/\lambda$, and have not inter-passband gaps. The precision, as evaluated by Ref. [30], is set at $\sigma_z^{phot} < 0.05(1+z)$.
In conjunction with ground-based photometry conducted in the optical passbands (also shown in Figure 3.4), the NISP's passbands enable the determination of average photometric redshifts (photo-z) for the tomographic redshift bins with the specified precision. The external data sources to be utilized include UNIONS (northern hemisphere) [42] and LSST (southern hemisphere)..

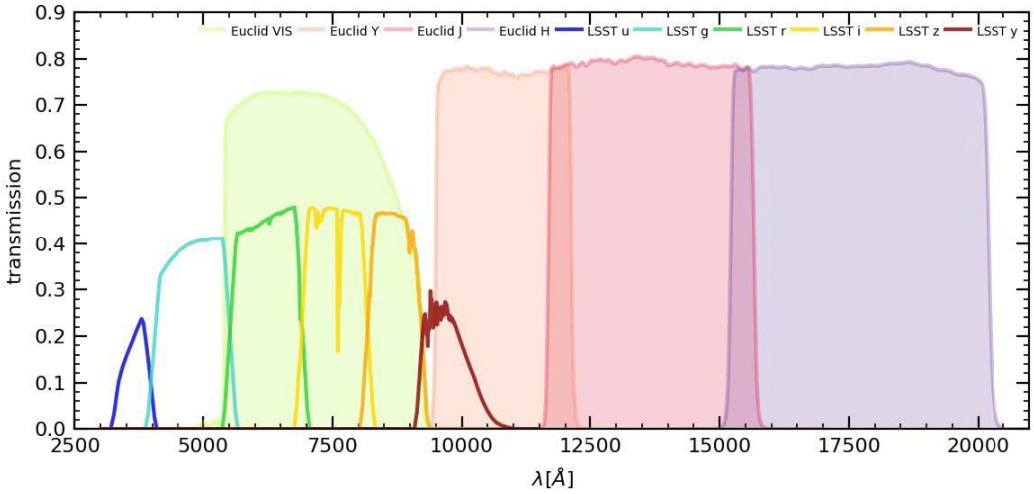


Figure 3.4: NISP photometric passbands and spectral response. There are shown also the VIS passband and the passabands of the ground-based experiment LSST. Images taken from OU:Photo-z Lectures at Euclid Summer School 2022.

Figure 3.5: First image of NISP in photometry mode taken after the *Euclid* launch. This was taken using the 'Y' band. This was an image to test if the instrument is on focus and works as expected. The image on the left is the NISP's full field of view, on the right there is the zoom in of the 4% of the full FoV. Credits: Euclid Consortium.

3.3.1 Photo-z distribution

Determining the distance to an extragalactic source is a crucial step in deriving meaningful physical quantities. The primary method for measuring the distance of an object is through its *Spectral Energy Distribution* (SED), which includes both a continuum and emission/absorption lines. As the Universe expands, the SED experiences stretching towards longer wavelengths by a factor of $1 + z$, where z represents the redshift. Estimating the distance presents challenges that involve identifying characteristic features in the SED and quantifying their amount of stretching. Subsequently, the measured redshift is associated with a proper distance, assuming a cosmological model.
Through the measurement of an object's flux, it becomes possible to obtain a sparse sampling of the SED. This sampling is sufficient for constraining the shape of the continuum, identifying the extragalactic nature of the sources, and estimating the redshift, focusing on prominent emission and absorption lines. This form of distance estimation, characterized by its lower resolution, is commonly referred to as a "photometric redshift" (photo-z).
The primary advantage of utilizing photo-z is the ability to derive distance measurements for all identified sources in an imaging survey. However, this advantage comes at the cost of lower redshift precision compared to spectroscopic measurements. For this reason, any study reliant on photo-z necessitates an accurate evaluation of photo-z performance. This assessment relies on various factors, including image properties (such as depth and wavelength coverage) and the nature of the sources (galaxies, active galaxy nuclei (AGN), or stars). Evaluating photo-z performance requires a deep and representative spectroscopic sample, highlighting the complementary nature of photometric and spectroscopic redshift surveys [43].
The first step in this process is to model the SED for different sources. Two commonly employed methods include defining SED templates, obtained through either theoretical models or observations, and employing Machine Learning algorithms capable of learning the mapping between colors and fluxes from training samples. Once the photometric redshift is determined, it is fundamental to retrieve the Probability Distribution Function of redshift (zPD) by comparing the model with the observed fluxes. This involves maximizing a likelihood function over a range of priors in the parameter space.
The template fitting method has the advantage of including rare galaxy models and is particularly favored when the spectroscopic redshift range is limited, and the field is deep. On the other hand, the ML approach is less favored with a limited training sample from spectroscopic measurements but proves useful when dealing with with a survey with wide FoV.
Euclid will handle a huge amount of data coming from the EWS, distributed within a restricted spectroscopic redshift range. It will use data from the EDS and from the external surveys for calibration. Consequently, the

template-fitting method will be employed to build a reference sample based on the data obtained from deep photometry in the calibration fields. This reference sample will then serve as the basis for the ML algorithm, from which the zPD for the EWS observations will be derived.

3.3.2 Redshift tomography

Photo-z has become a major tool for studying the properties of dark energy, particularly in the context of WL tomography, which has emerged as a key cosmological probe in current and future cosmological experiments. The idea of tomography involves slicing the redshift range into bins, to reconstruct the three-dimensional distribution of matter fields. In Ref. [31], it is noted that tomography can improve the measurements of the cosmological parameter estimates with respect to the non-tomographic case. Tomographic WL studies require accurate knowledge of the galaxy redshift distribution. Since these tomographic bins are commonly constructed from photo-z, they are affected by the uncertainties in the estimation of redshift from photometric data. Defining the mean of the redshift distribution as:

$$\langle z \rangle = \frac{\int_0^{\inf} z n(z) dz}{\int_0^{\inf} n(z) dz}, \tag{3.4}$$

where $n(z)$ is the true redshift distribution of the sample, $\langle z \rangle$ must be known with an accuracy (with respect to the spectroscopic redshift) better than $\sigma_{\langle z \rangle} = 0.002(1+z)$ to achieve the scientific objectives of the mission, as detailed in the previous section.
In the case of *Euclid*, according to the binning scheme proposed in Ref. [44], the survey is divided into 10 tomographic redshift bins, each containing an equal number of galaxies, with a median redshift of $z_m = 0.9$. The number density distribution per bin is defined as the convolution between the true redshift distribution and the probability distribution function $p_{ph}(z_p|z)$:

$$n_i(z) = \frac{\int_{z_i^-}^{z_i^+} dz_p n(z) p_{ph}(z_p|z)}{\int_{z_{min}}^{z_{max}} dz \int_{z_i^-}^{z_i^+} n(z) p_{ph}(z_p|z)}, \tag{3.5}$$

where (z_i^-, z_i^+) are the edges of the ith redshift bins. The underlying true distribution is, from Ref. [44]:

$$n(z) \propto \left(\frac{z}{z_0}\right)^2 \exp\left[-\left(\frac{z}{z_0}\right)^{3/2}\right], \tag{3.6}$$

with $z_0 = z_m/\sqrt{2}$, and the normalization is set such that $\int_0^\infty n(z)dz = N$, where N represents the mean surface density of sources in the survey. For

σ_b	c_b	z_b	σ_0	c_0	z_0	$f_{\rm out}$
0.05	1.0	0.0	0.05	1.0	0.1	0.1

Table 3.1: Baseline parameters for the photo-z distribution (3.7)

Euclid the value is $N = 30\,\text{galaxies}/\text{arcmin}^2$
The probability distribution function $p_{ph}(z_p|p)$ represents the probability that a galaxy with a true redshift z has a measured redshift z_p [44]:

$$\begin{aligned} p_{ph}(z_p|z) \;=\; & \frac{1-f_{\rm out}}{\sqrt{2\pi}\sigma_b\,(1+z)}\exp\left\{-\frac{1}{2}\left[\frac{z-c_b z_p-z_b}{\sigma_b\,(1+z)}\right]^2\right\} \\ & +\;\frac{f_{\rm out}}{\sqrt{2\pi}\sigma_0\,(1+z)}\exp\left\{-\frac{1}{2}\left[\frac{z-c_0 z_p-z_0}{\sigma_0\,(1+z)}\right]^2\right\}, \end{aligned} \tag{3.7}$$

where σ_b and z_b are the photometric redshift error distribution for a fraction of sources with reasonably well-measured redshifts and their additive bias, while σ_0 and z_0 are the corresponding quantities for catastrophic outliers. The parameter f_{out} represents the fraction of catastrophic outliers. To produce realistic results for *Euclid*, tephe values of these parameters are summarized in table 3.1 [44]. The total tomographic redshift distribution $n_i(z)$ as defined in eq. (3.5) is shown in Fig. 3.6.
Another possible binning scheme is described by Ref. [45]. They propose to divide the survey into 13 redshift bins with the same width in order to improve the FoM. This recipe is under investigation.
For the work described in this thesis, we follow the prescription described by Ref. [44].

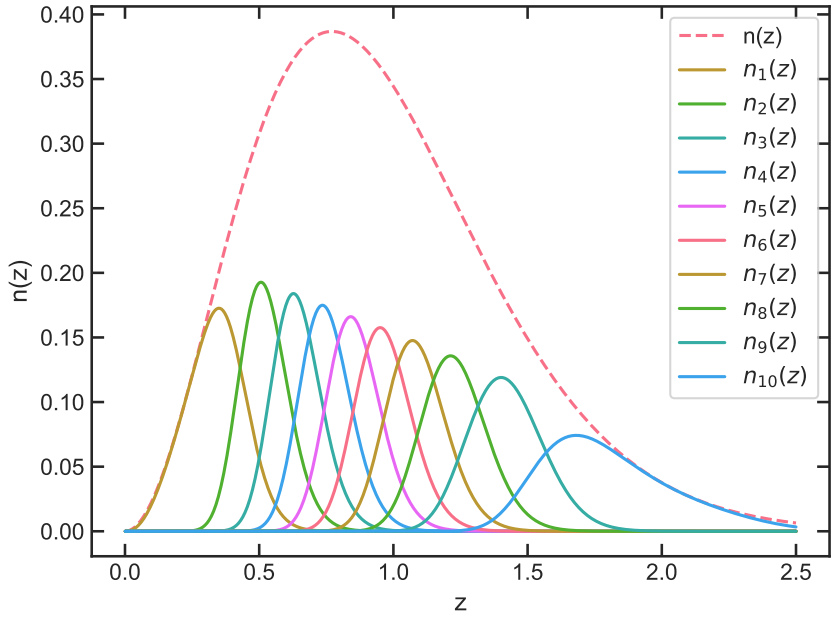


Figure 3.6: The redshift distribution assumed for the galaxies for the *Euclid* photometric survey. The dashed line indicated the true underlying redshift distribution $n(z)$. The solid lines are the redshift distribution $n_i(z)$ for each redshift bin.

Chapter 4

The integrated Sachs-Wolfe effect: a joint analysis of the CMB and the Euclid data

As discussed in 1.4.1, the late-iSW effect experienced by CMB photons arises from the cosmic acceleration at very low redshifts. Therefore, the iSW effect serves as a trace of Dark Energy's presence, which, according to the standard model, drives cosmic acceleration during late times. Or maybe, it might offer a subtle indication of phenomena lying beyond the ΛCDM model, e.g. Modified Gravity models.
The detection of this effect is challenging because the iSW signal is very low and, in addition, it affects very large scales of the CMB anisotropies which are dominated by the *cosmic variance*. On the other hand, the evolution of the gravitational potential perturbations that cause the iSW is related to the dark matter field, as shown in chapter 1. Following this, it is justifiable to assume that a certain correlation exists between the distribution of matter and the resultant pattern of the iSW effect observed in the CMB across the sky. Hence, the joint analysis of CMB and matter field at the epoch in which DE starts dominating is the most prominent analysis to have a significant detection of iSW reducing the effect of uncertainties.
The framework for a joint analysis of two fields is illustrated in Fig. 4.1 [46]. The observable that is exploited to measure the iSW from CMB temperature and galaxy overdensity field is the spherical harmonic cross-correlation power spectrum C_ℓ^{TG}.
The first step is to collect data from different cosmological probes, taken from separate surveys or even by the same survey. Then the data of each probe, after systematics removal, are projected into 2D maps. From the maps are extracted the cross- and auto-correlation power spectra, which capture the cosmological information contained in the two-point statistics of the map. The estimated observable is then fit to the theoretical expectation value by means of a statistical tool (Bayesian inference) to derive constraints

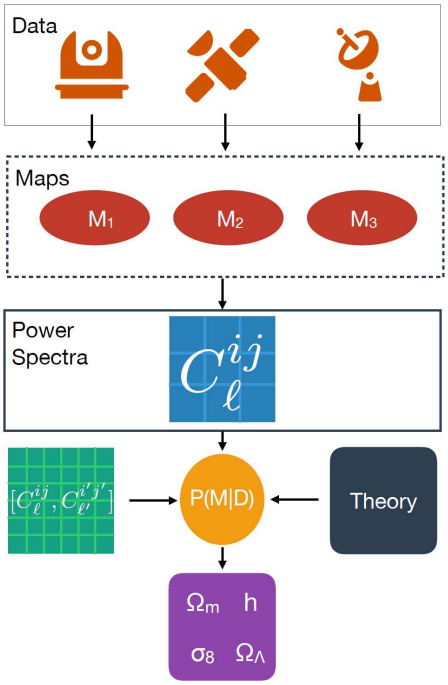


Figure 4.1: Synopsis of the cross-correlation analysis: from the data to the estimation of the cosmological parameters.[46]

on cosmological parameters.
This chapter introduces the work of the *Euclid* CMBxC Science Working Group, dedicated to the cross-correlations of *Euclid* and CMB data. My contribution to the CMBxC Science working group has been the development and validation of the likelihood framework. This means both the theoretical setup and the code to infer constraints on the cosmological parameters. I also worked on building a new estimator for the cross-correlation angular power spectrum (see chapter 5 and 6 for the discussion).
This chapter is organized as follows: in section 4.1 are summarized the advantages of a joint analysis between different probes, such as CMB and GC data from the *Euclid* survey; in section 4.2 I will show the theoretical prescription for cross-correlation observable, the cross-correlation angular power spectrum C_ℓ^{TG} and how it is related to the DE cosmological parameters. The sections 4.3 and 4.4 are dedicated to the estimation of the observable and the associated covariance from the 2D maps of the sky. In the last section 4.5, I define the Signal-to-Noise analysis ratio as a test for the detectability of the iSW effect, forecasting the results the *Euclid* photometric survey expects to reach.

4.1 Scientific motivations for probe combination for the iSW detection

The advantage of cross-correlation analysis has been widely demonstrated (see, for example, [47]). The cross-correlation of different probes helps break the degeneracy between cosmological parameters, as the constraints of each probe in the parameter space might be in different directions. Moreover, combining different datasets with different and/or uncorrelated systematic effects will enhance our knowledge and characterization of such effects. This will definitely boost the detection of the target signal with respect to the analysis of each probe alone.
The *Euclid* GC and the CMB are affected by the same distribution of matter, so we expect a high level of correlation [48]. In the case of the iSW probe, we expect that where the inhomogeneities in the matter field are evolving faster, the signal in the CMB spectrum is higher.
Over the last decades, the cross-correlation analysis between the large-scale CMB temperature anisotropies and GC has been explored with a large variety of different models: LSS surveys with different wavelength ranges (optical SDSS [26], radio NVSS [49], near-infrared 2MASS [50]) combined with different CMB surveys, from COBE to WMAP and Planck. Many works in the literature employ different estimators of the cross-correlated observable and different statistical methods to analyze the data. For example, the Planck 2015 [51] shows the detection of the iSW effect at $\sim 2.7\sigma$ significance when cross-correlating CMB temperature and polarization with the Sloan Digital Sky Survey (SDSS) catalog. The significance improves by combining different LSS surveys and even more by combining CMB temperature and polarization with CMB lensing. Taking advantage of this joint analysis, they also claim the detection of Ω_Λ at 3σ. What makes the iSW measurement difficult is the intrinsic weakness of the signal with respect to the other CMB anisotropies. Hence it is crucial to have perfect control of the noise and of the systematics.
Future experiments will put tighter constraints on the iSW effect and, more generally, on the cosmological parameters, thanks to both the large amount of data available and the higher statistical power of the analysis. Forecasted results on the iSW detection using future CMB and LSS have been largely explored in the literature (*Euclid*, LiteBird, Simons Observatory, LSST, SKA), see Refs. [52], [53] and [54]. For example, Fig. 4.3 from ref. [54] shows the improved constraints on the dark energy equation of state parameters, w_0 and w_a when including the cross-correlation between *Euclid* and CMB probes. In [53], it is shown that a joint analysis of *Euclid* with CMB probes (in this case, temperature and lensing), including the cross-correlation, will help substantially in reducing the impact of nuisance parameters on cosmological constraints and in improving the constraints on all cosmological parameters (see Fig. 4.2).

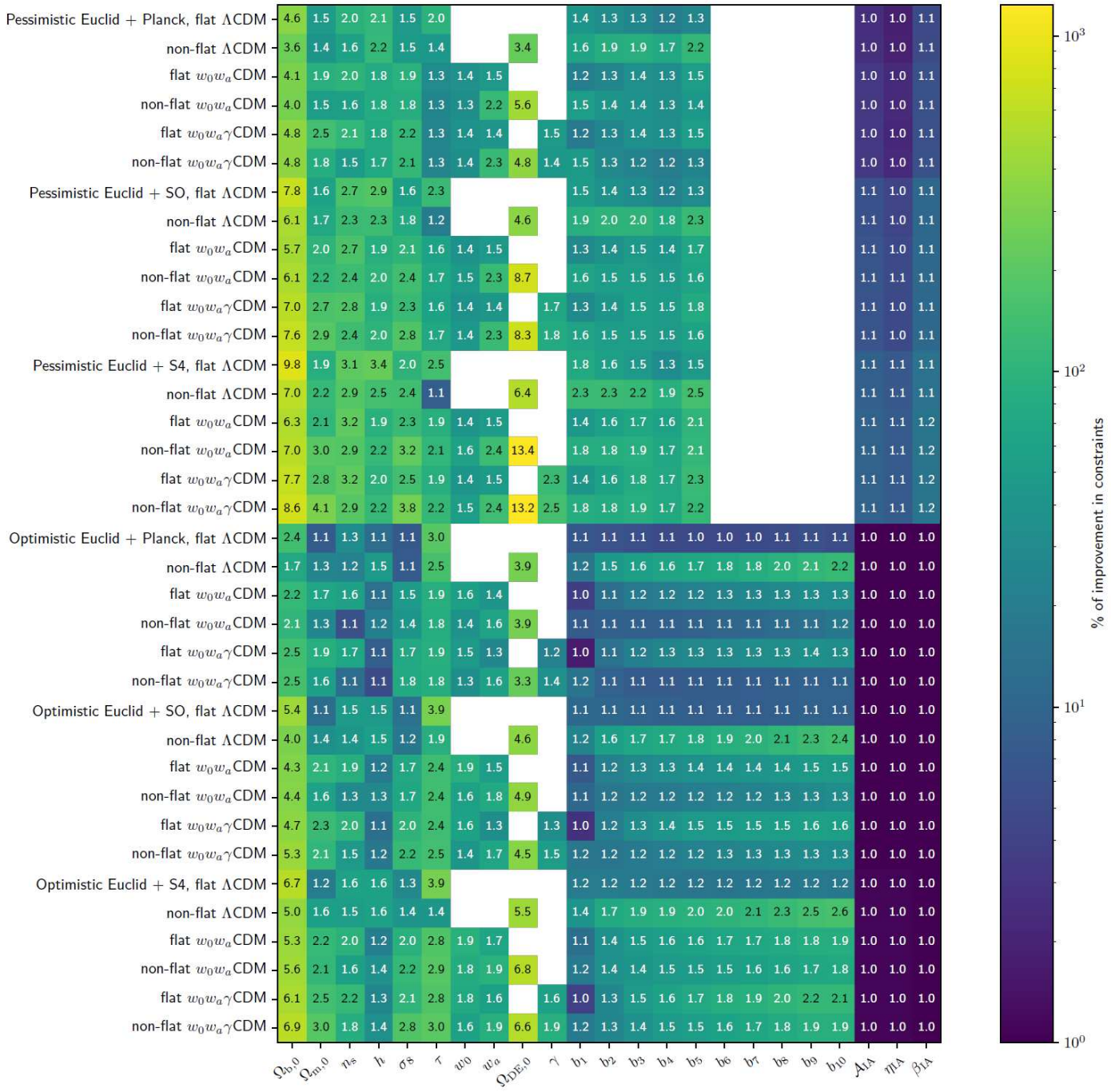


Figure 4.2: Predicted improvement of parameter constraints after adding CMB probes (temperature, polarization and lensing) to *Euclid*, including the cross-correlations. The color coding reflects the percentage improvement defined as $(\sigma_{\text{after}}/\sigma_{\text{before}} - 1) \times 100$, where σ_{after} and σ_{before} are the 1σ uncertainty before and after adding the CMB probes. The numbers in each square correspond to the factor of improvement $\sigma_{\text{after}}/\sigma_{\text{before}}$. Here are shown the results for different models (flat ΛCDM, non-flat ΛCDM, flat w_0w_aΛCDM, non-flat w_0w_aΛCDM. The parameters considered are the baseline ΛCDM parameters and its extension (Ω_Λ, w_0, w_a, γ), plus some nuisance parameters.) [53].

The detection of the iSW is highly dependent on the deepness of the survey considered and on its coverage, as well as the redshift distribution. Paper [54] shows that the signal-to-noise ratio for the cross-correlation analysis of *Euclid* and a CMB experiment (Planck [5], LiteBird [55], Simos Observatory (SO) [56] , CMB stage4 [57]) will reach $\sim 4\sigma$ value with the tomographic analysis with 10 redshift bins.

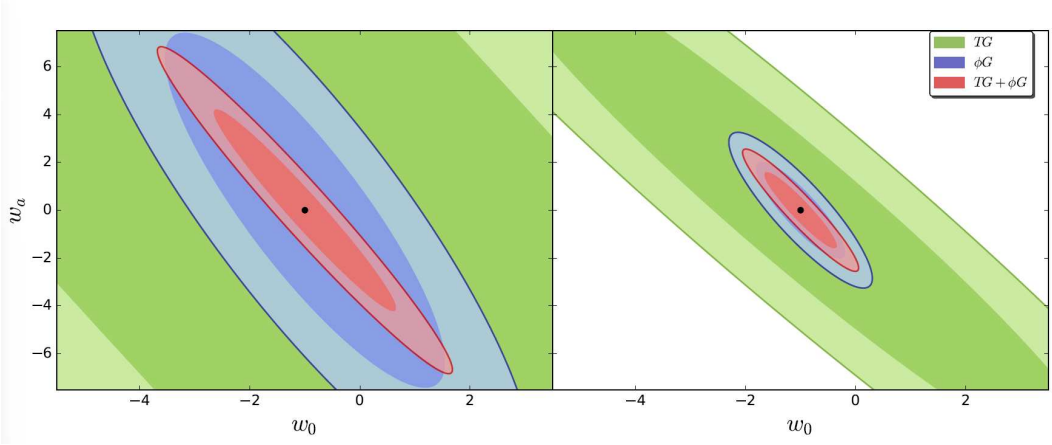


Figure 4.3: Marginalized 68% and 95% 2D confidence region for the constraints from the cross-correlation between an LSS survey and the CMB temperature (*TG*) and lensing (ϕG). The left panel corresponds to the analysis with Planck and *Euclid*, and the right panel to LiteBIRD+S4 and SKA1. The green contours correspond to the *TG* cross-correlation constraints (TG) the blue contours to the ϕG cross-correlation and the red contours to the sum of both. For this plot, all the other cosmological parameters are kept fixed to Planck 2018 values. [54].

4.2 The cross-correlation observable

Virtually all cosmological observations, spanning from CMB to galaxy surveys, return us with samples of 3D random fields projected onto the sphere, primarily presented as two-dimensional sky maps. This applies especially when it is complicated to retrieve information about the redshift of some sources. So it is natural to perform the spherical harmonics analysis on such products. A common observable is the angular power spectrum C_ℓ that, as said before, encodes all the statistical properties of a given field[1] The theoretical cross-correlation angular power spectrum C_ℓ^{TG} between the CMB temperature and galaxy overdensity is the ensemble average of the two fields, defined as: Given a generic 3D field $\delta^{3D,X}$, the 2D projected field

[1]This is true under the hypothesis of Gaussian random field, otherwise high-order statics is required.

$\Delta^X(\hat{\mathbf{n}})$ is defined as:

$$\Delta^X(\hat{\mathbf{n}}) = \int d\chi W^X(\chi)\delta^{3D,X}(\chi\hat{\mathbf{n}}, \eta(\chi))\,, \tag{4.1}$$

where $W^X(\chi)$ is some arbitrary weight function. Thus, the coefficient of the spherical harmonics expansion is:

$$\Delta^X_{\ell m} = \int d\Omega \Delta^X(\hat{\mathbf{n}}) Y^*_{\ell m}(\hat{\mathbf{n}}) \tag{4.2}$$

Using Rayleigh's expansion

$$e^{i\mathbf{k}\cdot\hat{\mathbf{n}}\chi} = 4\pi \sum_{\ell m} i^\ell j_\ell(k\chi) Y^*_{\ell m}(\hat{\mathbf{k}}) Y_{\ell m}(\hat{\mathbf{n}})\,, \tag{4.3}$$

we can Fourier transform $\delta^{3D,X}$ and obtain:

$$\begin{aligned}\delta^{3D,X}(\chi\hat{\mathbf{n}}, \eta(\chi)) &= \int \frac{d^3k}{(2\pi)^3}\delta^{3D,X}(\mathbf{k}, \eta(\chi)) e^{i\mathbf{k}\cdot\hat{\mathbf{n}}\chi} \\ &= 4\pi \sum_{\ell m} i^\ell \int \frac{d^3k}{(2\pi)^3}\delta^{3D,X}(\mathbf{k}, \eta(\chi)) j_\ell(k\chi) Y^*_{\ell m}(\hat{\mathbf{k}}) Y_{\ell m}(\hat{\mathbf{n}})\,.\end{aligned} \tag{4.4}$$

The projected field is:

$$\Delta^X(\hat{\mathbf{n}}) = 4\pi \sum_{\ell m} i^\ell \int d\chi W^X(\chi) \int \frac{d^3k}{(2\pi)^3}\delta^{3D,X}(\mathbf{k}, \eta(\chi)) j_\ell(k\chi) Y^*_{\ell m}(\hat{\mathbf{k}}) Y_{\ell m}(\hat{\mathbf{n}})\,, \tag{4.5}$$

and the spherical harmonics coefficients read as:

$$\Delta^X_{\ell m} = 4\pi^\ell \int d\chi W(\chi) \int \frac{d^3k}{(2\pi)^3}\delta^{3D,X}(\mathbf{k}, \eta(\chi)) j_\ell(k\chi) Y^*_{\ell m}(\hat{\mathbf{k}}) Y_{\ell m}\,, \tag{4.6}$$

where χ is the comoving distance, $\eta(\chi)$ the comoving horizon, $\hat{\mathbf{n}}$ the direction of observations, $\mathbf{k}$ the mode of the perturbations, $Y_{\ell m}$ the spherical harmonics, and j_ℓ the Bessel functions. Assuming isotropy, the theoretical cross-correlation angular power spectrum C_ℓ^{TG} between the CMB temperature and galaxy overdensity is:

$$C_\ell^{\mathrm{TG}} = \langle \Delta^{T,\mathrm{iSW}}_{\ell m} \Delta^{*\,G}_{\ell' m'} \rangle\,. \tag{4.7}$$

For the CMB field, we have already calculated the transfer function $\Theta_\ell(k, \eta_0)$ (see eq. 1.74). Taking only the iSW contribution and rewriting [2], we obtain:

$$\begin{aligned}\Theta_\ell(k, \eta_0) &= 2\int_0^{\eta_0} d\tilde{\eta} e^{-\tau(\tilde{\eta})} \dot{\psi}(k, \tilde{\eta}) j_\ell[k(\eta_0 - \tilde{\eta})] \\ &= 2\int_0^{\eta_0} d\tilde{\eta} e^{-\tau(\tilde{\eta})} \frac{3}{5}\mathcal{R}(\mathbf{k}) T(k) \frac{d}{d\tilde{\eta}}\left(\frac{D_+(a)}{a}\right) j_\ell[k(\eta_0 - \tilde{\eta})] \\ &= -3\frac{\Omega_m H_0^2}{k^2}\frac{1}{D_+(a)}\int_0^{\eta_0} d\tilde{\eta} e^{-\tau(\tilde{\eta})} \frac{d}{d\tilde{\eta}}\left(\frac{D_+(a)}{a}\right) \delta_m(\mathbf{k}, a(\tilde{\eta})) j_\ell[k(\eta_0 - \tilde{\eta})]\,,\end{aligned} \tag{4.8}$$

[2] We use the fact that the anisotropic stress is negligible, i.e. $\phi = \psi$

In the second equality, we divide the evolution of the gravitational potential into the transfer function $T(k)$ and the growth factor $D_+(a)$ (see the subsection 1.3.3) To obtain the last equality, we rewrite the Poisson equation (see (1.48)) as, with $\rho = \rho_m = \Omega_m \rho_{crit}/a^3$ in a matter-dominated Universe and $4\pi G\rho_{crit} = (3/2)H_0^2$:

$$\psi(\mathbf{k}, a) = -\frac{3\Omega_m H_0^2}{2k^2 a}\delta_m(\mathbf{k}, a(\eta))\,, \tag{4.9}$$

and thus, putting it all together,

$$\mathcal{R}(\mathbf{k}) = -\frac{5}{2}\frac{\Omega_m H_0^2}{k^2}\frac{1}{T(k)D_+(a)}\delta_m(\mathbf{k}, a)\,. \tag{4.10}$$

We can define the $W^{\rm iSW}$ as:

$$W^{\rm T,iSW}(k,\chi) \equiv 3\frac{\Omega_m H_0^2}{k^2}\frac{1}{D_+(a)}e^{-\tau}\frac{d}{d\chi}\left(\frac{D_+(a)}{a}\right)\,. \tag{4.11}$$

Now we can compare eq. (4.8) with (4.6) and obtain the first component for the $C_\ell^{\rm TG}$:

$$\Delta_{\ell m}^{\rm T,iSW}(k) = 4\pi^\ell \int \frac{d^3k}{(2\pi)^3} Y_{\ell m}^*(\hat{\mathbf{k}}) Y_{\ell m} \int d\chi W^{\rm T,iSW}(k,\chi) j_\ell(k\chi)\delta_m(\mathbf{k}, \chi(\eta))\,. \tag{4.12}$$

The galaxy overdensity $\delta_{\rm G,obs}(\mathbf{k},\eta)$ [3] is related to the cold dark matter field through the galaxy bias $b(\chi(\eta))$ as $\delta_{\rm G,obs}(\mathbf{k},\eta) = b(\chi(\eta))\delta_m(\mathbf{k},\chi(\eta))$ (see eq. (1.75)), so we can write eq. (4.6) as:

$$\Delta_{\ell m}^{\rm G}(k) = 4\pi^\ell \int d\chi W^{\rm G}(\chi) \int \frac{d^3k}{(2\pi)^3} b(\chi(\eta))\delta_m(\mathbf{k},\chi(\eta)) j_\ell(k\chi) Y_{\ell m}^*(\hat{\mathbf{k}}) Y_{\ell m}\,, \tag{4.13}$$

where we define $W^{\rm G}(\chi)$ as:

$$W^{\rm G}(\chi) = \frac{dN}{d\chi}\,, \tag{4.14}$$

which is the distribution of galaxies. Here N is the total number of galaxies. Finally, the cross-correlation angular power spectrum $C_\ell^{\rm TG}$ is:

$$\begin{aligned} C_\ell^{\rm TG} &= \langle \Delta_{\ell m}^{T,\rm iSW} \Delta_{\ell' m'}^{*\,G} \rangle \\ &= \frac{2}{\pi}\int_0^\infty dk k^2 \int_0^\infty d\chi W^{\rm T,iSW}(k,\chi) j_\ell(k\chi) \times \\ &\times \int_0^\infty d\chi' b(\chi'(\eta)) W^{\rm G}(\chi') j_\ell(k\chi') P(k,\chi,\chi') \end{aligned} \tag{4.15}$$

[3] Subscription *obs* stands for *observed*

The matter power spectrum $P(\mathbf{k}, \chi, \chi')$ comes from $\langle \delta_m(\mathbf{k}, \chi) \delta_m(\mathbf{k}, \chi') \rangle$, under the assumption of isotropy (see section 1.3.3) Changing the integral variable from χ to z [4] and defining the *kernel*s $G^{\mathrm{T,iSW}}(z)$, and $G^{\mathrm{G}}(z)$ as

$$G^{\mathrm{T,iSW}}(z) \equiv -3\Omega_m H_0^2 e^{-\tau} \frac{d}{dz} \left(\frac{D_+(z)}{a(z)} \right) \tag{4.16a}$$

$$G^{\mathrm{G}}(z) \equiv \frac{dN}{dz} D_+(z) b(z) \,, \tag{4.16b}$$

the C_ℓ^{TG} becomes:

$$C_\ell^{\mathrm{TG}} = \frac{2}{\pi} \int_0^\infty dk P_M(k, z=0) \int_0^\infty dz G^{\mathrm{T,iSW}}(z) j_\ell(k\chi(z)) \int_0^\infty dz' G^{\mathrm{G}}(z) j_\ell(k\chi(z')) \,, \tag{4.17}$$

where we use the scaling properties of the $P_M(k, z)$:

$$P_M(k, z=0) = \frac{P_M(k, z, z')}{D_+(z) D_+(z')} \,. \tag{4.18}$$

The term $\frac{dN}{dz}$ is the redshift distribution of galaxies. For a realistic survey, we know from section 3.3.2 that the measured redshift distribution is affected by uncertainties, so for *Euclid* $\frac{dN}{dz}$ will be eq. 3.5.
Here we are neglecting all the relativistic corrections that enter the galaxy overdensity (such as redshift space distortions, radial term, magnification bias, ...). For a complete analysis see Ref. [58]. In Ref. [54], the authors evaluate the impact of neglecting the correction to the number counts. They found that, for the cross-correlation power spectrum, the larger effect is at higher redshift ($z \sim 2.95$) and large scales ($\ell \lesssim 100$). Nevertheless, the constraints on the cosmological parameters do change by only $\sim 1 - 2\%$ with respect to the case when all corrections are included. The situation is much worse for the galaxy-galaxy power spectrum: the errors of the cosmological parameters can vary around $\sim 10\%$ when constrained by C_ℓ^{GG} alone. Fig. 4.4 shows the dimensionless[5] C_ℓ^{TG} angular power spectra for different redshift bins. The power spectra are computed given the *Euclid* prescription in eq. (3.5) as the redshift distribution for the galaxies, and the cosmological parameters are fixed to the values reported in table 4.1 [44], for a ΛCDM cosmology. For completeness, the galaxy auto-angular power spectrum C_ℓ^{GG} can be derived from eq. (4.16b) as:

$$C_\ell^{\mathrm{GG}} = \frac{2}{\pi} \int_0^\infty dk P_M(k, z=0) \int_0^\infty dz G^{\mathrm{G}}(z) j_\ell(k\chi(z)) \int_0^\infty dz' G^{\mathrm{G}}(z) j_\ell(k\chi(z')) \tag{4.19}$$

[4] $d\chi = \frac{1}{H(z)} dz = -d\eta$

[5] Usually, the angular power spectrum C_ℓ^{TG} can be expressed as $[\mu K]$ multiplying it by the CMB temperature.

H_0	$\Omega_b h^2$	$\Omega_c h^2$	$10^9\, A_s$	n_s	τ	$\sum m_\nu$
67	0.0224	0.12	2.1115	0.96	0.058	0.06

Table 4.1: Cosmological parameters for the fiducial cosmology (ΛCDM). This is the *Euclid* forecast prescription, see [44].

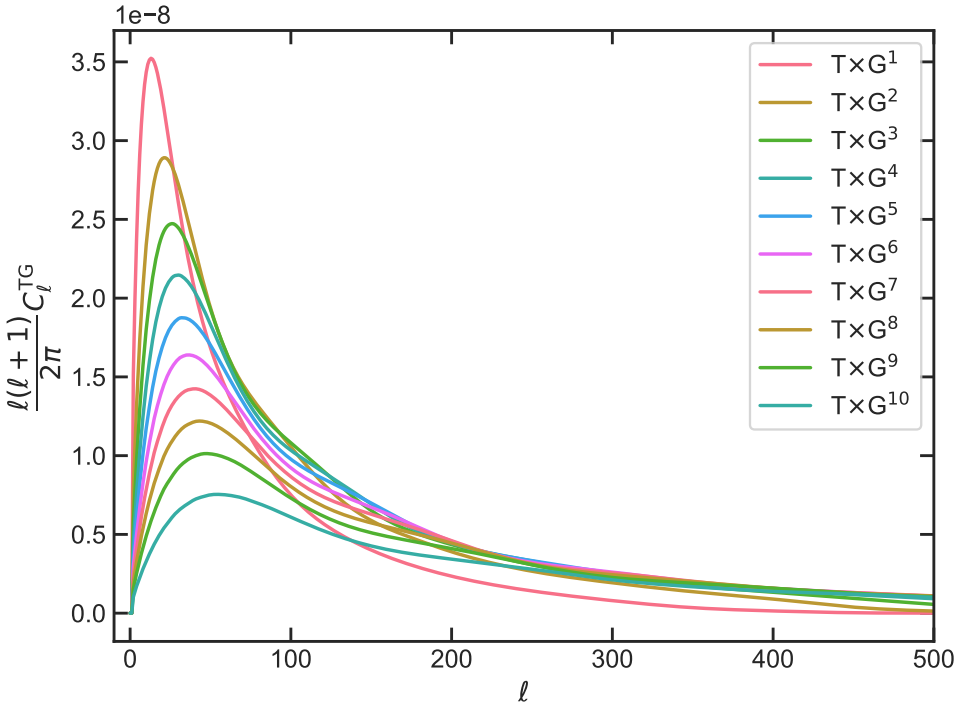


Figure 4.4: Cross-correlation angular power spectrum C_ℓ^{TG} for different redshift bins, according to eq. (3.5) . The cosmological parameters (ΛCDM cosmology) are fixed to the values reported in table 4.1 .

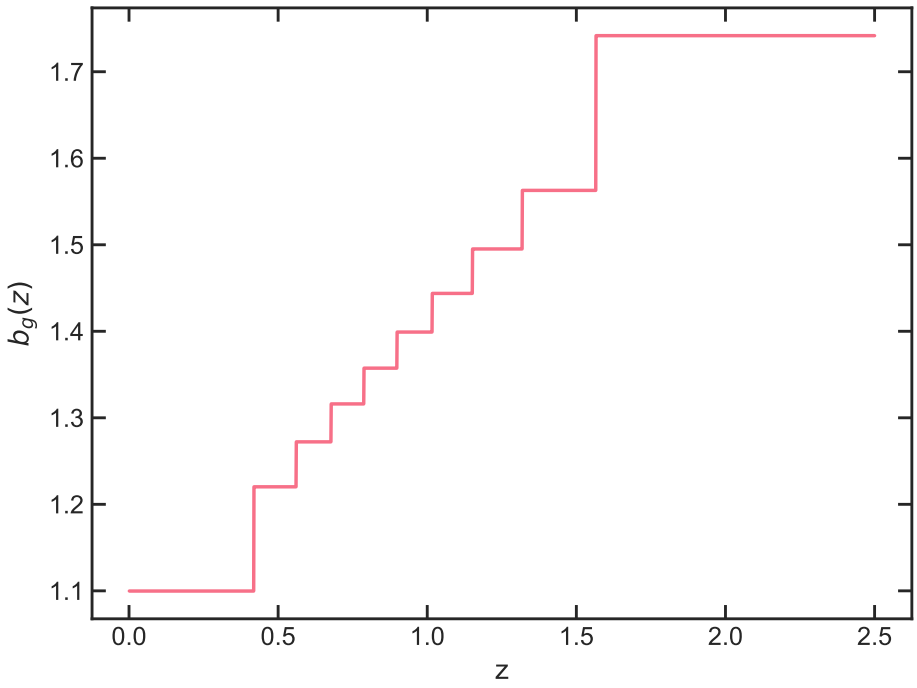


Figure 4.5: The galaxy bias $b_g(z)$ as defined in eq. (4.21)

Linear galaxy bias The relation between the underlying dark matter distribution δ_m and the observed galaxy overdensity $\delta_{\mathrm{G,obs}}$ is assumed to be linear and scale-independent:

$$\delta_{\mathrm{G,obs}}(\mathbf{k}, \eta) = b(\chi(\eta))\delta_m(\mathbf{k}, \chi(\eta))\,. \tag{4.20}$$

As said in the subsection 1.4.2, the linear bias approximation is sufficiently accurate at large scales. For the *Euclid* photometric survey, paper [44] assumes a constant galaxy bias in each tomographic redshift bin:

$$b_i = \sqrt{1 + \bar{z}}\,, \tag{4.21}$$

where i is the index of the redshift bin and $\bar{z}$ is the mean redshift on each bin. When producing a forecast, for example by mean of the Monte Carlo Markov Chain (MCMC), the values of b_i are considered nuisance parameters to marginalize over.
However, the main assumption on the galaxy bias in the scale-independence. Removing this would require the addition of nuisance parameters to control the uncertainties on the galaxy bias, leading to a decrease of the constraining power from the photometric galaxies clustering survey [44].
The galaxy bias for the tomographic case is shown in Fig. 4.5.

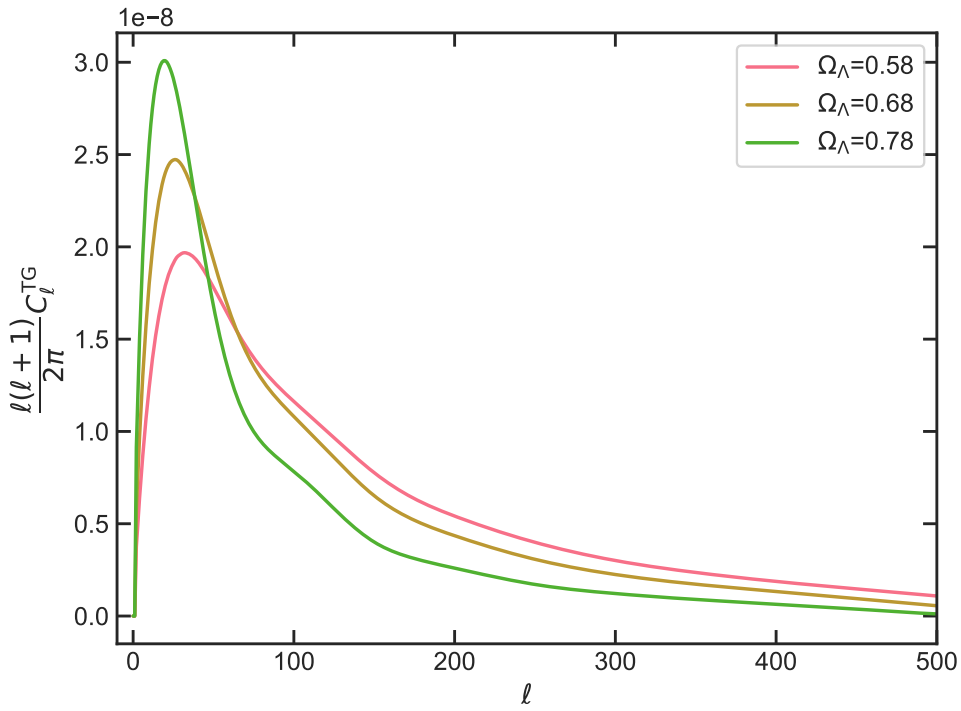


Figure 4.6: The angular power spectrum C_ℓ^{TG} for different values of the DE density parameter Ω_Λ. Higher values of Ω_Λ increase the height of the peak of the spectrum.

Dependence on the Dark Energy parameters The angular power spectrum C_ℓ^{TG} is dependent on the non-relativistic matter density parameter Ω_m and on the derivative of the growth $D_+(z)$ factor over the scale factor a. These parameters are indirectly related to the DE energy density parameter Ω_Λ and the DE equation of state $w(a) = \frac{\mathcal{P}}{\rho}$.
In fact, Ω_m is the total energy density contribution from the non-relativistic matter (cold dark matter + baryons) and thus, from eq. (1.19), neglecting the radiation contribution and curvature, we have that $\Omega_\Lambda = 1 - \Omega_m$.
As is shown in Fig. 4.6, the amplitude and height of the peak of the C_ℓ^{TG} change as Ω_Λ changes. Specifically, higher values of DE density correspond to higher peaks of the cross-correlation spectrum. This is expected since a large amount of DE tends to increase the evolution of the gravitational potential and of the dark matter perturbations, and thus the correlation with CMB photons. The behavior of $\frac{D_+(a)}{a}$ is shown in Fig. 4.7. It depends on both Ω_Λ and the equation of state $w(a)$. Increasing the amount of dark energy leads to greater suppression of the growth at late times. Fig. 4.8 shows the same dependence on w_0 and w_a. To conclude this discussion, we can also comment Fig. 4.9, which shows the percentage relative differences

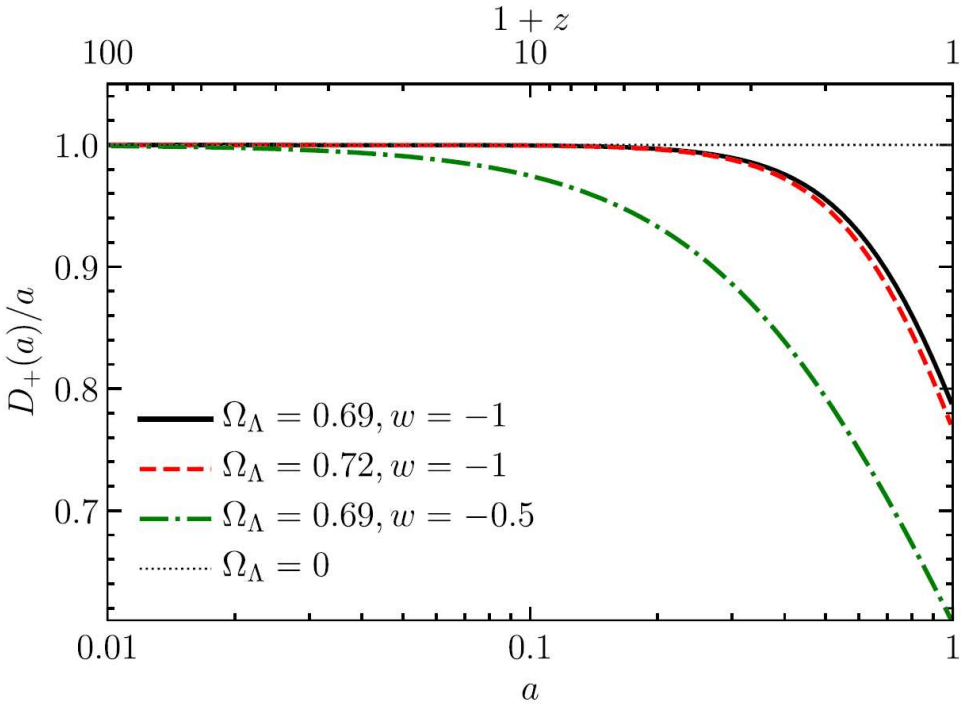


Figure 4.7: The growth factor $D_+(a)$ divided by the scale factor calculated at different values of Ω_Λ and $w(a)$. The solid line shows the fiducial cosmology (ΛCDM model with the cosmological values fixed at the Planck best-fit values). Plot taken from Ref. [15].

between $C_\ell^{\rm TG}$ calculated at the fiducial value $\Omega_\Lambda = 0.68$ and $C_\ell^{\rm TG}$ computed at $\Omega_\Lambda = 0.58$, slightly lower value, for different redshift bins. All the other cosmological parameters are fixed at the fiducial values reported in table 4.1. Qualitatively, it can be observed that varying DE density has a higher impact on the first bin from which we expect to have a higher signal. This behavior is expected since DE is dominating the energy content at $z_\Lambda \sim 1$.

Limber approximation The exact expression of the $C_\ell^{\rm TG}$ (4.17) is difficult to handle, because of the rapid oscillation of Bessel functions at high multipoles. However, the matter power spectrum, $P_M(k)$, varies slowly in a narrow range, Δk, with respect to the Bessel function. In this respect, we can consider $P_M(k)$ as independent of k and constant at the value $k \approx (\ell + 1/2)/\chi$. This is usually called *Limber approximation* [59]. Under this approximation, the integral over the spherical Bessel functions becomes:

$$\int k^2 j_\ell(k\chi) j_\ell(k\chi') = \frac{\pi}{2\chi^2}\delta_D(\chi - \chi')\,, \tag{4.22}$$

which means that the integrand is dominated by the region $\chi \approx \chi'$. In [60] it is demonstrated that the error is $\sim \mathcal{O}(\ell^{-2})$, so the approximation works for smaller scales (low values of ℓ). The approximated $C_\ell^{\rm TG}$, with the

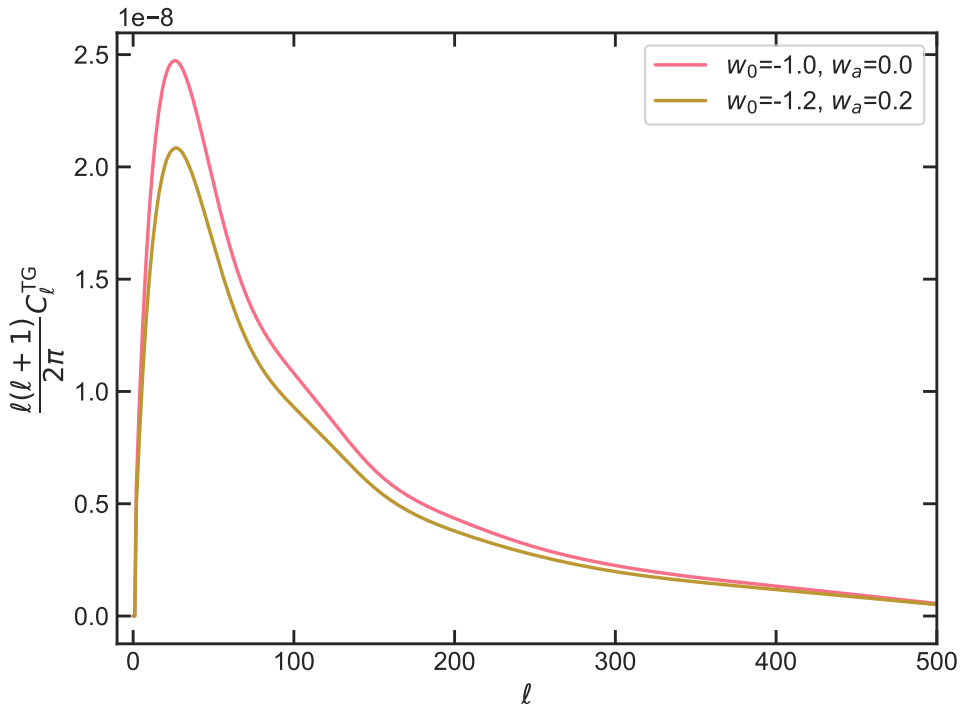


Figure 4.8: The angular power spectrum C_ℓ^{TG} for different values of the DE equation of state parameters w_0 and w_a. As for Ω_Λ, we see that these parameters change the shape of the spectrum.

integration in redshift space, is [61]:

$$C_\ell^{TG} = \frac{3\Omega_m H_0^2}{(\ell+\frac{1}{2})^2}\int dz b_g(z)\frac{dN}{dz}H(z)D(z)\frac{d}{dz}\left(\frac{D_+(z)}{a(z)}\right)P_M\left(k=\frac{\ell+\frac{1}{2}}{\chi(z)}\right). \tag{4.23}$$

The Limber approximation has the great advantage of reducing the computational time. Fig. 4.10 illustrates the percentage relative difference for each redshift bin between C_ℓ^{TG} computed without any approximation, i.e. as in eq. (4.7), and C_ℓ^{TG} computed with the Limber approximation turned on at every multipole (starting from $\ell = 1$), showing less than one per mille differences.

Hence, even if Limber approximation is not accurate for low ℓ, which is the scale at which we expect the iSW signal to peak, the error resulting from using this approximation is very low, especially when compared to the cosmic variance. In the next chapter (see section 5.5) will be discussed the impact of the Limber approximation on the CMBxC likelihood and it will be shown that the error induced by this approximation in subdominant with respect to the cosmic variance.

Another point of discussion could revolve around the behavior of the iSW

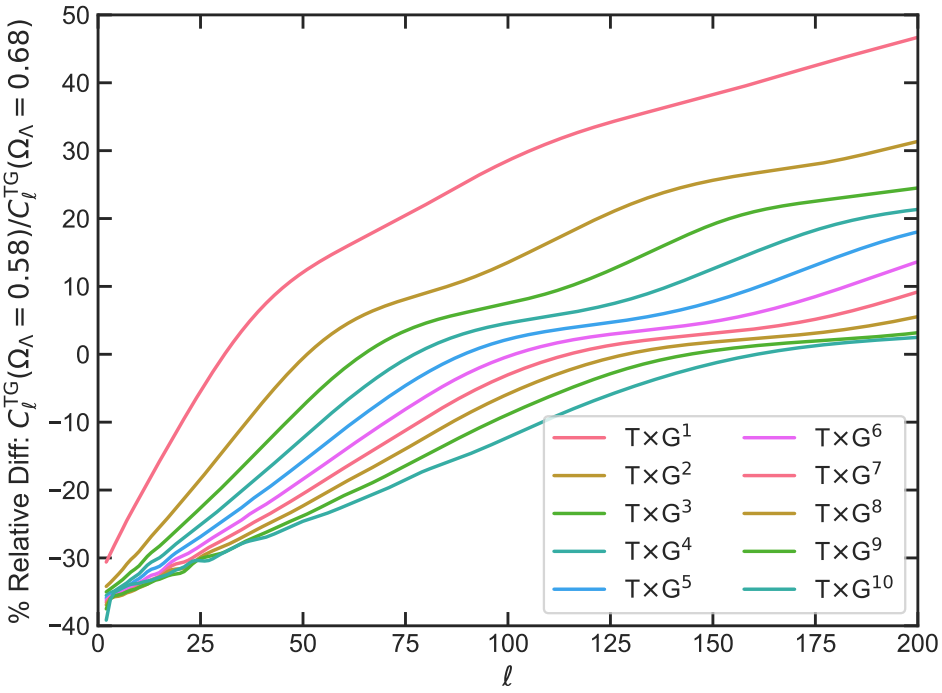


Figure 4.9: Percentage relative difference between the C_ℓ^{TG} calculated at $\Omega_\Lambda = 0.58$ and the C_ℓ^{TG} for Ω_Λ at the fiducial value, for different redshift bins. The cosmological parameters are fixed in table 4.1

weight function $W^{iSW}(k,\chi)$: both the $P_M(k)$ and the $W^{iSW}(k,\chi)$ are changing slower with respect to the product of the Bessel function and thus can be taken out of the integrals.

4.3 Power spectrum estimators

Given a 2D map of CMB temperature and the galaxy's overdensity, we are unable to extract the exact angular power spectrum as described in eq. (4.17). In analogy with the section 1.4.1, an unbiased estimator of the cross-correlation angular power spectrum is:

$$\hat{C}_\ell^{TG} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} a_{\ell m}^T a_{\ell m}^{*\,G}\,, \tag{4.24}$$

where the variables $(a_{\ell m}^T, a_{\ell m}^G)$ follow a Gaussian distribution. Thanks to their orthogonality properties, the spherical harmonics are particularly suitable when dealing with full-sky Gaussian distributed sky maps, with a diagonal covariance matrix, i.e. $\propto \delta_{\ell\ell'}^K$

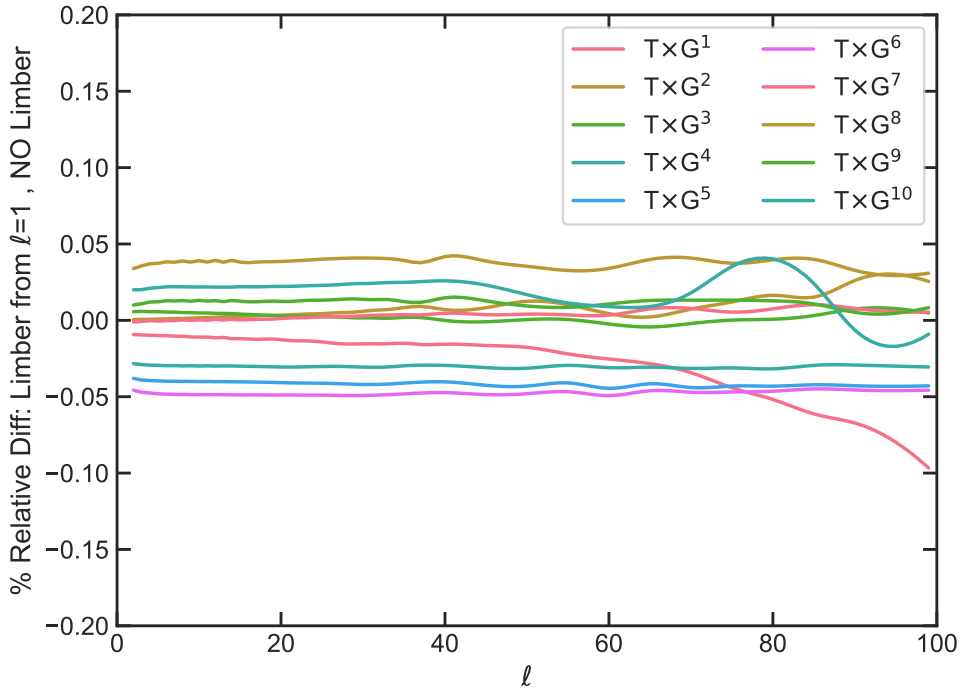


Figure 4.10: Relative percentage difference between the exact cross-correlation angular power spectrum $C_\ell^{\rm TG}$ in eq. (4.7) and the $C_\ell^{\rm TG}$ computed with the Limber approximation turned on from $\ell = 1$. The results are shown for each redshift bin.

However, in realistic experiments, only a portion of the sky is observable. The observational strategy scans a limited portion of the sky and introduces non-uniform and potentially correlated noise that contaminates the observed signal. The incomplete sky coverage induces a mode-coupling and thus the resulting $a_{\ell m}$ are correlated. The Galactic plane is usually masked in CMB experiments to cover the bright foreground emission that could contaminate the CMB signal.
In [62] and [63] the authors propose a method to extract the underlying true angular power spectrum from a CMB map in the case of partial sky coverage. The idea is to estimate from the masked maps the angular power spectrum, $\tilde{C}_\ell^{\rm TG}$, for the cross-correlations:

$$\tilde{C}_\ell^{\rm TG} = \frac{1}{2\ell+1}\sum_m \tilde{a}^T_{\ell m}\tilde{a}^G_{\ell m}\,, \tag{4.25}$$

that is related to the underlying true power spectrum $C_\ell^{\rm TG}$ as:

$$\langle \tilde{C}_\ell^{\rm TG}\rangle = \sum_{\ell'} M_{\ell\ell'} C_{\ell'}^{\rm TG}\,, \tag{4.26}$$

where $M_{\ell\ell'}$'is the *coupling matrix* that depends on the power spectrum of the mask applied to the map, $W(\hat{\mathbf{n}})$. Eq. (4.25) defines the so-called *pseudo-C_ℓ estimator*. If the coupling matrix is invertible, the recovered power spectrum is:

$$\hat{C}_\ell^{\mathrm{TG}} = M_{\ell\ell'}^{-1}\tilde{C}_{\ell'}^{\mathrm{TG}}\,. \tag{4.27}$$

This applies in the case of small enough sky cuts, i.e. until $\sim 30^\circ$above and below the Galactic plane [64], for larger values the matrix will be singular. On the contrary, if the portion of the observed sky is large, i.e. large values of the *sky fraction* $f_{\mathrm{sky}} = \frac{1}{4\pi}\int_S W^2(\hat{\mathbf{n}})d\Omega$, eq. (4.26) becomes:

$$\langle\tilde{C}_\ell^{\mathrm{TG}}\rangle \approx f_{\mathrm{sky}}C_\ell^{\mathrm{TG}}\,, \tag{4.28}$$

which is the so-called f_{sky} approximation [65].
Other possible estimators of the C_ℓ are, for example, the *Quadratic Maximum Likelihood* estimator (QML) [66] and the *needlets-based* estimator [67] .
We have developed and validated an estimator for the cross-correlation angular power spectrum between CMB photons and galaxy overdensity field based on the needlets function, a type of spherical wavelets. This will be the topic of the chapter 6.

4.4 Covariance estimators

The covariance matrix contains the relevant information about the error associated with the estimate of the power spectrum and the significance of the measurement. For this reason, as discussed in the chapter 5, it is one of the ingredients needed for the inference of the constraints on the cosmological parameters
Assuming a Gaussian distribution for the observables, the covariance matrix of the power spectrum estimator $\hat{C}_\ell^{\mathrm{TG}}$, defined in eq. (4.24) , is:

$$\begin{aligned}
\mathrm{Cov}(\hat{C}_\ell^{\mathrm{TG}},\hat{C}_{\ell'}^{\mathrm{TG}}) &= \langle(\hat{C}_\ell^{\mathrm{TG}} - C_\ell^{\mathrm{TG}})(\hat{C}_{\ell'}^{\mathrm{TG}} - C_{\ell'}^{\mathrm{TG}})\rangle \\
&= \frac{1}{(2\ell+1)(2\ell'+1)}\sum_{mm'}\langle a_{\ell m}^T a_{\ell m}^{G\,*} a_{\ell' m'}^T a_{\ell' m'}^{G\,*}\rangle + C_\ell^{\mathrm{TG}}C_{\ell'}^{\mathrm{TG}} \\
&- C_\ell^{\mathrm{TG}}C_{\ell'}^{\mathrm{TG}} \\
&= \frac{1}{(2\ell+1)(2\ell'+1)}(\langle a_{\ell m}^T a_{\ell m}^{G\,*}\rangle\langle a_{\ell' m}^T a_{\ell' m}^{G\,*}\rangle + \langle a_{\ell m}^T a_{\ell' m'}^{T\,*}\rangle\langle a_{\ell m}^{G\,*} a_{\ell' m}^G\rangle \\
&+ \langle a_{\ell m}^T a_{\ell' m'}^{G\,*}\rangle\langle a_{\ell' m'}^T a_{\ell m}^{G\,*}\rangle) - C_\ell^{\mathrm{TG}}C_{\ell'}^{\mathrm{TG}} \\
&= C_\ell^{\mathrm{TG}}C_{\ell'}^{\mathrm{TG}} + \frac{C_\ell^{\mathrm{TT}}C_{\ell'}^{\mathrm{GG}}}{(2\ell+1)}\delta_{\ell\ell'}^K + \frac{C_\ell^{\mathrm{TG}}C_{\ell'}^{\mathrm{TG}}}{(2\ell+1)}\delta_{\ell\ell'}^K - C_\ell^{\mathrm{TG}}C_{\ell'}^{\mathrm{TG}} \\
&= \frac{\delta_{\ell\ell'}^K}{(2\ell+1)}\left[C_\ell^{\mathrm{TG}}C_{\ell'}^{\mathrm{TG}} + C_\ell^{\mathrm{TT}}C_{\ell'}^{\mathrm{GG}}\right]\,.
\end{aligned} \tag{4.29}$$

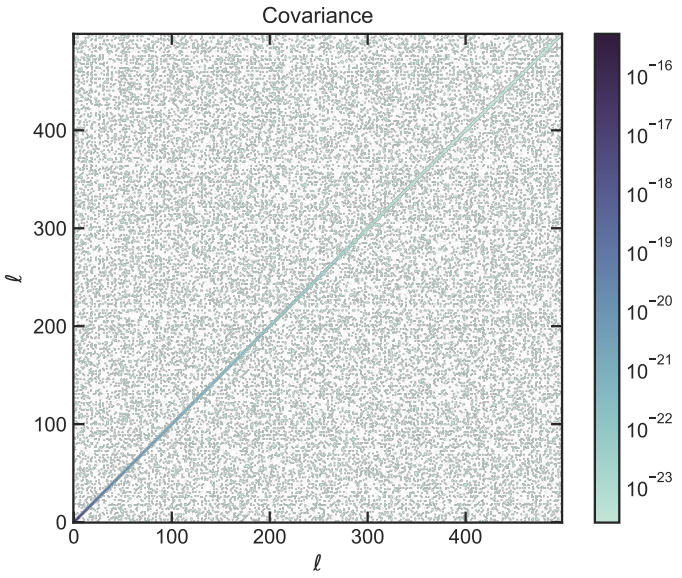


Figure 4.11: The covariance $\mathrm{Cov}(\hat{C}_\ell^{\mathrm{TG}}, \hat{C}_{\ell'}^{\mathrm{TG}})$ from eq. (4.29) calculated for the C_ℓ^{TG} shown in Fig. 4.4, with $\ell_{\mathrm{max}} = 500$ and $f_{\mathrm{sky}} = 1$.

To obtain eq. (4.29), we have used Wick's theorem applied to a Gaussian field. In Fig. 4.11 it is shown the covariance from eq. (4.29). The main contribution comes from the largest scale, as they are more affected by the cosmic variance. The variance for the cross-correlation power spectrum is, following eq. (4.29):

$$(\Delta C_\ell^{\mathrm{TG}})^2 = \frac{1}{2\ell+1}\left[C_\ell^{\mathrm{TG}} C_\ell^{\mathrm{TG}} + C_\ell^{\mathrm{TT}} C_\ell^{\mathrm{GG}}\right] . \tag{4.30}$$

Eq. (4.29) must include also the noise terms if present in the maps analyzed. In the tomographic case and thus the correlation between redshift bins, the covariance becomes:

$$\mathrm{Cov}(\hat{C}_\ell^{\mathrm{TG^i}}, \hat{C}_{\ell'}^{\mathrm{TG^j}}) = \frac{\delta^K_{\ell\ell'}}{(2\ell+1)}\left[C_\ell^{\mathrm{TG^i}} C_{\ell'}^{\mathrm{TG^j}} + (C_\ell^{\mathrm{TT}} + N_\ell^{\mathrm{TT}})(C_{\ell'}^{\mathrm{G^iG^j}} + N_{\ell'}^{\mathrm{G^iG^j}})\right] , \tag{4.31}$$

where the indexes i and j run over the redshift bin, N_ℓ^{TT} is the power spectrum of the noise associated with the CMB map (e.g. the instrumental noise of the CMB survey) and N_ℓ^{GG} is the shot noise of the galaxy overdensity map. The shot noise is the error due to the discrete sampling of galaxies and it is different from zero only when $G^i = G^j$. Assuming a Poisson distribution, the shot noise is the associated variance, i.e. the inverse of the mean number density per steradian $\bar{n}$ [68]:

$$N_\ell^{G^iG^j} = \frac{\delta^K_{ij}}{\bar{n}} . \tag{4.32}$$

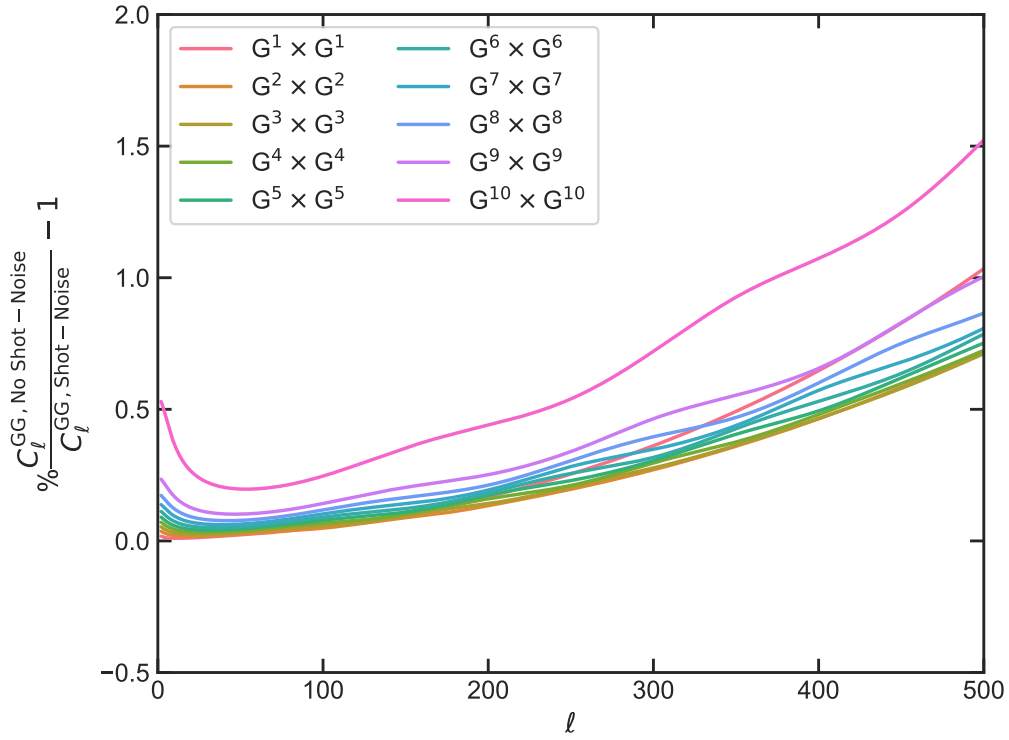


Figure 4.12: Percentage relative difference between the C_ℓ^{GG} calculated without the shot noise term N_ℓ^{GG} and the C_ℓ^{GG} with the shot noise, for different redshift bins (here we show only the C_ℓ^{GG} correlated in the same redshift bins). The cosmological parameters are fixed at the values in table 5.1

For the *Euclid* survey, the mean number density is 30 galaxies/arcmin, and the impact on the C_ℓ^{GG} is shown in Fig. 4.12.

Also, the noise term N_ℓ^{TT} is negligible at the scale of interest for this analysis, because of the high value of the cosmic variance [48]. For this reason, from now on we will drop the term N_ℓ^{TT} from the expression of the covariance. The covariance in eq. (4.31) is shown in Fig. 4.13, left panel. The right panel shows the percentage relative difference between the covariance calculated with and without the shot noise term. The contribution of the shot noise is subdominant with respect to the cosmic variance.
As said previously, in a realistic survey the observed sky is partially covered by a mask, to reduce the impact of the foreground and to account for the observation strategy scanning. An unbiased estimate of the covariance is unfeasible and several methods have been proposed. One approximation is to divide the covariance by a factor f_{sky}^{eff} [69], which is the effective fraction of the sky that is jointly mapped by the CMB and the galaxy survey. Thus,

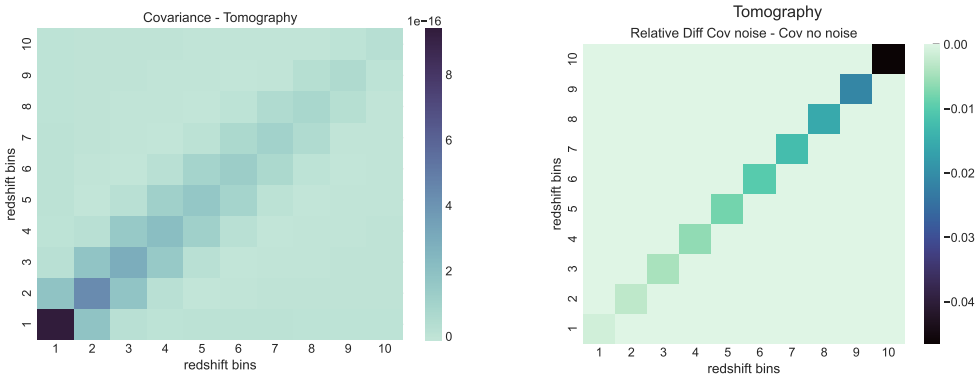


Figure 4.13: Left panel: covariance matrix calculated from eq. (4.31), for each redshift bin. Right panel: relative difference between the covariance calculated with shot noise and the covariance calculated without shot noise.

in the case of partial sky coverage, the covariance becomes:

$$\mathrm{Cov}(\hat{C}_\ell^{\mathrm{TG^i}}, \hat{C}_{\ell'}^{\mathrm{TG^j}}) = \frac{\delta^K_{\ell\ell'}}{f_{sky}^{\mathrm{eff}}(2\ell+1)} \left[C_\ell^{\mathrm{TG^i}} C_{\ell'}^{\mathrm{TG^j}} + C_\ell^{\mathrm{TT}} (C_{\ell'}^{\mathrm{G^iG^j}} + N_{\ell'}^{\mathrm{G^iG^j}}) \right] . \tag{4.33}$$

We can compute eq. (4.33) numerically. For example, given N_{sim} realizations of the CMB and galaxy overdensity maps, one can evaluate the covariance as the scatter of the simulations around the mean $\langle \hat{C}_\ell^{\mathrm{TG}} \rangle_{\mathrm{sim}}$ as:

$$\mathrm{Cov}_{\ell\ell'} = \frac{1}{N_{\mathrm{sim}} - 1} \sum_{i=1}^{N_{\mathrm{sim}}} (\hat{C}_\ell^{\mathrm{TG,i}} - \langle \hat{C}_\ell^{\mathrm{TG}} \rangle_{\mathrm{sim}})(\hat{C}_{\ell'}^{\mathrm{TG,i}} - \langle \hat{C}_{\ell'}^{\mathrm{TG}} \rangle_{\mathrm{sim}}) \tag{4.34}$$

Eq. (4.34) is used to validate the theoretical estimation of the covariance in (4.33). I will discuss this further in the chapter 6.

4.5 Signal-to-Noise analysis

One way to assess the detection significance of the iSW, given the covariance, consists of computing the so-called signal-to-noise ratio. The S/N for each multipoles can be defined as [48]:

$$\begin{aligned} \left(\frac{\mathrm{S}}{\mathrm{N}}\right)^2_\ell &\equiv \frac{d(S/N)^2}{d\ell} = \sum_{\ell'} \sum_{G_i G_j} C_\ell^{TG_i} \left[\mathrm{Cov}^{-1}_{TG_i, TG_j}\right]_{\ell\ell'} C_{\ell'}^{TG_j} \\ &= \sum_{\ell'} \sum_{G_i G_j} f_{sky}^{\mathrm{eff}}(2\ell+1) \frac{C_\ell^{\mathrm{TG^i}} C_{\ell'}^{\mathrm{TG^j}}}{\left[C_\ell^{\mathrm{TG^i}} C_{\ell'}^{\mathrm{TG^j}} + (C_\ell^{\mathrm{TT}} + N_\ell^{\mathrm{TT}})(C_{\ell'}^{\mathrm{G^iG^j}} + N_{\ell'}^{\mathrm{G^iG^j}}) \right]} \end{aligned} \tag{4.35}$$

where we are summing over redshift bins. The corresponding cumulative value over some range of multipoles $[\ell_{\rm min}, \ell_{\rm max}]$ is obtained as

$$\frac{\rm S}{\rm N} = \sqrt{\sum_{\ell=\ell_{\rm min}}^{\ell_{\rm max}} \left(\frac{\rm S}{\rm N}\right)_\ell^2}. \tag{4.36}$$

The expected cumulative S/N at $\ell_{\rm max}$=500 for the *Euclid* photometric survey, for the different sky fractions observed during the mission, is in table 4.2 . In

Data Release	Year	$f_{\rm sky}$	Photometric
DR1	Y1	6%	1.7
DR2	Y2	13%	2.4
	Y3	20%	3.0
DR3	Y4	26%	3.4
	Y5	32%	3.8
	Y6	36%	4.0

Table 4.2: Expected signal-to-noise for the photometric survey, as a function of the observational year, assuming a ΛCDM cosmology with the parameters fixed at the values in table 4.1 and the analysis settings set a the values in table 5.2.

Fig. 4.14, the cumulative S/N is plotted as a function of $\ell_{\rm max}$ (with $\ell_{\rm min} = 2$) for different values of the sky fraction $f_{\rm sky}$. The plot shows that we gain information about the iSW-GC signal up to $\ell_{\rm max} \lesssim 100$. Higher values of the maximum multipole of the analysis do not add more information about the signal. Moreover, as expected, the detected signal is higher for larger $f_{\rm sky}$. We can also analyze the amount of signal carried by each multipole by calculating $(S/N)^2_\ell$, as in eq. (4.35). The results for different $f_{\rm sky}$ are in Fig. 4.15, where it is shown that all the detected signal is accumulated up to $\ell \lesssim 100$, where we expected to find the larger scales affected by the iSW effects. The S/N is higher for larger sky fractions, as expected. The redshift bins where the iSW-GC signal is higher are the lower ones, as shown in Fig. 4.16, where we calculate the cumulative S/N for different values of $\ell_{\rm max}$ and different redshift bins. Furthermore, Fig. 4.17 shows how much S/N we gain when including tomography. The relative difference in the (total) S/N is 6.8%. Finally, Fig. 4.18 shows the impact of the shot noise on the S/N ratio. The relative difference of the total S/N is 0.3%. As aforementioned, the uncertainties are dominated by cosmic variance.

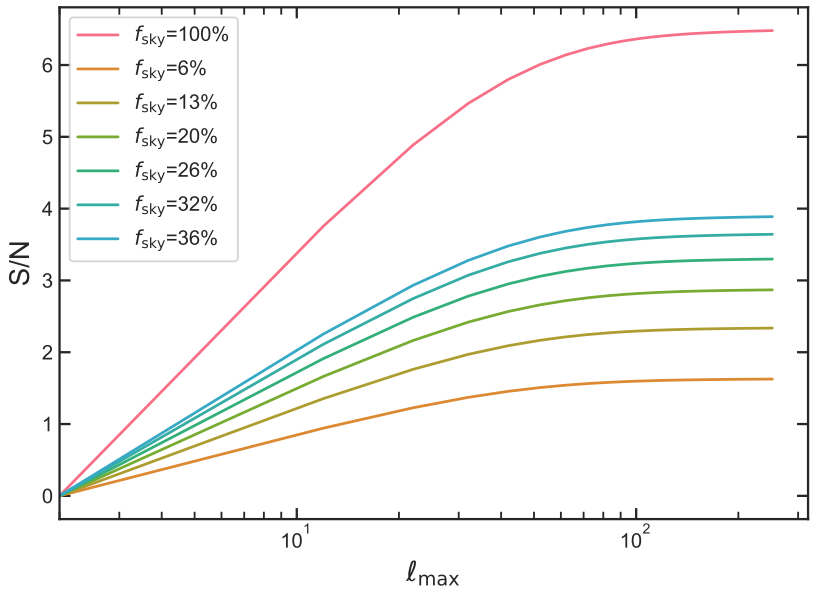


Figure 4.14: Cumulative Signal-to-Noise ratio computed as in eq. (4.36) as a function of ℓ_{max} (ℓ_{min} is fixed at 2). The S/N is evaluated at different sky fractions f_{sky}, keeping fixed all the other analysis settings at the values in the table 5.2

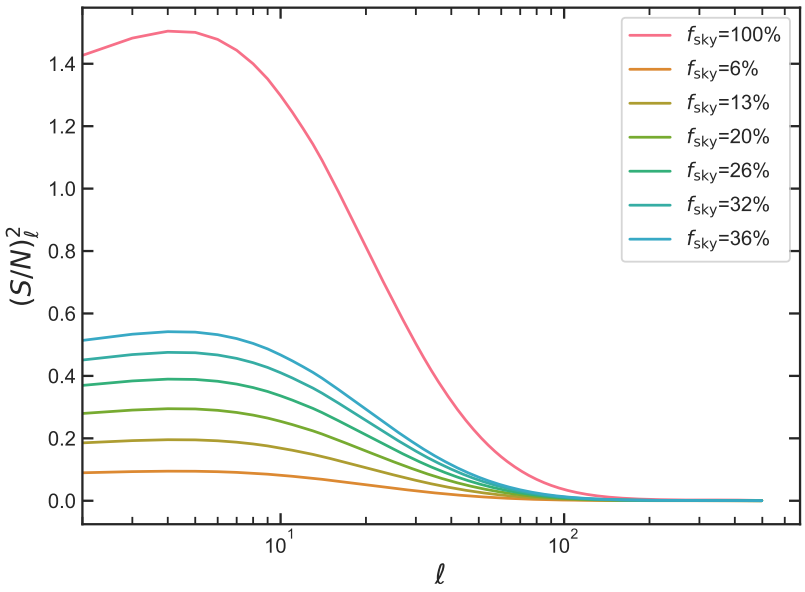


Figure 4.15: Signal-to-Noise ratio computed as in eq. (4.35) as a function of ℓ. The $(S/N)^2_\ell$ is evaluated at different sky fractions f_{sky}, keeping fixed all the other analysis settings at the values in the table 5.2

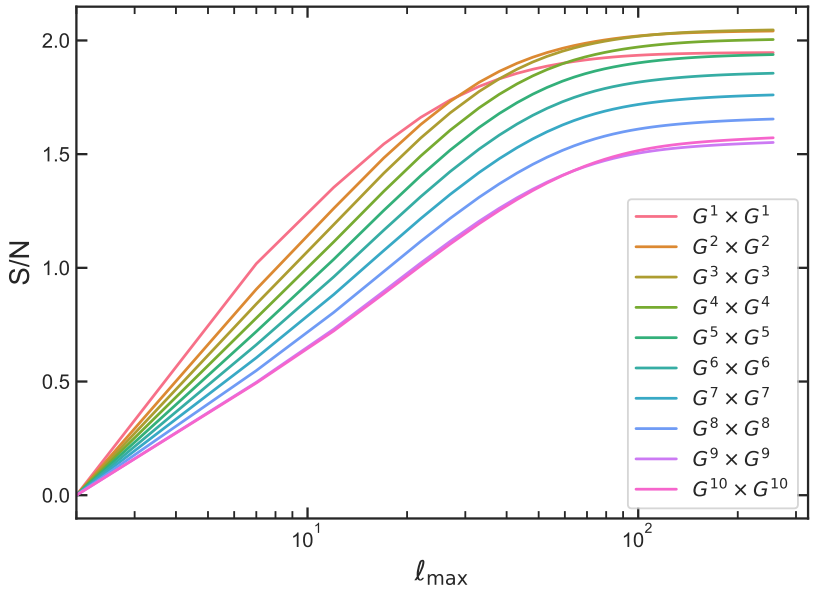


Figure 4.16: Cumulative Signal-to-Noise ratio as a function of ℓ_{max} (ℓ_{min} is fixed at 2). The S/N is evaluated at different redshift bins and here are shown only the auto-correlation terms. The analysis settings are fixed at the values in the table 5.2

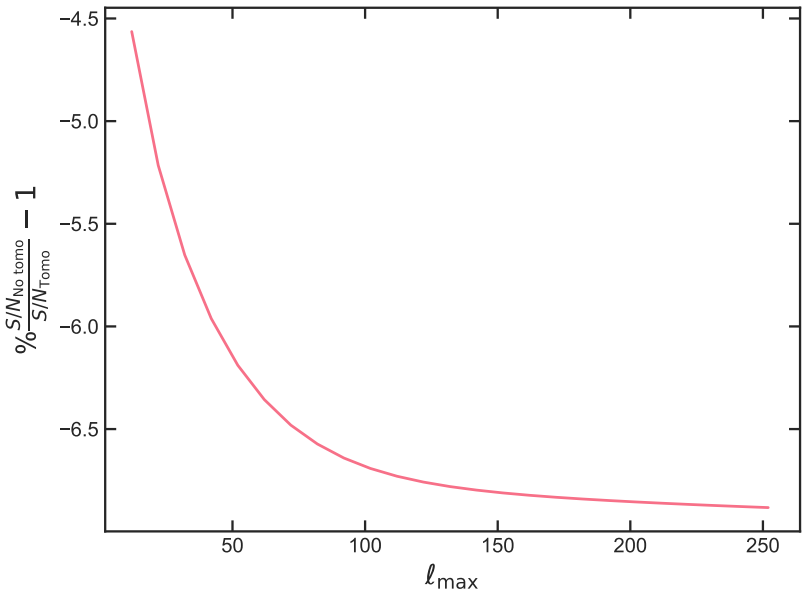


Figure 4.17: Percentage relative difference between the cumulative S/N calculated with tomography and the cumulative S/N ratio calculated without tomography, as a function of ℓ_{max} (ℓ_{min} is fixed at 2), for $f_{\mathrm{sky}} = 0.36$. The analysis settings are fixed at the values in the table 5.2.

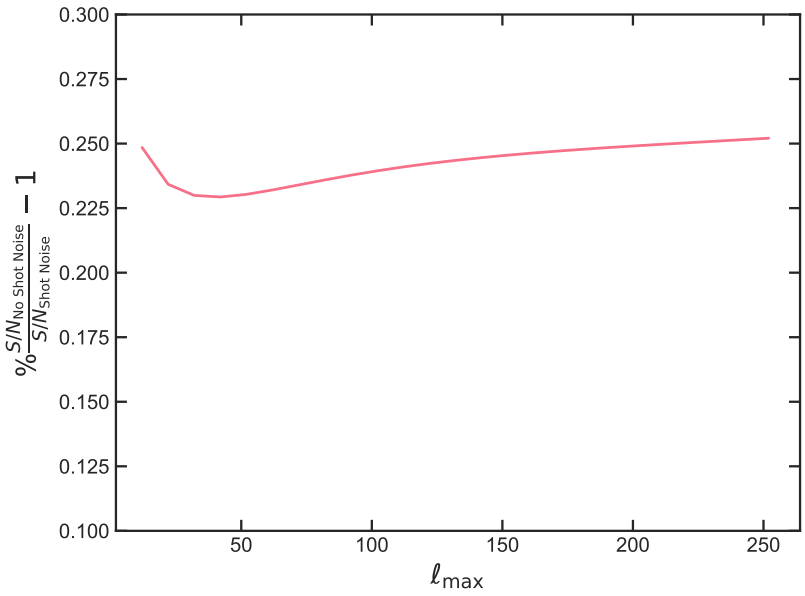


Figure 4.18: Percentage relative difference between the cumulative S/N calculated without shot noise and the cumulative S/N ratio calculated with the shot noise, as a function of ℓ_{max} (ℓ_{min} is fixed at 2), for $f_{\text{sky}} = 0.36$. The analysis settings are fixed at the values in the table 5.2.

Chapter 5

Likelihood validation for the CMBxGC cross-correlation analysis

This chapter focuses on the collaborative efforts of myself and all members of the CMBXC Science Working Group (SWG) within the Euclid collaboration, describing the likelihood function for the cross-correlation analysis between CMB large-scale temperature anisotropies and Galaxy Clustering (GC) . We have rigorously validated our likelihood implementation by inferring the probability distribution for a set of cosmological parameters assuming the ΛCDM model and minimal extension to it, while also assessing any biases introduced by the photometric settings of the Euclid survey. The primary goal of this activity is to integrate the CMBxGC likelihood into the official likelihood code for the Euclid probes, encompassing both the GC and WL components derived from the photometric and spectroscopic survey. This code is known as the Code for Cosmological Likelihood and Observables (CLOE).
This chapter is organized as follows: Section 5.1 provides a brief overview of the concept of likelihood theory and its application in inferring model parameters. In Section 5.2, I describe the likelihood function for the CMBxGC joint analysis and the approximation we have taken into account.
Section 5.3 and Section 5.4 will detail the methodology and discuss the results. It's worth noting that all of this content will be included in the forthcoming paper currently in preparation, titled 'Euclid Preparation: Estimators of the ISW-GC Cross-Correlation'.
Finally, I will conclude by discussing the implementation of the likelihood in CLOE.

5.1 The likelihood function

In the process of conducting an experiment, the ultimate objective is to extract all available information from the data to perform statistical infer-

ence regarding the underlying physical model. Typically, this begins with hypothesis 'H,' which serves as the 'theory' and can pertain to either the parameters of a given model or the model itself. Given 'H' and a set of parameters denoted as $\boldsymbol{\theta}$, which characterize the model, the probability of obtaining the observed data $\mathbf{d}$ is represented by the 'likelihood function', denoted as $\mathcal{L}$:

$$\mathcal{L} = P(\mathbf{d}|H(\boldsymbol{\theta})) = P(\mathbf{d}|\boldsymbol{\theta})\,. \tag{5.1}$$

The likelihood quantifies the probability of the data given $\boldsymbol{\theta}$, but it doesn't represent the probability of the theory itself. In other words, the area subtended the $\mathcal{L}$ as a function of $\boldsymbol{\theta}$ lacks meaning. The interpretation of the likelihood function differs depending on the statistical framework employed, whether frequentist or Bayesian.

In the frequentist approach, probabilities are interpreted as the frequencies of outcomes in an experiment in the limit of infinite repetitions. The probability distribution of certain parameters that describe the underlying physical process giving rise to the data is derived directly from the data itself. Since the data are considered random variables, the parameters $\boldsymbol{\theta}$ are also treated as random variables. The function that estimates the parameters from the data is referred to as an *estimator* of $\boldsymbol{\theta}$, denoted as $\hat{\boldsymbol{\theta}}(\mathbf{d})$. Together with the point estimate of the parameters, statistical inference also furnishes a confidence interval (CI) associated with the estimation. The CI represents an interval centered around $\hat{\boldsymbol{\theta}}$ where, at a certain confidence level, we can expect to find the true value of $\boldsymbol{\theta}$ if the experiment were to be repeated. Typically, the confidence level is expressed in terms of standard deviations from a Gaussian distribution, with '1σ' corresponding to a 68% confidence level, '2σ' to 95%, and so forth. Various methods, such as the 'Neymann' construction, can be employed to estimate the confidence interval from the data.

A commonly chosen estimator $\hat{\boldsymbol{\theta}}(\boldsymbol{d})$ is the *maximum likelihood* (ML) estimator denoted as $\hat{\boldsymbol{\theta}}_{\mathrm{ML}}$. This estimator is defined as the value of $\boldsymbol{\theta}$ that maximizes the likelihood:

$$\hat{\boldsymbol{\theta}}_{\mathrm{ML}}(\boldsymbol{\theta}) \equiv \arg\max_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta})\,. \tag{5.2}$$

In essence, the ML estimator identifies the parameter values that are most likely to produce the given data. It is also minimum variance, meaning that the variance of this estimator is the smallest possible. Moreover, as the number of data points approaches infinity, the ML estimator converges in probability to the true values of the parameters.

In the Bayesian framework, probabilities describe the degrees of belief in a given physical model. Scientists must make assumptions regarding unobserved quantities and then infer the conditional probability, denoted as $p(A|B)$, which represents the probability of event A occurring given that event B has already occurred.

This Bayesian approach is formally described by Bayes' theorem, which is

expressed as follows:

$$p(\boldsymbol{\theta}|\mathbf{d}) = \frac{p(\mathbf{d}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{d})}\,. \tag{5.3}$$

In this equation, the $p(\boldsymbol{\theta}|\mathbf{d})$ is the posterior distribution, which stands for the probability density function of the parameters $\boldsymbol{\theta}$ given the data. The term $p(\mathbf{d}|\boldsymbol{\theta})$ is the likelihood and $p(\boldsymbol{\theta})$ denotes the prior, which is the probability attached to the theory. The last term $p(\mathbf{d})$ is the evidence, and it is essentially a normalization factor for the posterior in the context of parameter estimation.
The ultimate aim of Bayesian statistical inference is to determine the posterior distribution. The form of this posterior distribution can be influenced by the degree of confidence or belief in our understanding of the theoretical model, as expressed through the prior probability. This prior probability can be informed by physical constraints or data from previous experiments. Notably, as more data is collected, the posterior becomes less sensitive to the choice of the prior.
Once the posterior over a set of N parameters $\boldsymbol{\theta}$ is established, it is possible to perform a 'marginalization' process with respect to $N-1$ parameters to recover the posterior for a specific parameter θ_i. This is achieved through the following integral:

$$p(\theta_i|\mathbf{d}) = \int p(\boldsymbol{\theta}|\mathbf{d}) d\boldsymbol{\theta}_{-i}\,. \tag{5.4}$$

In this equation, the index i runs over the total number of parameters N. Subsequently, it is common practice to estimate the value of θ (disregarding the index 'i') either as the expectation value $\langle\theta\rangle$ of the posterior distribution or as the mode θ_M:

$$\langle\theta\rangle = \int \theta p(\theta|\boldsymbol{d}) d\theta \tag{5.5a}$$

$$\theta_M \equiv \arg\max_{\boldsymbol{\theta}} p(\theta|\boldsymbol{d})\,. \tag{5.5b}$$

It is noteworthy that when a flat prior is employed, the estimator θ_M aligns with the maximum likelihood (ML) estimator discussed in equation (5.2). In Bayesian analysis, a point estimate must be accompanied by a credible interval, which is defined as interval I within which there is a probability X that the true value resides. This corresponds to the integral of the posterior within that interval being equal to X.

5.2 The likelihood function for the CMBxGC joint analysis

Before going into the details on the likelihood function that we have implemented for the CMBxGC joint analysis, let's recall the statistical properties of the observed $\hat{C}_\ell$. The reference papers are Refs. [70], [71] and [72].

5.2.1 Statistics of the angular power spectrum

The random variables $a_{\ell m}$ follow a Gaussian distribution described by the equation:

$$p(a_{\ell m}|C_\ell) = \frac{1}{\sqrt{2\pi C_\ell}} \exp\left(-\frac{|a_{\ell m}|^2}{2C_\ell}\right) \tag{5.6}$$

Here, $\langle a_{\ell m}\rangle = 0$ and the variance is C_ℓ. The estimated power spectrum $\hat{C}_\ell$ is obtained by averaging over the $a_{\ell m}$ and follows a Gamma distribution given by:

$$p(\hat{C}_\ell|C_\ell) \propto C_\ell^{-1} \exp\left(-\frac{\nu}{2}\frac{\hat{C}_\ell}{C_\ell}\right). \tag{5.7}$$

In this distribution, $\nu = (2\ell+1)$ represents the degree of freedom. While this expression describes the probability of observing the power spectrum $\hat{C}\ell$ given the fiducial value $C\ell$, it can also be regarded as the likelihood function $\mathcal{L}(C_\ell)$ given the observed data $\hat{C}_\ell$.
As ν approaches infinity, this distribution tends towards a Gaussian distribution with a mean $\langle\hat{C}\ell\rangle = C\ell$ and variance $(\Delta C_\ell)^2 = 2C_\ell/\nu$[1]. This analysis also highlights the recovery of the "cosmic variance" as mentioned in equation (1.70).
For the CMBxGC joint analysis, the observables $\mathbf{a}_{\ell m} = (a^{\rm T}_{\ell m}, a^{\rm G}_{\ell m})$ are drawn a Gaussian multivariate distribution characterized by the covariance matrix:

$$\mathrm{cov}[\mathbf{a}_{\ell m}, \mathbf{a}_{\ell m}] \equiv \mathbf{V}_\ell = \begin{pmatrix} C_\ell^{\rm TT} & C_\ell^{\rm TG} \\ C_\ell^{\rm TG} & C_\ell^{\rm GG} \end{pmatrix}. \tag{5.8}$$

The corresponding estimator is defined as $\mathbf{S}_\ell = \frac{1}{2\ell+1}\sum_m \mathbf{a}_{\ell m}\mathbf{a}^*_{\ell m}$. By performing similar calculations as in the case of single fields, it is determined that the full set of observed spectra $\mathbf{S}_\ell$, given the fiducial spectra, follows a Wishart distribution (which is a generalization of the Gamma distribution in multi-dimensional space) with $\nu = (2\ell+1)$ degrees of freedom and $p = 2$ dimensions. This distribution is given by[2]:

$$p(\mathbf{S}_\ell|\mathbf{V}_\ell/\nu) = \mathcal{L}(\mathbf{V}_\ell/\nu) = \frac{|\mathbf{S}_\ell|^{(\nu-p-1)/2}\exp\left[-\mathrm{Tr}((\mathbf{V}_\ell/\nu)^{-1}\mathbf{S}_\ell/2)\right]}{2^{p\nu}|\mathbf{V}_\ell/\nu|^{\nu/2}\Gamma_p(\nu/2)}. \tag{5.9}$$

Here, $\Gamma_p(\nu/2)$ represents the multivariate Gamma distribution. As ν tends towards infinity, the Wishart distribution converges to a multivariate Gaussian distribution for the variable $\hat{\mathbf{X}}_C = (\hat{C}_\ell^{\rm TT}, \hat{C}_\ell^{\rm TG}, \hat{C}_\ell^{\rm GG})$ with the following

[1] The maximum value of the Gamma distribution does not coincide with the mean because of the *skewness*, which is the asymmetry of the distribution about the mean.

[2] The symbol || indicates the determinant of the matrix.

covariance matrix:

$$\mathrm{cov}[\mathbf{X}_C, \mathbf{X}_C] \equiv \mathbf{C} = \begin{pmatrix} 2(C_\ell^{\mathrm{TT}})^2 & 2(C_\ell^{\mathrm{TT}} C_\ell^{\mathrm{TG}}) & 2(C_\ell^{\mathrm{TG}})^2 \\ 2(C_\ell^{\mathrm{TT}} C_\ell^{\mathrm{TG}}) & C_\ell^{\mathrm{TT}} C_\ell^{\mathrm{TG}} + (C_\ell^{\mathrm{TG}})^2 & 2(C_\ell^{\mathrm{TG}} C_\ell^{\mathrm{GG}}) \\ 2(C_\ell^{\mathrm{TG}})^2 & 2(C_\ell^{\mathrm{TG}} C_\ell^{\mathrm{GG}}) & 2(C_\ell^{\mathrm{GG}})^2 \end{pmatrix} . \tag{5.10}$$

Eq. (5.9) is the exact likelihood for the C_ℓ^{XY}, with XY = {TT, TG, GG}. We can calculate the marginal distribution for each element of $\mathbf{S}_\ell$. In this analysis, we are interested in the cross-correlation term, specifically the marginal distribution for $\hat{C}_\ell^{\mathrm{TG}}$, which is the likelihood function for C_ℓ^{TG}. Therefore we must integrate equation (5.9) over $\hat{C}_\ell^{\mathrm{TT}}$ and $\hat{C}_\ell^{\mathrm{GG}}$. The resulting likelihood is not a Gamma distribution, as seen in equation (5.7) for the single-field analysis. Instead, we find that the solution is a likelihood that depends not only on C_ℓ^{TG} but also on the fiducial C_ℓ^{TT} and C_ℓ^{GG}. The full calculation of the marginal distribution for the cross-correlation power spectrum between CMB temperature and *E*-mode CMB polarization field is provided in reference [70], and the same result applies to the cross-correlation with the galaxy overdensity. The resulting likelihood is as follows:

$$\mathcal{L}(C_\ell^{\mathrm{TG}}) \propto \left[\frac{(\nu \hat{C}_\ell^{\mathrm{TG}})^2}{C_\ell^{\mathrm{TT}} C_\ell^{\mathrm{GG}}}\right]^{(\nu-1)/4} \exp\left(\frac{\nu \hat{C}_\ell^{\mathrm{TG}} C_\ell^{\mathrm{TG}}}{|V_\ell|}\right) K_{(\nu-1)/2}\left(\frac{\nu |\hat{C}_\ell^{\mathrm{TG}}| \sqrt{C_\ell^{\mathrm{TT}} C_\ell^{\mathrm{GG}}}}{|V_\ell|}\right) . \tag{5.11}$$

In this equation, K_n is a modified Bessel function of the second kind.

5.2.2 Gaussian approximation for CMBxGC likelihood

Eq. (5.11) is derived under the assumptions of full-sky observation and no noise. However, these assumptions do not hold in realistic analyses, which must account for noise and sky-cuts. In such cases, the exact likelihood presented in (5.11) is no longer suitable. The introduction of a sky-cut, for example, leads to correlations between different multipoles (as seen in (4.25)), altering the distribution of the estimated spectra compared to (5.9) presented in Ref. [70]. Similar considerations apply, though they are not explicitly addressed in this analysis when anisotropic noise is present. Furthermore, (5.9) is computationally expensive, as it requires inverting complex covariance matrices at various stages of the analysis. An example of such a situation is the computation of the theoretical quantities C_ℓ^{TG}, C_ℓ^{TT}, and C_ℓ^{GG} from simulations for each model under examination, e.g. at each step of a Monte Carlo Markov Chain analysis.
As is often done in CMB experiments, a common approach is to approximate the likelihood for the cross-correlation observable the so-called "fiducial Gaussian" likelihood [72]. This approximation is given by:

$$\chi^2 \equiv -2 \log \mathcal{L}(\mathbf{X}_C) \propto (\hat{\mathbf{X}}_C - \mathbf{X}_C)^T \mathbf{C}_{fid}^{-1} (\hat{\mathbf{X}}_C - \mathbf{X}_C) , \tag{5.12}$$

H_0	$\Omega_b h^2$	$\Omega_c h^2$	$10^9\, A_s$	n_s	τ	$\sum m_\nu$
67	0.0224	0.12	2.1115	0.96	0.058	0.06

Table 5.1: Cosmological parameters for the fiducial cosmology (ΛCDM). This is the *Euclid* forecast prescription, see [44].

$f_{\rm sky}$	$\ell_{\rm min}$	$n_{\rm bins}$	σ_b	c_b	z_b	σ_0	c_0	z_0	$f_{\rm out}$
0.352	2	10	0.05	1.0	0.0	0.05	1.0	0.1	0.1

Table 5.2: Baseline analysis settings for the *Euclid* photometric survey. The parameters σ_b, c_b, z_b, σ_0, c_0, z_0, $f_{\rm out}$ are the parameters for the photo-z distribution in eq. (3.7)

where $\mathbf{C}_{fid}$ is the covariance in eq. (5.10), but it is calculated based on some fixed fiducial model assumed to be smooth and closely aligned with the underlying physical model [22]. For a comprehensive discussion of the likelihood approximation in CMB temperature and polarization analysis, please refer to Ref. [72].
The likelihood of the CMB-CG cross-correlation analysis, which we aim to validate under this approximation, is expressed as follows:

$$-2\log\mathcal{L}(C_\ell^{\rm TG}) = \sum_{\ell\ell'}\sum_{ij}\left(\hat{C}_\ell^{\rm TG^i} - C_\ell^{\rm TG^i}\right)\left[{\rm Cov}(\hat{C}_\ell^{\rm TG^i}, \hat{C}_{\ell'}^{\rm TG^j})\right]^{-1}_{fid}\left(\hat{C}_{\ell'}^{\rm TG^j} - C_{\ell'}^{\rm TG^j}\right), \tag{5.13}$$

where i and j index the redshift bins, and the covariance is calculated as shown in eq. (4.33), all based on a given fiducial model. This expression is the conditional probability of $\hat{C}_\ell^{\rm TG}$ calculated from eq. (5.13), assuming to know $\hat{C}_\ell^{\rm TT}$ and $\hat{C}_\ell^{\rm GG}$.
The assumed fiducial model corresponds to the ΛCDM model, with the cosmological parameters set to the values listed in table 4.1. These parameter values are once again summarized in table 5.1 for easier reference. The baseline settings for the *Euclid* photometric survey are presented in table 3.1, and these settings, along with the fiducial values for the minimum multipole $\ell_{\rm min}$, the number of redshift bins $n_{\rm bins}$, and the fraction of the observed sky $f_{\rm sky}$, are summarized once more in table 5.2.

5.3 Validation over a grid of smooth spectra

We validate the CMBxGC likelihood by employing smooth power spectra, which are the theoretical power spectra computed by a Boltzmann code under the assumption of a fiducial cosmological model. To validate our like-

lihood recipe, we estimate the uncertenties at the level of some cosmological parameters, to identify any potential bias introduced by this approximation. Additionally, we assess and analyze the influence of various analysis settings employed in the *Euclid* photometric survey.
We consider the following cosmological models to explore:

- ΛCDM+A_{iSW}, where we vary the phenomenological motivated parameter $A_{\rm iSW}$, which serves as an overall scaling factor for the expected amplitude of the fiducial $C_\ell^{\rm TG}$ in the likelihood formula. Its expected value is one. The introduction of such a parameter enables consistency checks of our theoretical prescription;
- ΛCDM, where we vary the primordial amplitude of scalar perturbation A_s and the cold dark matter energy density $\Omega_c h^2$;
- w_0CDM, where we vary only the parameter w_0, derived from the Dark Energy equation of state, as specified in eq. (1.21);
- $w_0 w_a$CDM, where we vary the parameters w_0 and w_a obtained from the same equation;

All the other cosmological parameters that are not subject to variation are held constant at their fiducial values as shown in 5.1. For each case, we also explore different values and combinations of the survey analysis settings.

5.3.1 Methodology

I developed a Python code that interfaced with the Boltzmann code CAMB in order to compute the theoretical angular power spectra, which we referred to as "smooth", given a cosmological model. This code allows us to compute the smooth power spectra at different values of one or two cosmological parameters (all the others are kept fixed to the fiducial cosmology) and store them in a "grid". Practically, at each point of a grid of cosmological parameters corresponds a set of smooth spectra ($C_\ell^{\rm TG}$, $C_\ell^{\rm TT}$, $C_\ell^{\rm GG}$), calculated at the values of the parameters
The code allows also varying the analysis settings of the photometric survey: the minimum multipole $\ell_{\rm min}$, the number of redshift bins $n_{\rm bins}$, the fraction of the surveyed sky $f_{\rm sky}$ and σ_b, z_b, $f_{\rm out}$, σ_0, z_0, listed in table 5.2.
Additionally, the code computes the fiducial $C_\ell^{\rm fid}$ at fiducial cosmology.
The output of this code is a "grid" of $C_\ell^{\rm grid}$'s, along with the fiducial spectra $C_\ell^{\rm fid}$'s. For each point of the grid, we compute the likelihood function using eq. (5.13). Here, we set the "estimated" power spectrum, $\hat{C}_\ell^{\rm TG}$, to $C_\ell^{\rm TG,fid}$, and the theoretical spectrum, $C_\ell^{\rm TG}$, to each spectrum $C^{\rm grid}\ell$. The covariance is calculated based on the fiducial $C_\ell^{\rm fid}$.
Now, one of the advantages of the fiducial Gaussian approximation becomes evident: it allows us to compute the covariance matrix once for all

the models under analysis, with the assumption that the underlying model is closely aligned with the fiducial one.
From the likelihood function, we determine the parameter value that maximizes the function (referred to as the "best-fit" value). Once the likelihood is appropriately normalized, we calculate the confidence interval at a 68% confidence level to identify any biases in the distribution. We repeat this procedure while varying several analysis settings. Subsequently, we compare the resulting confidence intervals with those of the fiducial case to assess any bias on parameter constraints.
The analysis settings explored in this analysis are listed in the following:

- $\ell_{\rm min} = [2, 5, 10, 30]$;
- $n_{\rm bins} = [1, 3, 10]$;
- $f_{\rm sky} = [19.7, 30.7, 35.2]$;
- $z_b = [0, 0.05, 0.1]$;
- $\sigma_b = [0.001, 0.05, 0.1]$;
- $\sigma_0 = [0.01, 0.05, 0.1]$;
- $z_0 = [0, 0.1]$;
- $f_{\rm out} = [0, 0.1]$.

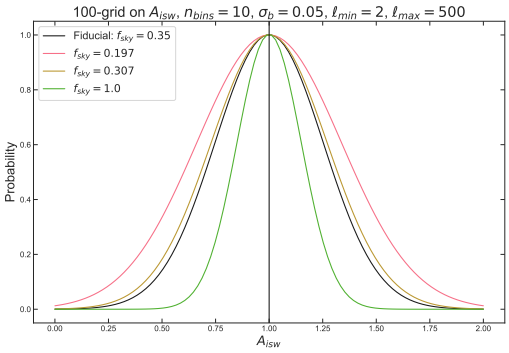
5.3.2 Results

This subsection summarizes the results of the validation of the likelihood approximation over a set of smooth spectra.

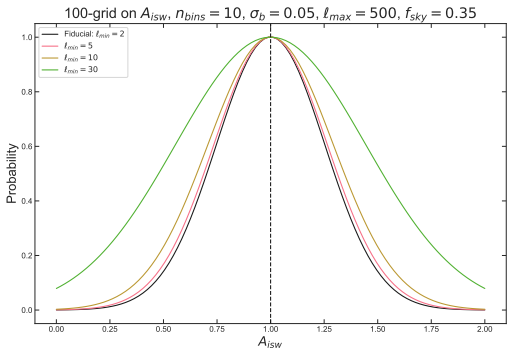
ΛCDM + $\mathbf{A}_{\rm iSW}$ The fiducial value is $A_{\rm iSW,\,fid} = 1$. To compute the distribution for $A_{\rm iSW}$, we maximize the following likelihood:

$$\mathcal{L}(A_{\rm iSW}) \propto \exp\left[-\frac{1}{2}\sum_{\ell}\sum_{ij}\left(C_\ell^{\rm TG^i,\,fid} - A_{\rm iSW}C_\ell^{\rm TG^i,\,fid}\right) \times \left[{\rm Cov}(C_\ell^{\rm TG^i,\,fid}, C_{\ell'}^{\rm TG^j,\,fid})\right]^{-1}\left(C_{\ell'}^{\rm TG^j,\,fid} - A_{\rm iSW}C_{\ell'}^{\rm TG^j,\,fid}\right)\right]. \tag{5.14}$$

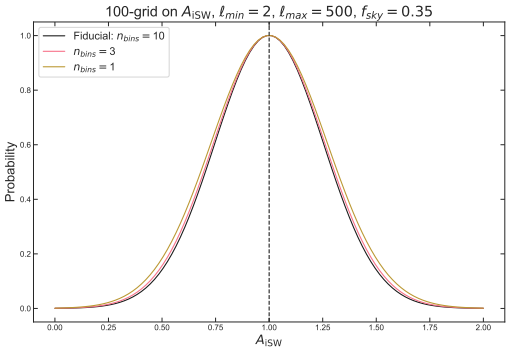
We vary the parameter in the range $A_{\rm iSW} \in [0, 2]$. The results for different analysis settings are displayed in Fig. 5.1. The plots show a symmetrical distribution centered in the fiducial value, as expected. The width of these distributions is notably influenced by the minimum multipole value, $\ell_{\rm min}$, in the analysis. As seen in Fig. 5.1b, a decrease in the $\ell_{\rm min}$ value results in increased constraining power, effectively narrowing the distribution. This behavior can be attributed to the fact that the iSW signal is predominantly concentrated at very large scales, and excluding multipoles $\ell \lesssim 30$ reduces



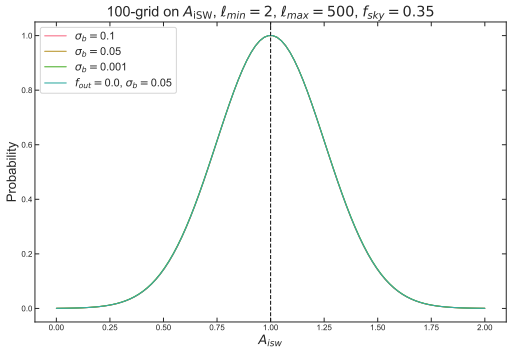
(a) The probability distribution for $A_{\rm iSW}$ varying $f_{\rm sky}$.



(b) The probability distribution for $A_{\rm iSW}$ varying $\ell_{\rm min}$.



(c) The probability distribution for $A_{\rm iSW}$ varying $n_{\rm bins}$.



(d) The probability distribution for $A_{\rm iSW}$ varying $\sigma_{\rm b}$ and $f_{\rm out}$.

Figure 5.1: The probability distribution for the parameter $A_{\rm iSW}$, varying the analysis settings. The fiducial value is indicated by the black line. All the parameters of the analysis (table 5.2), not displayed in the plots are kept constant. The cosmological parameters are fixed according to the values listed in table 5.2. The underlined fiducial model is ΛCDM + $A_{\rm iSW}$.

the constraining power.
Furthermore, the amplitude $A_{\rm iSW}$ is significantly affected by the fraction of the sky observed during the survey. Fig. 5.1a illustrates the probability distribution for different values of $f_{\rm sky}$, representing the fraction of the sky observed by the *Euclid* mission. In the ideal scenario of full sky observation with $f_{\rm sky} = 1$, the iSW signal exhibits enhanced constraining power since all scales are accessible. However, as $f_{\rm sky}$ decreases, the covariance decreases proportionally ($\propto f_{\rm sky}$), resulting in higher uncertainties due to the loss of information from larger scales.
Conversely, the tomographic analysis offers a slight improvement in constraining the amplitude, as evidenced in Fig. 5.1c when increasing the number of redshift bins in the analysis.
Moving on to the last plot, Fig. 5.1d, it is evident that varying the photometric error σ_b has no discernible effect on the distribution, nor does the fraction

$\Lambda\text{CDM} + A_{\text{iSW}}$	
	$\%\left(\text{CI}_{f_{\text{out}}=0}/\text{CI}_{f_{\text{out}}=0.1}\right)_{A_{\text{iSW}}} - 1$
$\sigma_b = 0.001$	0%
$\sigma_b = 0.05$	0.08%
$\sigma_b = 0.10$	0.08%

Table 5.3: The percentage relative difference of the CI at 68% CL of the A_{iSW} probability distribution calculated with $f_{\text{out}} = 0$ and $f_{\text{out}} = 0.1$ (our baseline), varying σ_b. All the other analysis settings are kept fixed.

$\Lambda\text{CDM} + A_{\text{iSW}}$	
	$\%\left(\text{CI}_{\text{No noise}}/\text{CI}_{\text{Noise}}\right)_{A_{\text{iSW}}} - 1$
$n_{\text{bins}} = 1$	0.07%
$n_{\text{bins}} = 3$	0.15%
$n_{\text{bins}} = 10$	0.24%

Table 5.4: The percentage relative difference of the CI at 68% CL of the A_{iSW} probability distribution with and without the shot noise, for different values of n_{bins}.

of catastrophic outliers f_{out}. Further exploring this aspect, Fig. 5.3 shows the impact of modifying f_{out}. This is achieved by plotting the percentage relative difference between the CI at 68% CL of the A_{iSW} probability distribution A_{iSW} with $f_{\text{out}} = 0$ and $f_{\text{out}} = 0.1$ (our baseline) for various values of σ_b,to spot any potential degeneracies between the two parameters. The negligible difference between the two cases suggests that a 10% probability of catastrophic outliers in the photo-z distribution does not affect the results. Next, we assess the effect of shot noise on the uncertainties of A_{iSW}. Fig. 5.4 shows the percentage relative difference of the CI with and without shot noise for various values of the number of redshift bins. As previously discussed (see the discussion about Fig. 4.13), shot noise does not impact the distribution of A_{iSW} since uncertainties are primarily dominated by cosmic variance. It is noteworthy that the difference is more pronounced in the tomographic analysis with $n_{\text{bins}} = 10$. This behavior can be attributed to the fact that each bin receives an equal contribution from shot noise, as every bin is populated by an identical number of galaxies. This can amplify the impact of shot noise. To conclude the discussion on A_{iSW}, we calculate the 68% confidence intervals (CI) for the different analysis cases and compare them to the baseline. This is done by computing the ratio between the CIs, and the results are summarized in table 5.5. The results in the table reaffirm

$\Lambda\text{CDM} + A_{\text{iSW}}$	
	$(\text{CI}/\text{CI}_{\text{fid}})_{A_{\text{iSW}}}$
fiducial	1
$\ell_{\text{min}} = 5$	1.047
$\ell_{\text{min}} = 10$	1.156
$\ell_{\text{min}} = 30$	1.697
$n_{\text{bins}} = 3$	1.027
$n_{\text{bins}} = 1$	1.074
$f_{\text{sky}} = 0.197$	1.331
$f_{\text{sky}} = 0.307$	1.071
$f_{\text{sky}} = 1.000$	0.594
$\sigma_b = 0.001$	1.004
$\sigma_b = 0.10$	1.000
$f_{\text{out}} = 0.0$	1.001

Table 5.5: Here are reported the 68% confidence intervals (CI) on the distribution of A_{iSW} with respect to different analysis settings of the *Euclid* photometric survey. The fiducial values refer to our baseline in table 5.2, all the other analysis settings are taken accordingly unless otherwise stated in the first column. All the values are normalized by the 68% confidence interval of the fiducial baseline. The underlined fiducial model is $\Lambda\text{CDM} + A_{\text{iSW}}$.

the previous discussions: A_{iSW} remains insensitive to variations in σ_b and f_{out}, while it is influenced by changes in f_{sky} and ℓ_{min}. Specifically, the CI decreases as f_{sky} and n_{bins} increase. The CI decreases as ℓ_{min} decreases, as previously noted.

ΛCDM Within the framework of the ΛCDM model, we conducted a likelihood test over a grid consisting of 20x20 points for A_s and $\Omega_c h^2$. We varied these parameters within the ranges $A_S \in \left[1.65 \times 10^{-10}, 5.46 \times 10^{-9}\right]$ and $\Omega_c h^2 \in [0, 0.4]$, investigating the degeneracy between these cosmological parameters while holding the others constant at their fiducial values. The joint probability distribution in the parameter space is shown in Fig. 5.2. We repeat the same analysis as for the A_{iSW} parameter, exploring various combinations and values for the analysis settings. The results are presented in Fig. 5.3. The CMBxGC signal effectively constrains the marginal distributions of $\Omega_c h^2$ and A_s, with their maximum values coinciding with the fiducial parameters $\Omega_c h^2 = 0.12$ and $A_S = 2.1115 \times 10^{-9}$. The tightest constraints on the parameters are achieved when analyzing with $n_{\text{bins}} = 10$ and

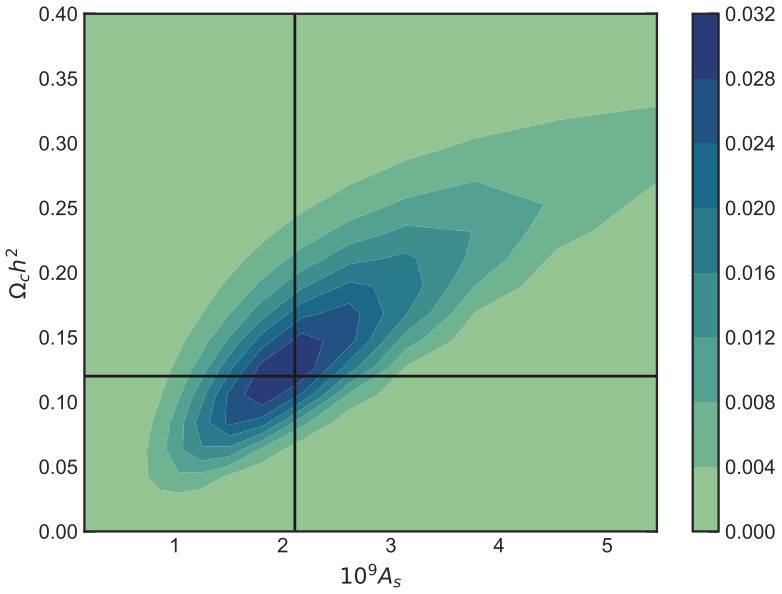
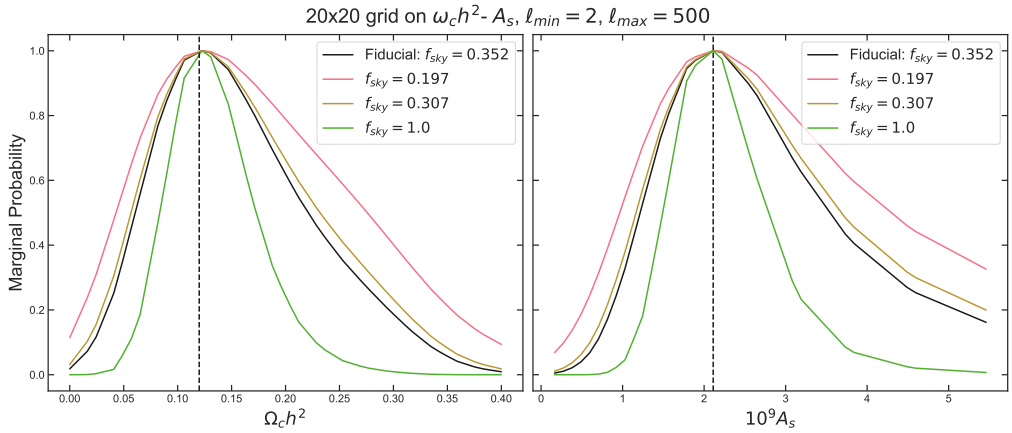


Figure 5.2: The density distribution of the joint probability function in the $A_s - \Omega_c h^2$ parameter space. The black lines represent the fiducial values (see table 5.1). The probability distribution peaks at the fiducial values of A_s and $\Omega_c h^2$. All other cosmological parameters remain fixed, in accordance with the values provided in table 5.1. The analysis settings are also held constant at the values listed in table 5.2. The underlined fiducial model is ΛCDM.

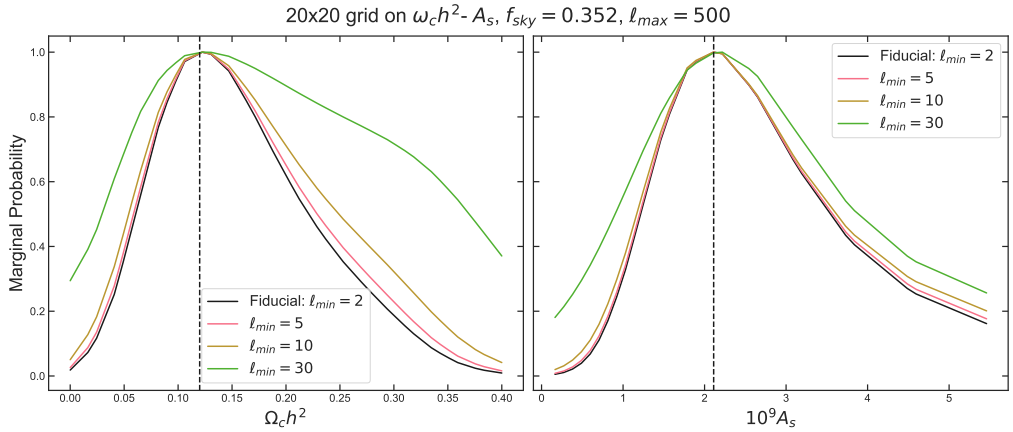
$\ell_{\rm min} = 2$. As expected, full-sky observation provides greater constraining power since it allows access to all relevant scales for the iSW signal. Similar to the previous analysis, variations in σ_b for different values of $f_{\rm out}$ do not impact the constraints on A_s and $\Omega_c h^2$, see Table 5.6. Shot noise affects the

	ΛCDM	
	$\%\left(\mathrm{CI}_{f_{\rm out}=0}/\mathrm{CI}_{f_{\rm out}=0.1}\right)_{A_s} - 1$	$\%\left(\mathrm{CI}_{f_{\rm out}=0}/\mathrm{CI}_{f_{\rm out}=0.1}\right)_{\Omega_c h^2} - 1$
$\sigma_b = 0.001$	0.99%	2.34%
$\sigma_b = 0.05$	0.97%	1.55%
$\sigma_b = 0.10$	0.14%	0.0.8%

Table 5.6: The percentage relative difference on the CI at 68% CL of the $\Omega_c h^2$ and A_s marginal probability distributions calculated with $f_{\rm out} = 0$ and $f_{\rm out} = 0.1$ (our baseline), varying σ_b. All the other analysis settings are kept fixed at table 5.2. The cosmological model is ΛCDM with the parameters fixed at table 5.1.



[h] (a) The probability distributions for $\Omega_c h^2$ and A_s varying $f_{\rm sky}$.



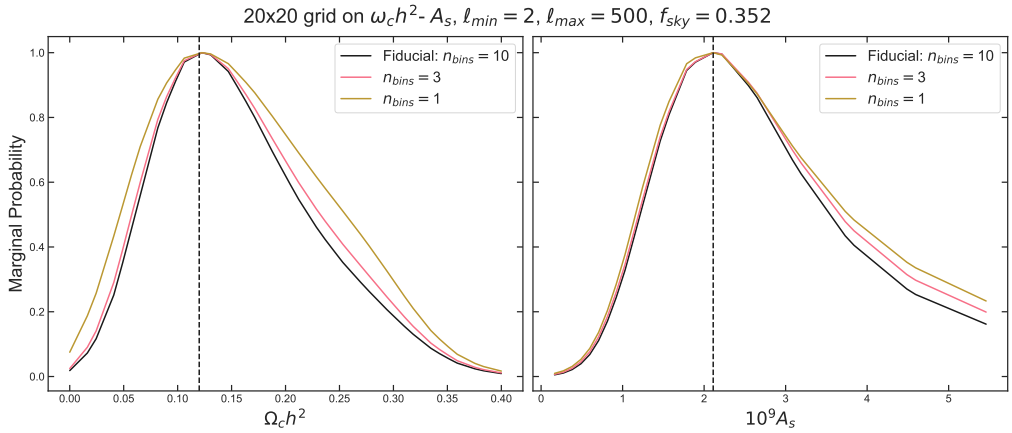
(b) The probability distributions for $\Omega_c h^2$ and A_s varying $\ell_{\rm min}$.

Figure 5.3: The marginal probability distribution for the parameters $\Omega_c h^2$ and A_s, varying the analysis settings. The fiducial case is represented by the black line. All other analysis parameters (not depicted in the plots) are maintained at their specified values in table 5.2. The cosmological parameters remain fixed, adhering to the values provided in table 5.2. The underlined fiducial model is ΛCDM.

CI of marginal distributions less than 1%, as shown in Table 5.7, indicating that uncertainties are dominated by cosmic variance. Moreover, as for the $A_{\rm iSW}$, differences are more pronounced in the tomographic case.

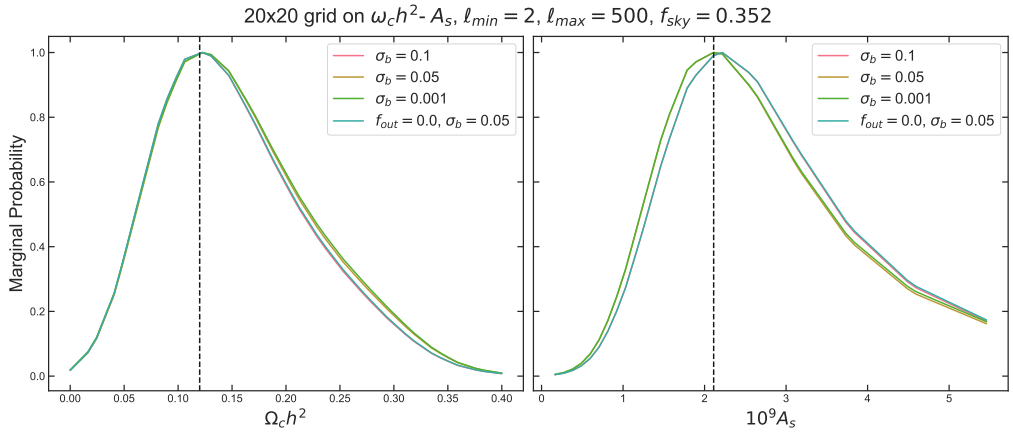
Table 5.8 presents the CI at the 68% CL for the marginal distribution of A_s and $\Omega_c h^2$, divided by the CI of the fiducial baseline, summarizing the results discussed above.

w_0CDM In the framework of the ΛCDM + w_0 model, we infer the probability distribution for w_0 by constructing a grid comprising 50 points spanning the range $w_0 \in [-1.5, -0.5]$. We fixed all other cosmological parameters



[h]

(c) The probability distributions for $\Omega_c h^2$ and A_s varying $n_{\rm bins}$.



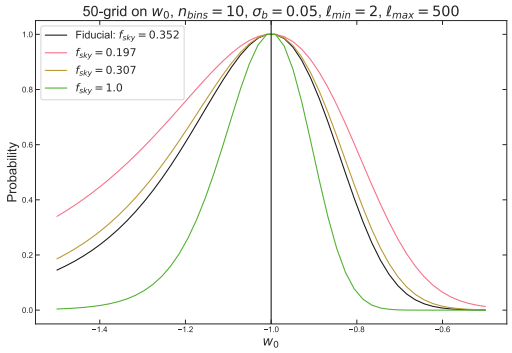
(d) The probability distributions for $\Omega_c h^2$ and A_s varying $\sigma_{\rm b}$ and $f_{\rm out}$.

Figure 5.3: The marginal probability distribution for the parameters $\Omega_c h^2$ and A_s, varying the analysis settings. The fiducial case is represented by the black line. All other analysis parameters (not depicted in the plots) are maintained at their specified values in table 5.2. The cosmological parameters remain fixed, adhering to the values provided in table 5.2. The underlined fiducial model is ΛCDM.

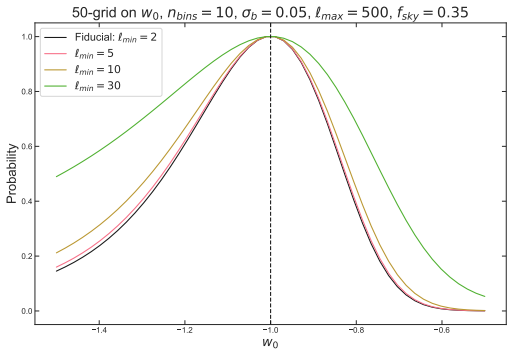
at their fiducial values as outlined in table 5.1. Again, we evaluate this distribution while varying the analysis settings, as in the previous analysis. The considered parameters encompass $f_{\rm sky}$, $\ell_{\rm min}$, $f_{\rm out}$, σ_b, and $n_{\rm bins}$.
In Fig. 5.4 are shown the main results. These plots show analogous trends to those observed in the case of the $A_{\rm iSW}$ parameter. The fiducial value $w_0 = -1$ is recovered as the best-fit value of the distribution. The distribution is mainly affected by $f_{\rm sky}$ (Fig. 5.4a) and $\ell_{\rm min}$ (Fig. 5.4b): a larger fraction of the sky corresponds to the tighter constraints on the parameter, whereas a larger $\ell_{\rm min}$ leads to a wider distribution. As for $A_{\rm iSW}$, $n_{\rm bins} = 10$, which is our baseline, improves the constraints on the parameter (Fig. 5.4c), surpassing the performance of both the non-tomographic and 3-redshift bin

	ΛCDM	
	$\%\,(\mathrm{CI}_{\text{No noise}}/\mathrm{CI}_{\text{Noise}})_{A_s} - 1$	$\%\,(\mathrm{CI}_{\text{No noise}}/\mathrm{CI}_{\text{Noise}})_{\Omega_c h^2} - 1$
$n_{\text{bins}} = 1$	0.16%	0.57%
$n_{\text{bins}} = 3$	0.24%	0.96%
$n_{\text{bins}} = 10$	0.43%	0.44%

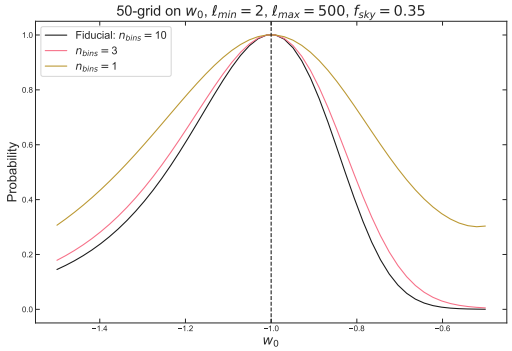
Table 5.7: The percentage relative difference of the CI at 68% CL of the marginal probability distributions of $\Omega_c h^2$ and A_s with and without the shot noise, for different values of n_{bins}. All the other analysis settings are kept fixed at table 5.2. The cosmological model is ΛCDM with the parameters fixed at table 5.1.



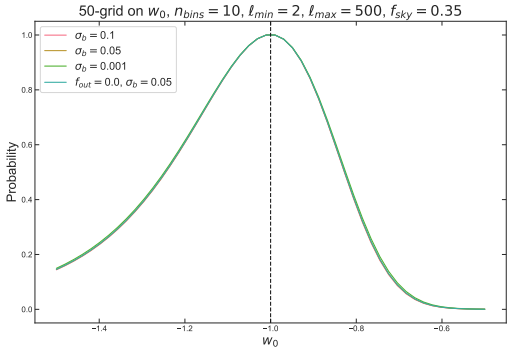
(a) The probability distribution for w_0 varying f_{sky}.



(b) The probability distribution for w_0 varying ℓ_{min}.



(c) The probability distribution for w_0 varying n_{bins}.



(d) The probability distribution for w_0 varying σ_{b} and f_{out}.

Figure 5.4: The probability distribution for the parameter w_0, varying the analysis settings. The fiducial value is indicated with the black line. All the parameters of the analysis (table 5.2), not displayed in the plots are kept constant. The cosmological parameters are fixed according to the values listed in table 5.2. The underlined fiducial model is w_0CDM.

ΛCDM		
	$(\mathrm{CI}/\mathrm{CI}_{\mathrm{fid}})_{A_s}$	$(\mathrm{CI}/\mathrm{CI}_{\mathrm{fid}})_{\Omega_c h^2}$
fiducial	1	1
$\ell_{\mathrm{min}} = 5$	1.033	1.064
$\ell_{\mathrm{min}} = 10$	1.062	1.184
$\ell_{\mathrm{min}} = 30$	1.216	1.628
$n_{\mathrm{bins}} = 3$	1.050	1.061
$n_{\mathrm{bins}} = 1$	1.093	1.176
$f_{\mathrm{sky}} = 0.197$	1.53	1.309
$f_{\mathrm{sky}} = 0.307$	1.243	1.080
$f_{\mathrm{sky}} = 1.000$	1.064	0.566
$\sigma_b = 0.001$	1.008	1.017
$\sigma_b = 0.10$	1.006	0.980
$f_{\mathrm{out}} = 0.0$	1.023	0.984

Table 5.8: Here are reported the 68% confidence intervals (CI) on the distributions of A_s and $\Omega_c h^2$ with respect to different analysis settings of the *Euclid* photometric survey. The fiducial values refer to our baseline in table 5.2, all the other analysis settings are taken accordingly unless otherwise stated in the first column. All the values are normalized by the 68% confidence interval of the fiducial baseline. The underlined fiducial model is ΛCDM.

analyses.
Furthermore, Fig. 5.4d demonstrates that changing the photometric error value, σ_b exerts no discernible effect on the distribution. The same holds true for f_{out}: variations in this parameter scarcely affect the distribution. To provide a clearer perspective on this effect, Table 5.9 shows the percentage relative difference between the distributions computed with $f_{\mathrm{out}} = 0.1$ (our baseline) and $f_{\mathrm{out}} = 0$ for varying σ_b values.

We check the impact of shot noise on the constraints for w_0 (refer to Table 5.10). As in the previous cases, the shot noise contributes negligibly compared to cosmic variance. Additionally, the discrepancies are more pronounced in the tomographic analysis.

Table 5.11 reports the 68% confidence intervals (CI) of the w_0 distributions, normalized to the CI of the fiducial baseline. The analysis settings are adjusted according to the initial column, while all other parameters, including the cosmological ones, remain fixed at their fiducial values. The results reinforce the previous results: constraints on w_0 are influenced by the frac-

w_0CDM	
	$\%\left(\mathrm{CI}_{f_{\mathrm{out}}=0}/\mathrm{CI}_{f_{\mathrm{out}}=0.1}\right)_{w_0}-1$
$\sigma_b = 0.001$	0.18%
$\sigma_b = 0.05$	0.13%
$\sigma_b = 0.10$	0.27%

Table 5.9: The percentage relative difference on the CI at 68% CL of the w_0 probability distribution calculated with $f_{\mathrm{out}} = 0$ and $f_{\mathrm{out}} = 0.1$ (our baseline), varying σ_b. All the other analysis settings are kept fixed at table 5.2. The cosmological model is w_0CDM with the parameters fixed at table 5.1.

w_0CDM	
	$\%\left(\mathrm{CI}_{\mathrm{No\ noise}}/\mathrm{CI}_{\mathrm{Noise}}\right)_{w_0}-1$
$n_{\mathrm{bins}} = 1$	0.05%
$n_{\mathrm{bins}} = 3$	0.21%
$n_{\mathrm{bins}} = 10$	0.47%

Table 5.10: The percentage relative difference on the CI at 68% CL of the w_0 probability distribution with and without the shot noise, for different values of n_{bins}. The cosmological model is w_0CDM.

tion of the sky observed and the minimum scale analyzed while showing minimal sensitivity to the fraction of outliers in the photo-z distribution and the photometric error.

w_0w_aCDM We conduct an analysis involving w_a and w_0 by generating a 20x20 grid spanning the ranges $w_0 \in [-1.5, -0.5]$ and $w_a \in [-0.2, 0.2]$. As in previous cases, we maintain the other parameters at their fiducial values, and the cosmological model remains ΛCDM. We explore various combinations and values for the analysis settings, extracting the probability distribution of each parameter after marginalizing with respect to the other one. The primary findings are summarized in Fig. 5.5. Clearly, the CMBxGC cross-correlation signal is unable to constrain w_a, as can be seen also in Fig. 5.6, which depicts the joint probability distribution of w_0 and w_a. However, it successfully recovers the marginal probability distribution for w_0, with its peak at $w_0 = -1$. Similar to the previous case, the tightest constraints on w_0 are achieved with $n_{\mathrm{bins}} = 10$, $\ell_{\mathrm{min}} = 2$, and $f_{\mathrm{sky}} = 1$. Changing f_{out} has negligible impact on the distribution of w_0, as Table 5.12 shows for different values of σ_b. The effect of shot noise is negligible

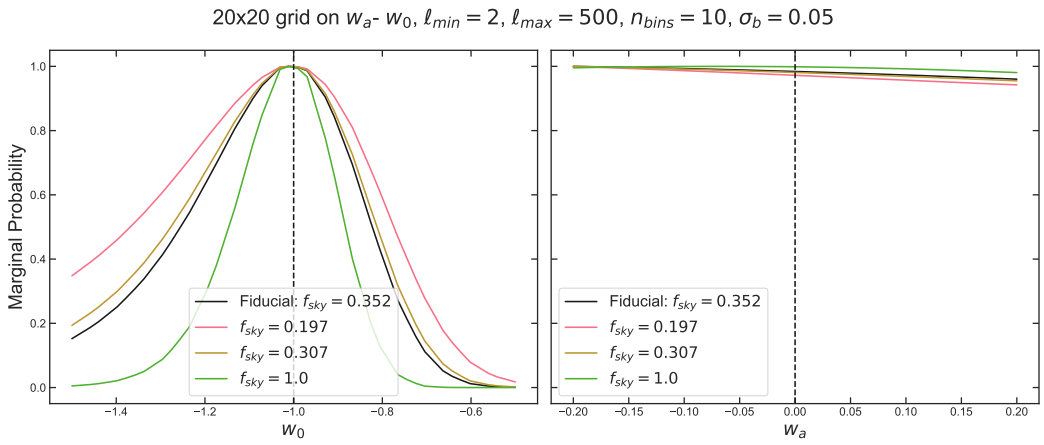
w_0CDM	
	$(\mathrm{CI}/\mathrm{CI}_{\mathrm{fid}})_{w_0}$
fiducial	1
$\ell_{\mathrm{min}} = 5$	1.023
$\ell_{\mathrm{min}} = 10$	1.091
$\ell_{\mathrm{min}} = 30$	1.392
$n_{\mathrm{bins}} = 3$	1.083
$n_{\mathrm{bins}} = 1$	1.442
$f_{\mathrm{sky}} = 0.197$	1.245
$f_{\mathrm{sky}} = 0.307$	1.060
$f_{\mathrm{sky}} = 1.000$	0.585
$\sigma_b = 0.001$	1.011
$\sigma_b = 0.10$	0.997
$f_{\mathrm{out}} = 0.0$	1.001

Table 5.11: Here are reported the 68% confidence intervals (CI) on the distribution of w_0 with respect to different analysis settings of the *Euclid* photometric survey. The fiducial values refer to our baseline in table 5.2, all the other analysis settings are taken accordingly unless otherwise stated in the first column. All the values are normalized by the 68% confidence interval of the fiducial baseline. The underlined fiducial model is w_0CDM.

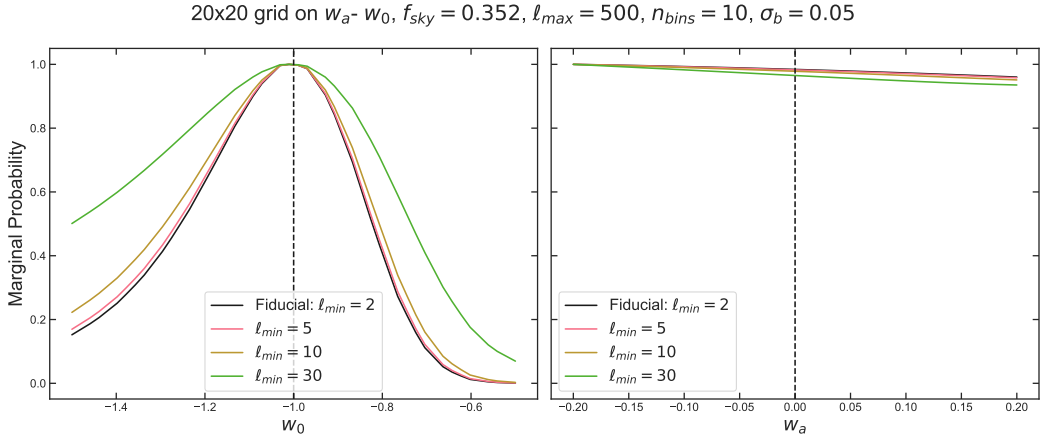
as in the previous analysis, as shown in Tab. 5.13, indicating that cosmic variance is predominantly in the uncertainties on w_0. In conclusion, Table 5.14 presents the 68% confidence intervals (CI) for the marginal distribution of w_0, normalized by the CI found in the baseline case, summarizing the results discussed above.

5.4 Validation on 'wrong' fiducials

We assess the errors in the estimation of the cosmological parameters induced by a "wrong" fiducial cosmology assumed when computing the covariance matrix in eq. (5.13). This test involves calculating the covariance matrix from power spectra computed at the "wrong" fiducial model, meaning spectra that differ from the fiducial spectra we are assuming. Practically, we insert in eq. (5.13) the power spectra C_ℓ^{TG} computed with the fiducial cosmology and the covariance computed at the "wrong" fiducial, i.e. $\left[\mathrm{Cov}(C_\ell^{\mathrm{TG^i}}, C_{\ell'}^{\mathrm{TG^j}})\right]_{\mathrm{wrong\,fid}}$. We then assess the bias in the likelihood compared to the fiducial case.



(a) The probability distributions for w_0 and w_a varying $f_{\rm sky}$.



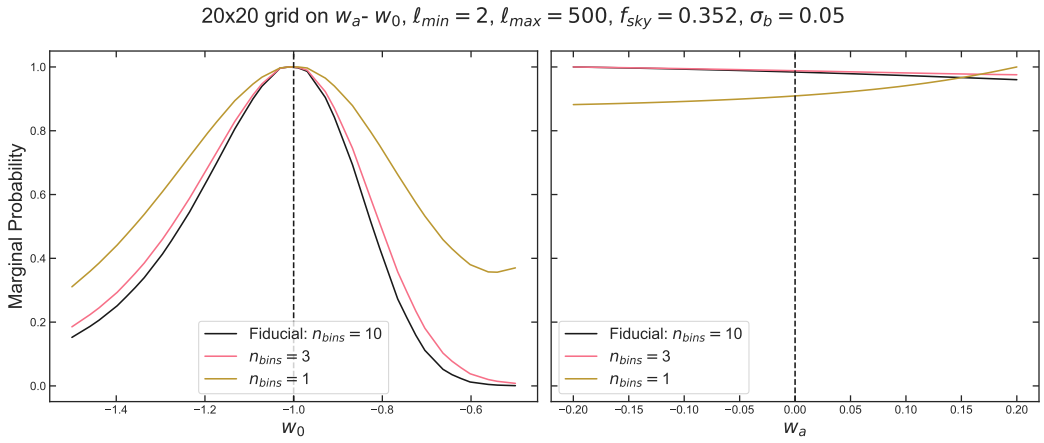
(b) The probability distributions for w_0 and w_a varying $\ell_{\rm min}$.

Figure 5.5: The marginal probability distribution for the parameters w_0 and w_a, varying the analysis settings.The fiducial case is represented by the black line. All other analysis parameters (not depicted in the plots) are maintained at their specified values in table 5.2. The cosmological parameters remain fixed, adhering to the values provided in table 5.2. The underlined fiducial model is w_0w_aCDM.

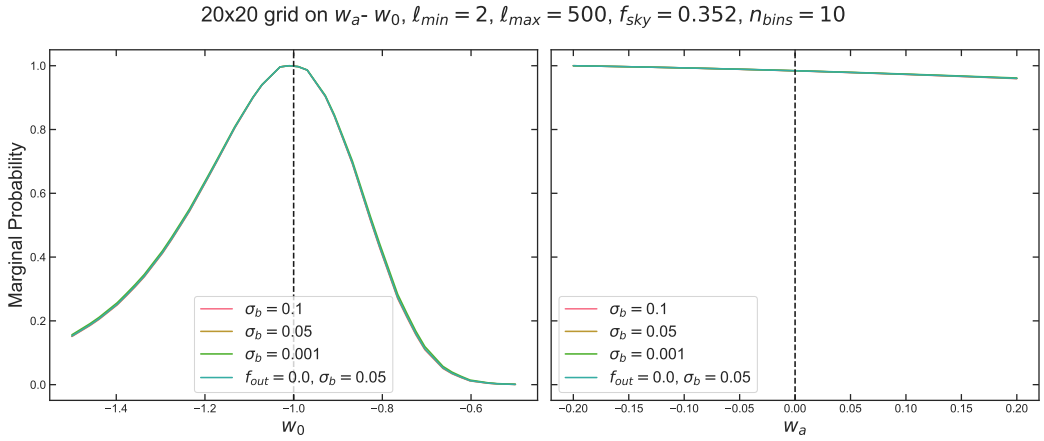
The wrong fiducial is chosen to differ from the "correct" one by adding a bias to the value of one cosmological parameter. We test two cosmological models: ΛCDM with the "wrong" fiducial value for $\Omega_c h^2$, and w_0CDM with the "wrong" fiducial value for w_0.
We choose as bias the 1σ uncertainties for $\Omega_c h^2$ and w_0 as measured by Planck 2018 (refer to the table in `https://wiki.cosmos.esa.int/planck-legacy-archive/images/b/be/Baseline_params_table_2018_68pc.pdf`):

$$\sigma^{\rm Planck}_{\Omega_c h^2} = +0.0010 \tag{5.15}$$

$$\sigma^{\rm Planck}_{w_0} = +0.0035\,. \tag{5.16}$$



(c) The probability distributions for w_0 and w_a varying n_{bins}.



(d) The probability distributions for w_0 and w_a varying σ_{b} and f_{out}.

Figure 5.5: The marginal probability distribution for the parameters w_0 and w_a, varying the analysis settings.The fiducial case is represented by the black line. All other analysis parameters (not depicted in the plots) are maintained at their specified values in table 5.2. The cosmological parameters remain fixed, adhering to the values provided in table 5.2. The underlined fiducial model is $w_0 w_a$CDM.

These values are derived from CMB temperature and polarization data along with BAO measurements. As the distribution for w_0 is asymmetric, here we take the mean value of the left and righ uncertainties from the table. To obtain the "wrong" values for these parameters, we add 3 times the 1σ uncertainty from eq. (5.15) to our fiducial values (as a conservative approach):

$$\Omega_c h^2_{\text{wrong}} = \Omega_c h^2_{\text{fid}} + 3 \times \sigma^{\text{Planck}}_{\Omega_c h^2} \tag{5.17}$$

$$w_{0\,\text{wrong}} = w_{0\,\text{fid}} + 3 \times \sigma^{\text{Planck}}_{w_0}. \tag{5.18}$$

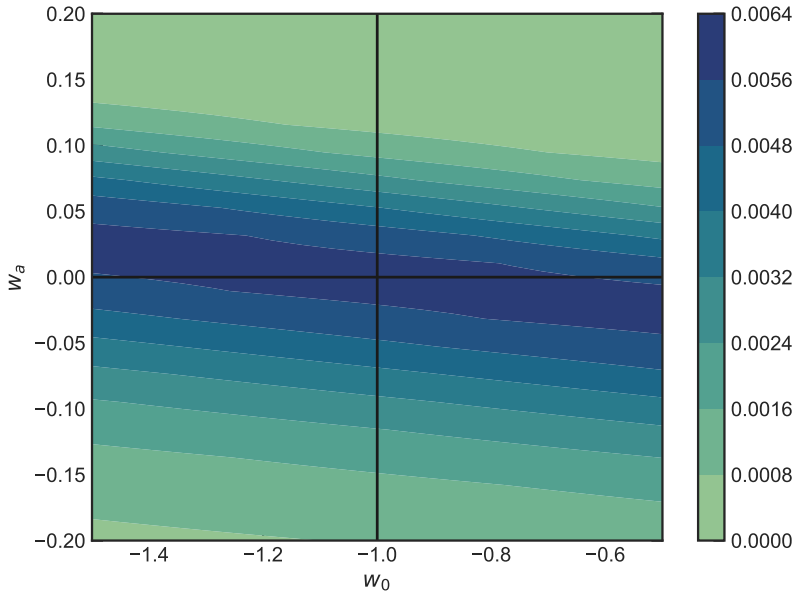


Figure 5.6: The density distribution of the joint probability function in the $w_o - w_a$ parameter space. The black lines represent the fiducial values (see table 5.1). All other cosmological parameters remain fixed, in accordance with the values provided in table 5.1. The analysis settings are also held constant at the values listed in table 5.2. The w_a parameter is not constrained by the CMBxGC signal. The cosmological model is $w_0 w_a$CDM.

$w_0 w_a$CDM	
	$\%\left(\mathrm{CI}_{f_{\mathrm{out}}=0}/\mathrm{CI}_{f_{\mathrm{out}}=0.1}\right)_{w_0} - 1$
$\sigma_b = 0.001$	0.21%
$\sigma_b = 0.05$	0.11%
$\sigma_b = 0.10$	0.24%

Table 5.12: The percentage relative difference on the CI at 68% CL of the marginal w_0 probability distribution with and without the shot noise, for different values of n_{bins}. All the other analysis settings are kept fixed at table 5.2. The cosmological model is $w_0 w_a$CDM with the parameters fixed at table 5.1.

We calculate the probability distribution of w_0 over a grid of values, as explained in section 5.3, by minimizing the χ^2:

$$\begin{aligned}\chi^2(w_0) = \sum_{\ell\ell'}\sum_{ij}&\left(C_\ell^{\mathrm{TG^i,fid}} - C_\ell^{\mathrm{TG^i},\mathrm{w_0}}\right)\\ &\times\left[\mathrm{Cov}(C_\ell^{\mathrm{TG^i,wrong}}, C_{\ell'}^{\mathrm{TG^j,wrong}})\right]^{-1}\left(C_{\ell'}^{\mathrm{TG^j,fid}} - C_{\ell'}^{\mathrm{TG^j},\mathrm{w_0}}\right).\end{aligned} \tag{5.19}$$

w_0w_aCDM	
	$\%\,(\mathrm{CI}_{\text{No noise}}/\mathrm{CI}_{\text{Noise}})_{w_0} - 1$
$n_{\text{bins}} = 1$	0.04%
$n_{\text{bins}} = 3$	0.20%
$n_{\text{bins}} = 10$	0.43%

Table 5.13: The percentage relative difference of the CI at 68% CL of the marginal probability distribution of w_0 when considering shot noise compared to when excluding it, for different values of n_{bins}. This distribution is obtained after marginalizing over w_a. All the other analysis settings are kept fixed at table 5.2. The cosmological model is w_0w_aCDM with the parameters fixed at table 5.1.

Here, C_ℓ^{TG,w_0} is the power spectrum computed at different values of w_0. We follow a similar procedure for $\Omega_c h^2$. In this case, the grid is computed over $\Omega_c h^2$ and A_s simultaneously, and we obtain the distribution for $\Omega_c h^2$ after marginalizing over A_s. The results for w_0 are presented in Fig. 5.7. In Fig. 5.7a and Fig. 5.7b we vary the number of redshift bins n_{bins} and the photometric error σ_b. Fig. 5.7c shows the percentage relative difference between the probability distribution of w_0 computed with the covariance at the wrong fiducials and the probability computed with the covariance at the exact fiducials, in the tomographic and non-tomographic case. The relative difference is below the order of $\sim 10^{-6}\%$ and thus negligible, although more pronounced in the tomographic case. Fig. 5.7a and 5.7b shows that the 'wrong' distribution for w_0 has its maximum in the fiducial value $w_0 = -1$. In table 5.15 are reported the best-fit values and the CI at 68% CL for the analysis with 'wrong' and 'exact' fiducials, both for the tomographic and non-tomographic cases. Both the best-fit values and the confidence intervals calculated with the wrong and exact fiducial values of w_0 are the same for both the tomographic and non-tomographic analyses.
The marginal distribution of $\Omega_c h^2$ is affected at most by approximately 60% when computing the covariance with the 'wrong' fiducial, as shown in Fig. 5.8c. These differences occur at the edges of the distribution, with the largest disparity observed in the tomographic case. Table 5.16 reports the best-fit values and the 68% confidence intervals of the distribution when the likelihood is computed with the 'wrong' and 'exact' fiducials, for both the tomographic and non-tomographic cases. The best-fit value remains the same when computed with both the 'wrong' and 'exact' fiducials in both analysis cases. However, the difference in the CI is 4% for the non-tomographic analysis and 6% for the tomographic one.
Fig. 5.8a and Fig. 5.8b display the distribution of $\Omega_c h^2$ when the covariance is computed using the 'wrong' fiducial values while varying n_{bins} and σ_b.

w_0w_aCDM	
	$(\mathrm{CI}/\mathrm{CI}_{\mathrm{fid}})_{w_0}$
fiducial	1
$\ell_{\mathrm{min}} = 5$	1.022
$\ell_{\mathrm{min}} = 10$	1.086
$\ell_{\mathrm{min}} = 30$	1.370
$n_{\mathrm{bins}} = 3$	1.076
$n_{\mathrm{bins}} = 1$	1.387
$f_{\mathrm{sky}} = 0.197$	1.230
$f_{\mathrm{sky}} = 0.307$	1.056
$f_{\mathrm{sky}} = 1.000$	0.616
$\sigma_b = 0.001$	1.009
$\sigma_b = 0.10$	0.997
$f_{\mathrm{out}} = 0.0$	1.001

Table 5.14: Here are reported the 68% confidence intervals (CI) on the distribution of w_0 with respect to different analysis settings of the *Euclid* photometric survey, after marginalizing over w_a. The fiducial values refer to our baseline in table 5.2, all the other analysis settings are taken accordingly unless otherwise stated in the first column. All the values are normalized by the 68% confidence interval of the fiducial baseline. The underlined fiducial model is w_0w_aCDM.

5.5 Validation of Limber approximation

In paragraph 4.2, we discuss the validity of the Limber approximation and its impact on the angular power spectra C_ℓ^{TG}. Here, we can assess this approximation using the likelihood. We compute the $\Delta\chi^2$ defined as:

$$\begin{aligned}\Delta\chi^2 \equiv& \chi^2_{\mathrm{Limber}}(\ell) - \chi^2_{\mathrm{No\ Limber}}(\ell) \\ =& \sum_{ij}\sum_{\ell'} C_\ell^{\mathrm{TG^i\,Limber}} \left[\mathrm{Cov}_{\ell\ell'}\right]^{-1}_{ij\,\mathrm{fid}} C_{\ell'}^{\mathrm{TG^j\,Limber}} \\ &- \sum_{ij}\sum_{\ell'} C_\ell^{\mathrm{TG^i\,No\ Limber}} \left[\mathrm{Cov}_{\ell\ell'}\right]^{-1}_{ij\,\mathrm{fid}} C_{\ell'}^{\mathrm{TG^j\,No\ Limber}},\end{aligned} \tag{5.20}$$

where the covariance is fixed at the fiducial cosmology, without the Limber approximation, and the power spectra $C_\ell^{\mathrm{TG^i\,Limber}}$ and $C_\ell^{\mathrm{TG^i\,No\ Limber}}$ are calculated with and without Limber approximation, respectively. We divide $\Delta\chi^2$ for $\chi^2_{\mathrm{No\ Limber}}$ and the result is Fig. 5.9. The percentage relative difference of χ^2 is less than 10^{-1}% when switching to the Limber approximation

Tomographic			
$w_{0,\text{best-fit}}^{\text{wrong}}$	CI^{wrong}	$w_{0,\text{best-fit}}^{\text{exact}}$	CI^{exact}
-1.010	0.209	-1.010	0.209
Non-Tomographic			
$w_{0,\text{best-fit}}^{\text{wrong}}$	CI^{wrong}	$w_{0,\text{best-fit}}^{\text{exact}}$	CI^{exact}
-1.010	0.209	-1.010	0.209

Table 5.15: Best-fit values and 68% confidence intervals obtained from the probability distribution of w_0 in the 'wrong fiducial analysis', in the tomographic and non-tomographic case. The covariance matrix is computed using the fiducial values for the cosmological parameters, as listed in Table 5.1, except for w_0, which is set to the incorrect value defined in Equation (5.17). Additionally, we present the results obtained from the analysis with the exact fiducial value for w_0.

Tomographic			
$\Omega_c h^2\|_{\text{best-fit}}^{\text{wrong}}$	CI^{wrong}	$\Omega_c h^2\|_{\text{best-fit}}^{\text{exact}}$	CI^{exact}
0.122	0.151	0.122	0.142
Non-Tomographic			
$\Omega_c h^2\|_{\text{best-fit}}^{\text{wrong}}$	CI^{wrong}	$\Omega_c h^2\|_{\text{best-fit}}^{\text{exact}}$	CI^{exact}
0.122	0.173	0.122	0.166

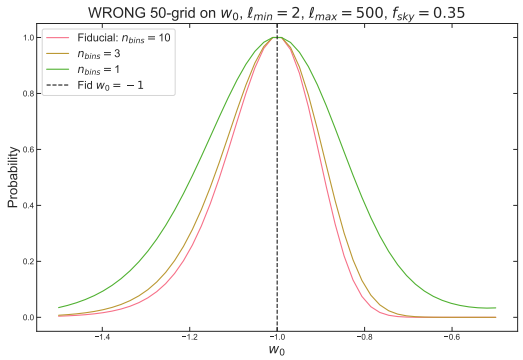
Table 5.16: Best-fit values and 68% confidence intervals obtained from the marginal probability distribution of $\Omega_c h^2$ in the 'wrong fiducial analysis', in the tomographic and non-tomographic case. The covariance matrix is computed using the fiducial values for the cosmological parameters, as listed in Table 5.1, except for $\Omega_c h^2$, which is set to the incorrect value defined in Equation (5.17). Additionally, we present the results obtained from the analysis with the exact fiducial value for $\Omega_c h^2$. The distribution is obtained after marginalizing over A_s.

at $\ell_{\text{min,Limber}} = 1$, approximating the C_ℓ^{TG} at each multipole. The difference is similar when $\ell_{\text{min,Limber}} = 30$, as the scales where the iSW signal is most pronounced are computed without any approximation. Therefore, although the Limber approximation is less accurate at the relevant multipoles for the iSW analysis, as discussed in paragraph 4.2, the error introduced by the Limber approximation is negligible with respect to the cosmic variance.

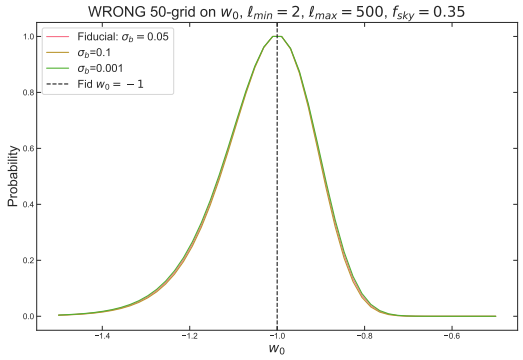
5.6 Code for the Cosmological Likelihood and Observables

The Cosmological Likelihood and Observables (CLOE) code is a Python software tool developed by the Inter-Science Taskforce Likelihood group (IST: Likelihood) within the *Euclid* collaboration. The IST: Likelihood team is responsible for the development and delivery of CLOE and collaborates closely with other Science Working Groups (SWG) within the project.
CLOE serves as a repository for likelihood functions associated with various observables. It interfaces with a Boltzmann code, which computes theoretical predictions for a given cosmological model, and employs a Bayesian sampler to calculate the likelihood function at multiple points within the parameter space to be explored.
Within the CMBX SWG, we first validated the theoretical recipe of the cross-correlation observable, i.e. the cross-correlation angular power spectrum described in eq. (4.23). Here we used the Limber approximation. To validate the power spectrum, we estimated the constraints on the A_{iSW}. We consider as "theory" the theoretical recipe we want to validate and as "data" the power spectrum computed by the Boltzmann code. To recover the posterior distribution for the A_{iSW} parameter we performed an MCMC analysis using the likelihood provided by Planck and the Gaussian likelihood (see eq. (5.13)).
Second, we worked on the implementation of the likelihood framework for the cross-correlation between CMB temperature anisotropies and GC data, particularly from the photometric analysis. In particular, we worked in collaboration with the CMB lensing-GC cross-correlation analysis group of the CMBX SWG in order to build a full analytical covariance that contains the terms of the cross-correlation between CMB probes (temperature and lensing) and the GC. Then, this covariance was added to the covariance implemented by the IST:Likelihood in CLOE that contains the terms from the main probes from *Euclid*, i.e. the spectroscopic GC, the photometric GC and the cross-correlation between the two, the so-called 3x2pt block of the covariance. The full covariance matrix that includes the cross-correlation between the photometric GC and the CMB cross is called the 7x2pt covariance. It is loaded in the code whether the CMB cross-correlation probes are asked by the user. I had the task to review the implementation of the 7x2pt

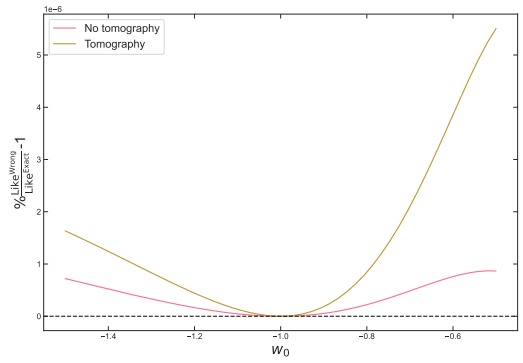
covariance. We made sure our code is flexible and with the same structure as CLOE to avoid any possible conflicts.
Now that the CMBxGC spectra and likelihood are validated, we are currently merging our code into CLOE. We split our modifications of CLOE into three blocks to facilitate the review process by the iST:Likelihood group. The first block concerns all the extra files that we want to add to CLOE; the second block concerns the additions to the existing code, without modifying the initial structure; the third blocks concerns all the actual modifications to CLOE.
Our implementation is currently undergoing a review and evaluation process.



(a) The probability distribution for w_0 varying n_{bins} computed with the covariance at the wrong fiducial.

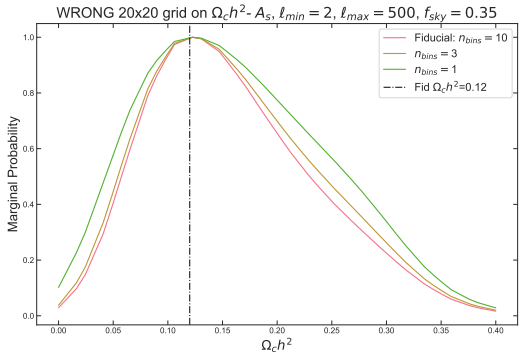


(b) The probability distribution for w_0 varying σ_b computed with the covariance at the wrong fiducial.

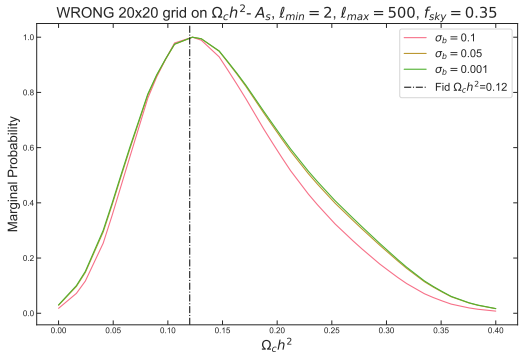


(c) Percentage difference of the probability distribution of w_0, computed with the covariance at the wrong fiducial and exact fiducial, in the tomographic and non-tomographic case.

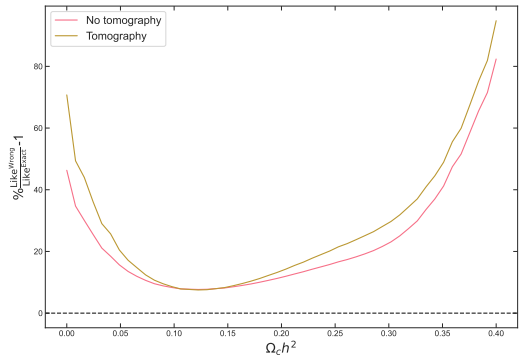
Figure 5.7: The probability distribution for the parameter w_0, calculating the covariance using a wrong fiducial value of w_0. The fiducial value for w_0 is shown with the black line. All the analysis parameters (not shown in the plots, see table 5.2) are held constant, and the cosmological parameters are fixed according to the values in table 5.2.



(a) The probability distribution for $\Omega_c h^2$ varying n_{bins} computed with the covariance at the wrong fiducial.



(b) The probability distribution for $\Omega_c h^2$ varying σ_b computed with the covariance at the wrong fiducial.



(c) Percentage difference of the probability distribution of $\Omega_c h^2$, computed with the covariance at the wrong fiducial and exact fiducial, in the tomographic and non-tomographic case.

Figure 5.8: The marginal probability distribution for the parameter $\Omega_c h^2$, calculating the covariance using a wrong fiducial value of $\Omega_c h^2$. The fiducial value for $\Omega_c h^2$ is shown with the black line. All the analysis parameters (not shown in the plots, see table 5.2) are held constant, and the cosmological parameters are fixed according to the values in table 5.2.

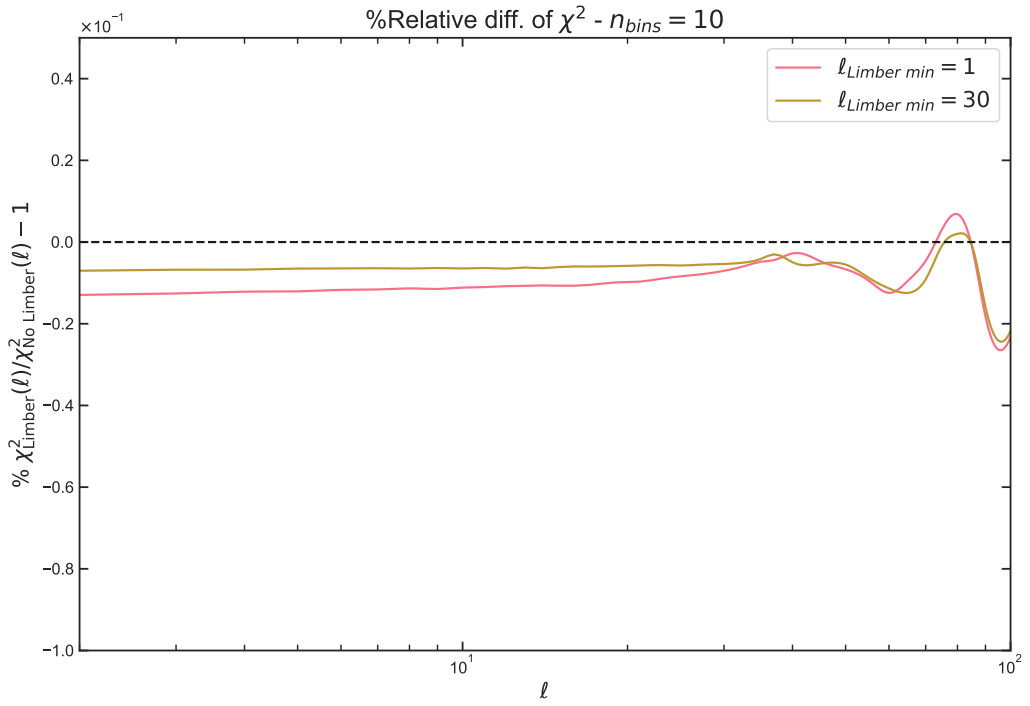


Figure 5.9: This figure illustrates the percentage relative difference in χ^2 when computed with and without the Limber approximation, as described in equation (5.20). The analysis is performed for two different values of the minimum multipole at which the Limber approximation is applied, $\ell_{\text{Limber}} = 1$, $\ell_{\text{Limber}} = 30$. All cosmological parameters are held fixed at the baseline values listed in table 5.1.

Chapter 6

Needlet estimator for the cross-correlation observable

In this chapter, I introduce a promising estimator for the cross-correlation observable, which is based on spherical wavelet functions, called needlets, which share unique properties with other spherical wavelets, both in the harmonic and real space. In particular, during my PhD I have worked on the validation of this estimator, analyzing realistic sky maps and estimating the cross-correlation observable.
Section 6.1 describes the advantages of needlets that make them well-suited for the analysis of cosmological data. Section 6.2 focuses on the construction of needlets and their mathematical properties.
The needlet-based cross-correlation estimator is presented in Section 6.3, while Section 6.4 describes its validation over a set of simulations, together with the methodology and the results obtained.
Lastly, in Section 6.5, we present preliminary results from a forecast analysis using simulations for the Euclid survey.

6.1 Needlets for cosmological data analysis

As extensively discussed in Chapter 4 and Chapter 5, the standard cross-correlation analysis between Cosmic Microwave Background (CMB) and galaxy clustering reconstructs the two-point statistics in harmonic space. When dealing with full-sky observations and isotropic noise, the harmonic estimator proves to be optimal and unbiased, as previously discussed. Additionally, each multipole is independent, resulting in a diagonal covariance matrix. This allows us to estimate the cross-correlation observable with minimal error, primarily driven by cosmic variance.
However, realistic has to handle masked data and missing observations due to foreground contamination and the survey's observation strategy. In such scenarios, the harmonic estimator is no longer optimal, and approximations

are employed, such as the $f_{\rm sky}$ approximation .
In this chapter, I introduce and discuss an alternative analysis approach for CMB and GC data, based on the use of needlet functions, a specific type of spherical wavelet. Spherical wavelets are functions defined on the 2D sphere and possess the unique property of double localization, both in harmonic and real space. This is an advantage when analyzing maps with incomplete data. In the harmonic analysis, the lack of information in the real space leads to spreading the effect on the spherical harmonics observables, introducing correlations that are difficult to handle. Thus, the spherical wavelets provide a very natural alternative for this kind of analysis. In particular, the needlets have very peculiar asymptotic properties, which make the needlet coefficients uncorrelated for any fixed angular scales as the frequency (the scale at which we analyze the map) increases. These properties are valid also in the presence of a masked region.
Several studies and analyses have been produced in this direction, in particular for the CMB field analysis. Ref. [73] describes the asymptotic properties of needlets, while Ref. [67] describes the application of needlet analysis to CMB data. Needlets have been employed, for instance, in estimating non-Gaussianity in the CMB field (Ref. ([74]); Planck. ([75])).
This work focuses on the estimation of the cross-correlation observable between CMB and LSS using the needlets. In literature, this application was studied in Ref. [1] for the detection of the iSW signal, while Ref. [2] reports the analysis of the CMB lensing- GC cross-correlation signal estimated with needlets and in the case of masked sky.

6.2 Building the needlet functions

The purpose of this section is to provide an overview of the mathematical formalism behind the needlet function and to show its primary properties. The discussion is based on Ref. [76]) and Ref. ([67]).
To construct the needlets, it is essential to introduce two components: the "cubature points" and the "cubature weights," which discretize the sphere. The goal is to choose a suitable discretization scheme for the sphere, in order to replace an integral of a function with a summation without loss of any accuracy. In practice, we define a set of cubature weights λ_{jk} corresponding to a set of cubature points ξ_{jk}, such that over the surface of the sphere S^2 we have:

$$\int_{S^2} dx, f(x) \rightarrow \sum_k f(\xi_{jk})\lambda_{jk}\,. \tag{6.1}$$

Here, $f(x)$ is a function defined on the sphere, and λ_{jk} are real, positive numbers. The meaning of the indices j and k will soon become clear.
An example of discretization of the sphere is the grid defined by the standard package HEALpix ([77]). This code divides the sphere into a set of N_{pix} pixels, each with an identical area and equispaced distribution. The number

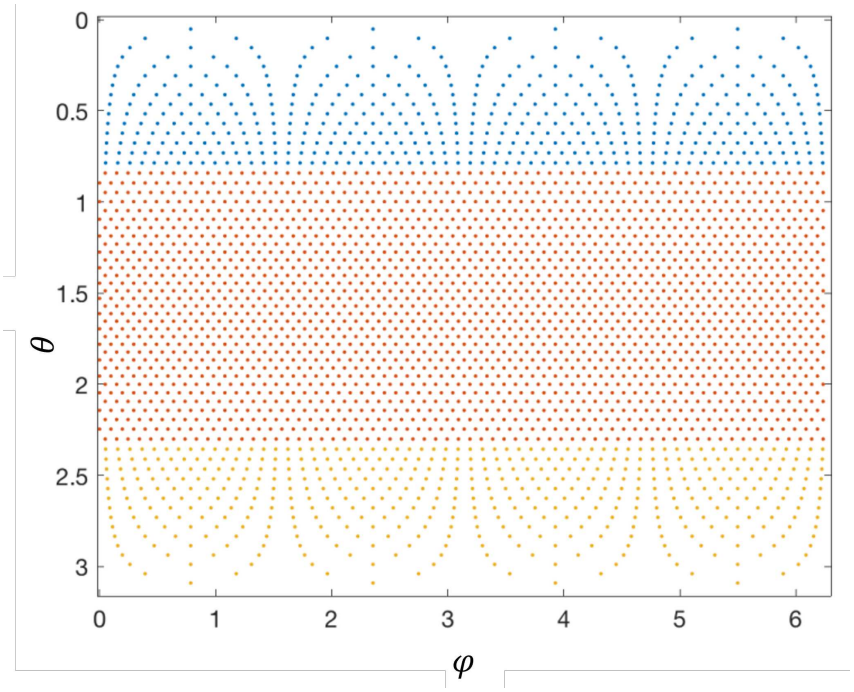


Figure 6.1: HEALPix grid where θ is the latitude and φ is the longitude, in the range $[0, 2\pi] \times [0, \pi]$. The pixels in the northern, equatorial, and southern regions are shown in blue, red, and yellow, respectively. Image from [78]

of pixels in these grids is determined by the resolution, expressed as $N_{pix} = 12N_{side}^2$, where $N_{side} = 2^k$ ($k \in \mathbb{N}$). Each cubature point is specified by the coordinates θ_k, φ_k, denoting the longitude and latitude of a pixel in the grid, while the cubature weight corresponds to the pixel's area:

$$\begin{aligned} \{\xi_{jk}\} &= \{\theta_k, \varphi_k\} \\ \{\lambda_{jk}\} &= \frac{4\pi}{N_{pix}}. \end{aligned} \tag{6.2}$$

With this definition, the index k represents the k-th pixel on the grid and ranges from 1 to N_{pix}, while the index j is associated with the scale or frequency at which the HEALPix map is analyzed. In the general case, the number of pixels might be different for different frequencies, but here it is considered the simplest and most common pixelization. HEALPix pixelization is the most used pixelization for the data analysis of cosmological fields that are projected into the 2D sphere, such as the CMB.
Figure 6.1 shows the projected HEALPix grids, featuring equispaced pixels divided into three regions: equatorial, northern, and southern. At each cubature point of the HEALPix grid, for a given frequency j, the spherical

needlets ψ_{jk} are defined as follows:

$$\begin{aligned}\psi_{jk}(\hat{\mathbf{n}}) &\equiv \sqrt{\lambda_{jk}} \sum_{\ell=D^{j-1}}^{D^{j+1}} w\Big(\frac{\ell}{D^j}\Big) \sum_{m=-\ell}^{\ell} Y_{\ell m}(\xi_{jk}) Y^*_{\ell m}(\hat{\mathbf{n}}) \\ &= \sqrt{\lambda_{jk}} \sum_{\ell=D^{j-1}}^{D^{j+1}} w\Big(\frac{\ell}{D^j}\Big) \frac{2\ell+1}{4\pi} P_\ell(\langle \xi_{jk}, \hat{\mathbf{n}} \rangle)\,,\end{aligned} \tag{6.3}$$

Here, $\hat{\mathbf{n}}$ denotes a direction on the sphere, and $\{\xi_{jk}, \lambda_{jk}\}$ is a set of cubature points and weights, as defined previously. Additionally, D is a real and positive constant, and $w(\cdot)$ represents a weight function. Eq. (6.3) shows that the needlets can be seen as a convolution of the projection operator $\sum_{m=-\ell}^{\ell} Y_{\ell m}(\xi_{jk}) Y^*_{\ell m}(\hat{\mathbf{n}})$ with a weight function $w(\cdot)$. The weight function is defined across the multipole space, with its width regulated by the free parameter D. This function is constructed to satisfy three critical conditions:

1. $w\Big(\frac{\ell}{D^j}\Big)$ is different from zero for $\ell \in \big[D^{j-1}, D^{j+1}\big]$, ensuring that needlets have bounded support in the harmonic domain:
2. $w(\cdot)$ is infinitely differentiable, i.e. is a smooth function, a crucial property for the localization properties in real space;
3. $$\sum_{j=1}^{\infty} b^2\Big(\frac{\ell}{D^j}\Big) \equiv 1 \text{ for all } \ell > D$$

   is essential for reconstructing a field after it has been expanded into needlet functions, as will be discussed.

One way to build the weight function in order to satisfy the condition above is the following:

- STEP 1: Construct the function

  $$\phi_1(t) = \begin{cases} \exp\Big(-\frac{1}{1-t^2}\Big) & -1 \le t \le 1 \\ 0 & \text{otherwise} \end{cases}$$

  This function is C^∞ and compactly supported in $[-1, 1]$.

- STEP 2: Build the C^∞ function

  $$\phi_2(u) = \frac{\int_{-1}^{u} f(t)dt}{\int_{-1}^{1} f(t)dt},$$

  which is normalized such that $\phi_2(-1) = 0$ and $\phi_2(1) = 0$;

- STEP 3: Create the function

$$\phi_3(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq \frac{1}{D} \\ \phi_2\Big(1 - \frac{2D}{D-1}\Big(t - \frac{1}{D}\Big)\Big) & \text{if } \frac{1}{D} \leq t \leq 1 \\ 0 & \text{if } t > 1 \end{cases}$$

- STEP 4: Define the weight function

$$w^2(x) = \phi_3\Big(\frac{x}{D}\Big) - \phi_3(x),$$

  which satisfies all the aforementioned conditions.

The weight functions are plotted in fig 6.2, for two different values of the parameter D and different frequency bands. Once D is fixed, the number of frequencies and multipoles are related by

$$\ell_{\max} = D^{j_{\max}}\,. \tag{6.4}$$

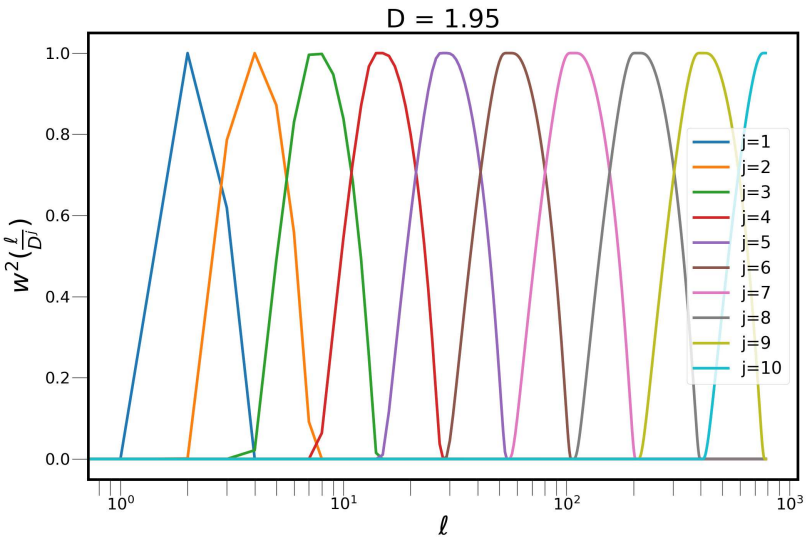
A smaller value of D corresponds to a broader localization in ℓ-space, as shown in Fig. 6.2b, while a larger value results in a narrower localization. Conversely, in real space, a larger D leads to a more precise localization of the functions.
With D fixed, each needlet can capture the signal only from a specific range of multipoles within a given frequency band j. There are alternative methods for constructing suitable weight functions for the needlets, such as employing the cosine function and utilizing a B-spline approach (see, for example, Ref. [76]).
For a field $T(\hat{\mathbf{n}})$ defined on the sphere, the coefficients in the expansion using the needlet basis $\psi_{jk}(\hat{\mathbf{n}})$ are as follows:

$$\begin{aligned} a_{jk} &= \int_{S^2} d\Omega T(\hat{\mathbf{n}})\psi_{jk}(\hat{\mathbf{n}}) \\ &= \sqrt{\lambda_{jk}}\sum_{\ell m} w\Big(\frac{\ell}{D^j}\Big)Y_{\ell m}(\xi_{jk})\underbrace{\int_{S^2} d\Omega T(\hat{\mathbf{n}})Y^*_{\ell m}(\hat{\mathbf{n}})}_{a_{\ell m}} \\ &= \sqrt{\lambda_{jk}}\sum_{\ell m} a_{\ell m} w\Big(\frac{\ell}{D^j}\Big)Y_{\ell m}(\xi_{jk})\,. \end{aligned} \tag{6.5}$$

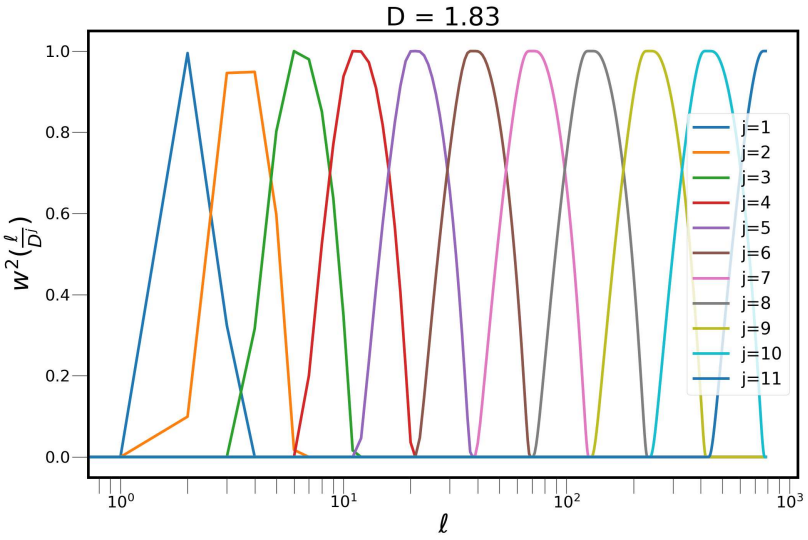
Here, $a_{\ell m}$ represents the spherical harmonics coefficient of the field T.

6.2.1 Properties of the needlet functions

As previously mentioned, spherical needlets offer various advantages for data analysis applications. From the computational point of view, needlets are extremely convenient because are very easy to implement with the



(a) Weight functions with $D = 1.95$.



(b) Weight functions for $D = 1.83$.

Figure 6.2: Weight function $w^2(\cdot)$ in the multipole space for difference frequency j. The needlet scale parameter is set to $D = 1.95$ in the upper panel and to $D = 1.83$ in the bottom panel.

HEALpix pixelization. Specifically, for a fixed j, the needlet coefficients a_{jk} are a HEALPix map itself, with k representing the index of the pixels, differently from the spherical harmonic coefficients.
Furthermore, while not constituting an orthonormal basis, needlets constitute a "tight frame". This means that expanding a function in terms of needlet coefficients preserves the norm of the function. This property is evident when summing over all k and j for the a_{jk} terms defined in eq. (6.5), taking advantage of the third condition for the weight functions. Consequently, needlets allow for a straightforward reconstruction formula:

$$T(\hat{\mathbf{n}}) = \sum_{jk} a_{jk} \psi_{jk}(\hat{\mathbf{n}}) \,. \tag{6.6}$$

This formula allows the transformation of the field back to its original state without any loss of information.
The properties of the weight functions not only localize the needlets in ℓ-space but, as previously mentioned, also yield significant localization in real space. More specifically, it can be demonstrated (as shown in Ref. [79]) that the needlets exhibit quasi-exponential localization around any cubature point ξ_{jk}. Due to their smoothness, the following expression holds:

$$|\psi_{jk}(\hat{\mathbf{n}})| \leq \frac{c_M D^j}{(1 + D^j\, d(\xi_{jk}, \hat{\mathbf{n}})} \,, \tag{6.7}$$

for any number M and constant c_M, where $d(,)$ is the angular distance. This means that for a fixed value of the angular distance, the value of $\psi_{jk}(\hat{\mathbf{n}})$ goes to zero quasi-exponentially as D decreases. The anticipated trade-off between harmonic and real spaces is as follows: smaller D values lead to a more concentrated localization in harmonic space (with fewer multipoles involved in each needlet), whereas choosing larger values guarantees a faster decay in real space.
Other properties of the needlets are the correlation properties. Consider a random field T, with spherical harmonics coefficients drawn from a distribution with a mean of $\langle a_{\ell m} \rangle = 0$ and variance C_ℓ. Firstly, for two different frequency bands j and j', the needlet coefficients are uncorrelated by design when $|j - j'| \geq 2$ (this is evident from Fig. 6.2, as the weight functions only overlap with the closest frequency bands).
For the case of $j = j'$, the variance of the needlet coefficients is given by:

$$\begin{aligned} \langle a_{jk} a^*_{jk} \rangle &= \lambda_{jk} \sum_{\ell = D^{j-1}}^{D^{j+1}} w^2\Big(\frac{\ell}{D^j}\Big) C_\ell \underbrace{\sum_{m=-\ell}^{\ell} Y_{\ell m}(\xi_{jk}) Y^*_{\ell m}(\xi_{jk})}_{\frac{2\ell+1}{4\pi}} \\ &= \lambda_{jk} \sum_{\ell = D^{j-1}}^{D^{j+1}} \frac{2\ell+1}{4\pi} w^2\Big(\frac{\ell}{D^j}\Big) C_\ell \,, \end{aligned} \tag{6.8}$$

using the property of the spherical harmonics in the last equality. This expression is nearly uniform across all the pixels, expressed as:

$$\sigma_j^2 \equiv \frac{1}{N_{pix}} \sum_{\ell=D^{j-1}}^{D^{j+1}} \frac{2\ell+1}{4\pi} w^2\Big(\frac{\ell}{D^j}\Big) C_\ell \,. \tag{6.9}$$

The off-diagonal term $j = j' \pm 1$ is:

$$\sigma_{jj'}^2 = \frac{1}{N_{pix}} \sum_{\ell=D^{j'-1}}^{D^{j'+2}} \frac{2\ell+1}{4\pi} w\Big(\frac{\ell}{D^{j'+1}}\Big) w\Big(\frac{\ell}{D^{j'}}\Big) C_\ell P_\ell(d(\xi_{j'+1k}, \xi_{j'k}))\,, \tag{6.10}$$

where P_ℓ is the Legendre polynomial.
Another noteworthy property is that as the frequency j increases, needlet coefficients remain asymptotically unaffected by the presence of missing observations or masked regions on the sphere. Being $\bar{a}_{jk}$ the coefficients evaluated on the partially observed sphere and a_{jk} the coefficients evaluated on the fully observed sphere, one has:

$$\sqrt{\langle \bar{a}_{jk} - a_{jk} \rangle^2} \propto \frac{D^j}{(1 + D^j \epsilon)} \sqrt{\langle T^2(\hat{\mathbf{n}}) \rangle}\,. \tag{6.11}$$

In this equation, ϵ represents the distance from pixel k to the masked region. This expression shows that for a fixed distance ϵ, $\bar{a}_{jk}$ becomes closer to a_{jk} as the frequency increases.

6.2.2 Data analysis with the needlets

In the previous section, I discussed the primary properties of needlets. Now, apply these functions are applied to data analysis.
The analysis operation yields an approximation to the spherical harmonic coefficients from the data T on the sphere, following the HEALPix pixelization:

$$\hat{a}_{\ell m} = \frac{4\pi}{N_{npix}} \sum_{i=1}^{N_{pix}-1} Y^*_{\ell m}(\theta_i, \varphi_i) T(\theta_i, \varphi_i), \quad 0 \le \ell \le \ell_{\max}, -\ell \le m \le \ell\,. \tag{6.12}$$

Here, the indices i and j iterate over the number of pixels on the grid, and $\ell_{\max}$ represents the maximum multipoles at which the spherical harmonics expansion is truncated. This truncation is forced by the finite resolution of the map[1].
Inserting the above expression into eq. (6.5), the needlet coefficients resulting

[1]The HEALPix code limits the multipoles to $\ell_{\max} = 3N_{side} - 1$, for a given fixed value of N_{side}.

from the analysis of T on the HEALPix map are given by:

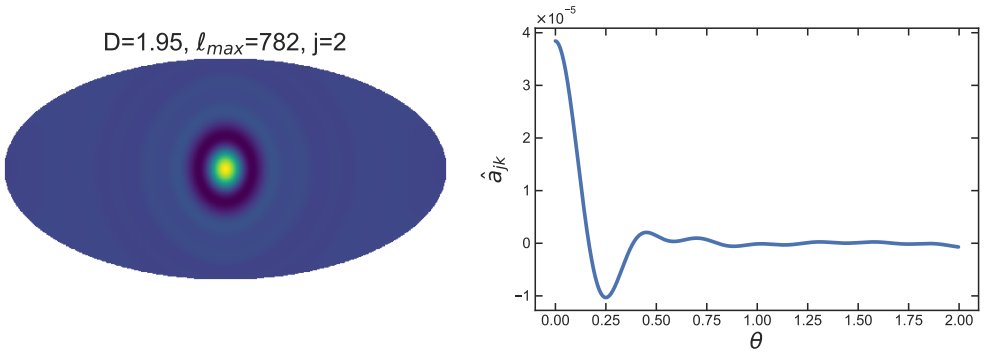
$$\begin{aligned}\hat{a}_{jk} &= \sqrt{\lambda_{jk}} \sum_{\ell=D^{j-1}}^{D^{j+1}} w\Big(\frac{\ell}{D^j}\Big) \sum_{m=-\ell}^{\ell} \hat{a}_{\ell m} Y_{\ell m}(\xi_{jk}) \\ &= \frac{4\pi}{N_{pix}} \sqrt{\lambda_{jk}} \sum_{\ell=D^{j-1}}^{D^{j+1}} w\Big(\frac{\ell}{D^j}\Big) \sum_{m=-\ell}^{\ell} \sum_{i=1}^{N_{pix}-1} T(\theta_i, \varphi_i) Y^*_{\ell m}(\theta_i, \varphi_i) Y_{\ell m} \,.\end{aligned} \tag{6.13}$$

Fig. 6.3 shows an example of signal analysis using needlets. Considering a signal that is zero in every pixel except for the central one, which is set to one, the analysis of the map with needlets for various values of j is shown in the left panels of the figures. These panels show the needlets coefficients on a HEALPix map. The parameters D and $\ell_{\max}$ are kept fixed. As the frequency increases, the localization in pixel space improves. This improvement is also evident in the right panels, where the needlets coefficients are plotted as functions of the angles.
Another example of needlet-based data analysis involves analyzing the signal of the CMB temperature anisotropies. we generate a CMB map using the theoretical angular power spectrum C_ℓ^{TT}, assuming the ΛCDM cosmological model, and fixing the cosmological parameters to the values specified in Table 1.1 from the Planck mission. Subsequently, we analyze this CMB map with needlets, considering two different values for D and varying frequencies j. The results of this analysis are presented in Fig. 6.4. The interplay between pixel space and harmonic space is evident: for higher values of j, the signal becomes more localized within pixels, while it extends over a broader range of ℓ values in harmonic space. Conversely, increasing D at a fixed j results in greater localization of the signal in pixel space and a broader spectrum of multipoles captured by the needlets. This property is useful, for instance, in testing for the presence of systematics or spurious signals, knowing the scales at which the signal we want to extract peaks.

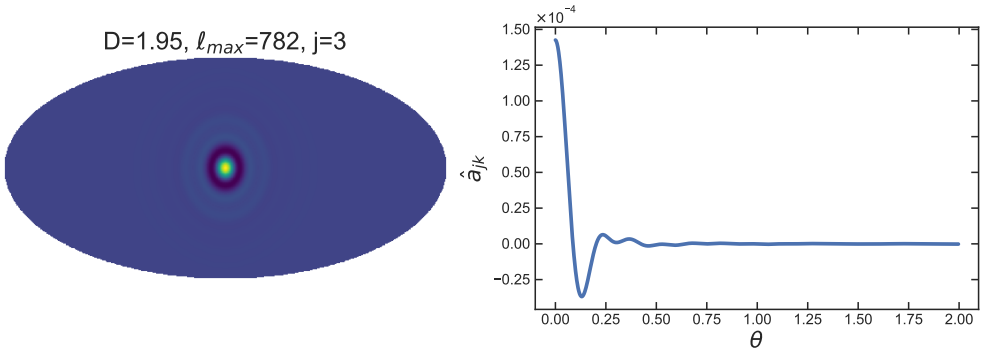
6.3 The cross-correlation estimator

In this analysis, our goal is to estimate the cross-correlation observable between the large-scale CMB temperature anisotropy and the GC. For this purpose, we consider two maps: one of the CMB temperature field and the other of the galaxy overdensity field. we proceed by extracting the needlets coefficients a^T_{jk} for the CMB and a^G_{jk} for the galaxy overdensity field. To build an estimator, we use the following formula:

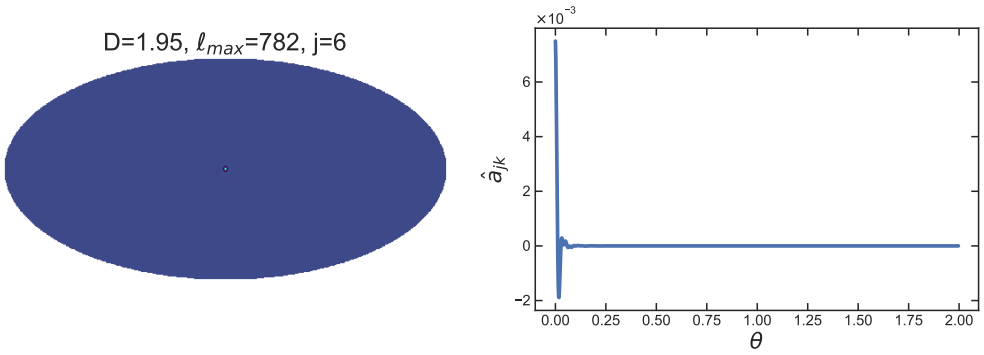
$$\hat{\beta}_j^{\mathrm{TG}} = \frac{1}{N_{pix}} \sum_k a^T_{jk} a^{*G}_{jk} \,. \tag{6.14}$$



(a) Needlets coefficients of a signal centered in one pixel for $j = 2$, in pixel space.

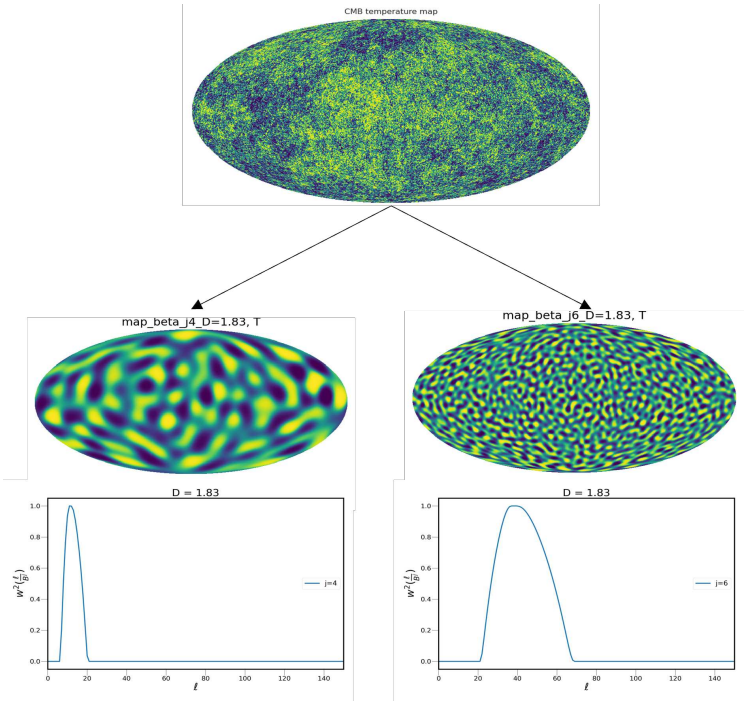


(b) Needlets coefficients of a signal centered in one pixel for $j = 3$, in pixel space.

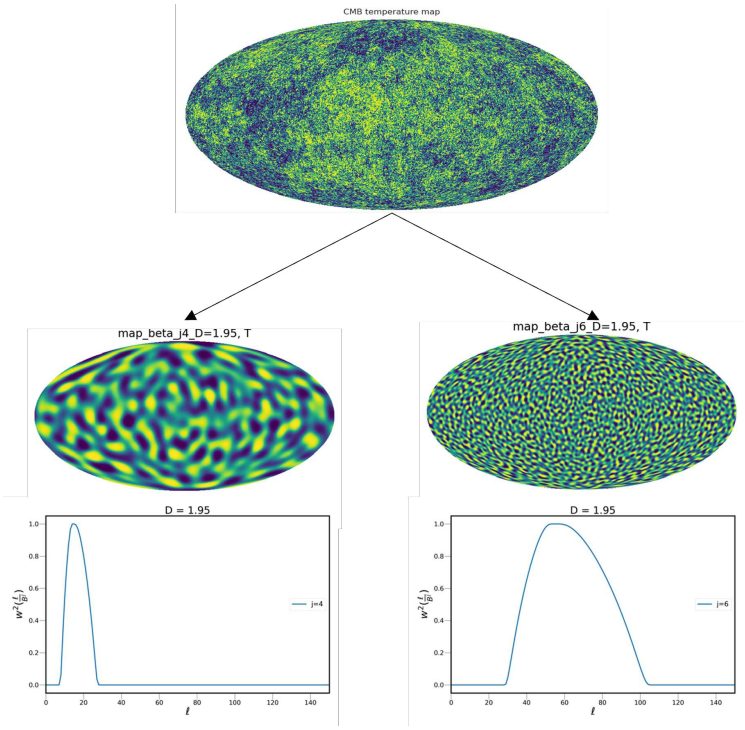


(c) Needlets coefficients of a signal centered in one pixel for $j = 2$, in pixel space.

Figure 6.3: Needlets coefficients of a signal equal to zero in each pixel but one in the center, for different values of j, in pixel space. In each figure, the left panel represents the HEALPix map of $\hat{a}_{jk}$, while the right panel is the projection in angle. The parameter D is fixed at $D = 1.95$ and $\ell_{\max} = 782$



(a) Analysis of the CMB temperature anisotropies with $D = 1.83$ and $j = 4, 6$



(b) Analysis of the CMB temperature anisotropies with $D = 1.95$ and $j = 4, 6$

Figure 6.4: Analysis of the CMB temperature anisotropies with the needlets, for two values of D and two different frequency bands. The CMB map is shown on the left, the maps of the needlet coefficient are in the center, and on the right are plotted the needlet weight functions for each frequency band considered.

The expectation value of $\hat{\beta}_j^{\rm TG}$ is expressed as:

$$\langle \hat{\beta}_j^{\rm TG} \rangle \equiv \beta_j^{\rm TG} = \sum_\ell \frac{2\ell+1}{4\pi} w\Big(\frac{\ell}{D^j}\Big) C_\ell^{\rm TG} \,. \tag{6.15}$$

This implies that it is an unbiased estimator of the angular power spectrum. To be precise, the needlets estimator provides an unbiased estimate of a binned version of the $C_\ell^{\rm TG}$, with each bin being weighted by the needlets' weight functions and the factor $(2\ell+1)/4\pi$.
The variance of the estimator, when $j = j'$, can be expressed as:

$$\begin{aligned}(\Delta\beta_j^{\rm TG})^2 \equiv {\rm Var}\left[\hat{\beta}_j^{\rm TG}\right] &= \sum_\ell \frac{(2\ell+1)^2}{16\pi^2} w^4\Big(\frac{\ell}{D^j}\Big) (\Delta C_\ell^{\rm TG})^2 \\ &= \sum_\ell \frac{2\ell+1}{16\pi^2} w^4\Big(\frac{\ell}{D^j}\Big) \left[C_\ell^{\rm TT} C_\ell^{\rm GG} + (C_\ell^{\rm TG})^2\right] \,.\end{aligned} \tag{6.16}$$

In the last expression, we employ the formula for the variance of the harmonic estimator of $C_\ell^{\rm TG}$ (as seen in eq. (4.29)). Similar to the harmonic analysis, we can substitute $C_\ell^{\rm GG} \to C_\ell^{\rm GG} + N_\ell^{GG}$ to account for shot noise.
As discussed in section 6.2, the application of needlet functions in cosmological studies is particularly interesting due to their behavior over a sphere with masked regions. Specifically, the coefficients a_{jk} remain asymptotically unaffected by the presence of masked regions as $j \to \infty$. However, it's worth noting that a study presented in Ref. [2] demonstrates that the estimator defined in eq (6.14) becomes biased when aggressive masking is applied[2]. They propose an approach similar to the one used for the harmonic estimator to mitigate this bias. Once the needlet coefficients have been extracted from a masked sky using the formula:

$$\tilde{a}_{jk} = \sqrt{\lambda_{jk}} \sum_\ell w\Big(\frac{\ell}{D^j}\Big) \sum_m \tilde{a}_{\ell m} Y_{\ell m}(\xi_{jk}) \,, \tag{6.17}$$

we can construct the estimator for the cross-correlation observable as:

$$\hat{\Gamma}_j^{\rm TG} = \frac{1}{N_{pix}} \sum_k \tilde{a}_{jk}^{\rm T} \tilde{a}_{jk}^{*,{\rm G}} \,. \tag{6.18}$$

Using the equation for the pseudo-C_ℓ as given in eq. (4.25), the expectation value is expressed as:

$$\begin{aligned}\langle \hat{\Gamma}_j^{\rm TG} \rangle \equiv \Gamma_j^{\rm TG} &= \sum_{\ell m} w^2\Big(\frac{\ell}{D^j}\Big) \langle \tilde{a}_{\ell m}^{\rm T} \tilde{a}_{\ell m}^{*\,{\rm G}} \rangle \\ &= \sum_{\ell\ell'} \frac{2\ell+1}{4\pi} w^2\Big(\frac{\ell}{D^j}\Big) M_{\ell\ell'} C_{\ell'}^{\rm TG} \,,\end{aligned} \tag{6.19}$$

[2]The analysis in their study utilizes a map with a sky fraction $f_{\rm sky}$=0.013, but in section 6.5, we will demonstrate that even with $f_{\rm sky}$=0.36, the estimator in eq. (6.14) remains biased

where $M_{\ell\ell'}$ is the coupling matrix determined by the mask. The expression above means that $\hat{\Gamma}_j^{\prime\,\mathrm{TG}}$ is a biased estimator for C_ℓ^{TG} but an unbiased estimator for a binned version of the pseudo-C_ℓ, weighted by the needlets weight function. To make a comparison, $\hat{\Gamma}_j^{\mathrm{TG}}$ behaves like a pseudo-needlet estimator. For a large fraction of the observed sky, the expression in Eq. (6.19) simplifies to $\hat{\Gamma}_j^{\mathrm{TG}} \approx f_{\mathrm{sky}} \langle \hat{\beta}_j^{\mathrm{TG}} \rangle$. This is called the f_{sky} approximation, as discussed in section 4.3. Finally, the associated variance for $j = j'$ is given by:

$$\begin{aligned}(\Delta\Gamma_j^{\mathrm{TG}})^2 &= \sum_{\ell\ell'} \frac{(2\ell+1)^2}{16\pi^2} w^4\Big(\frac{\ell}{D^j}\Big) M_{\ell\ell'} (\Delta C_{\ell'}^{\mathrm{TG}})^2 \\ &= \sum_{\ell\ell'} \frac{2\ell+1}{16\pi^2} w^4\Big(\frac{\ell}{D^j}\Big) M_{\ell\ell'} \left[C_{\ell'}^{\mathrm{TG}} + C_{\ell'}^{\mathrm{TT}} C_{\ell'}^{\mathrm{GG}}\right] .\end{aligned} \tag{6.20}$$

The pseudo-C_ℓ approach typically employs a "backward modeling" approach in which measurements are deconvolved from systematic effects, like the cut sky, to obtain an unbiased estimate of the power spectrum. The recovered spectrum in eq. (4.27) can be directly compared to the theoretical prediction. A similar approach for the needlets has not been developed yet, it would require computing an invertible coupling matrix, either numerically or analytically, such that $\hat{\Gamma}_j^{\mathrm{TG}} = \sum_{jj'} M_{jj'} \beta_j^{\mathrm{TG}}$ and $\beta_j^{\mathrm{TG}} = \beta_j^{\mathrm{TG}}(\hat{\Gamma}_j^{\mathrm{TG}})$. However, the expression in (6.19) can always be compared to raw measurements and can estimate the pseudo-spectra, which are affected by the presence of cut-sky, without introducing bias, following a "forward modeling" approach.

6.4 Validation of the estimator

In this section, I describe the validation process for the needlets estimator and its variance.

I have developed a Python code to simulate correlated HEALPix maps of the CMB temperature anisotropies and galaxy overdensity. These simulations are based on the fiducial angular power spectra (C_ℓ^{TT}, C_ℓ^{TG}, C_ℓ^{GG}) computed using CAMB. The fiducial model follows the ΛCDM cosmological parameters from Planck 2018 (see table 1.1). For this analysis, we consider the non-tomographic case with a Gaussian redshift distribution of galaxies centered at $z = 1$ and having a width of $\sigma_z = 0.25$. The linear bias is set to one ($b_z = 1$). we exclude all relativistic corrections to the projected galaxy overdensity field, such as redshift-space distortions and lensing effects, and focus solely on the density term, which accounts for dark matter perturbations.

Each validation run uses fixed values of $\ell_{\mathrm{max}} = 782$ and $j_{\mathrm{max}} = 12$. The parameter D is determined by inverting the relationship given in eq. (6.4).

The resolution of the HEALPix maps is set to $N_{\rm side} = 512$, and we simulate a total of $N_{\rm sim} = 500$ maps for each field.
Each simulated map is generated using a HEALPix routine. It starts with generating a realization of $a_{\ell m}^{\rm T(G)}$ drawn from a Gaussian distribution with a mean of $\langle a_{\ell m}^{\rm T(G)} \rangle = 0$ and a variance equal to the fiducial $C_\ell^{\rm TT(GG)}$ computed using CAMB.
To correlate the CMB and galaxy maps, we follow the prescription outlined in [80]. we draw two complex numbers, (ζ_1, ζ_2), from a Gaussian distribution with unit variance. Using these numbers, we compute the correlated fields $a_{\ell m}^{\rm T}$ and $a_{\ell m}^{\rm G}$ as follows:

$$a_{\ell m}^{\rm T} = \zeta_1 \sqrt{(C_\ell^{\rm TT})} \tag{6.21}$$

$$a_{\ell m}^{\rm G} = \zeta_1 \frac{C_\ell^{\rm TG}}{\sqrt{C_\ell^{\rm TT}}} + \zeta_2 \sqrt{C_\ell^{\rm GG} - \frac{(C_\ell^{\rm TG})^2}{C_\ell^{\rm TT}}}\,. \tag{6.22}$$

Here, $C_\ell^{\rm TG}$ represents the fiducial cross-correlation angular power spectrum obtained from the CAMB code. After computing these correlated $a_{\ell m}^{\rm T}$ and $a_{\ell m}^{\rm G}$, we synthesize them into CMB and galaxy maps using an HEALPix routine.
From the simulated maps, the code extracts the corresponding needlets coefficients a_{jk}^{T} and a_{jk}^{G}. Then it calculates the cross-correlation estimator as defined in equation (6.15). Subsequently, we compute the expectation value and the covariance of the estimator using the following expressions:

$$\langle \hat{\beta}_j^{\rm TG} \rangle = \frac{1}{N_{\rm sim-1}} \sum_{i}^{N_{\rm sim}-1} \hat{\beta}_{j,i}^{\rm TG} \tag{6.23a}$$

$$\mathrm{Cov}_{jj'} = \frac{1}{N_{\rm sim}-1} \sum_{i=1}^{N_{\rm sim}} (\hat{\beta}_{j,i}^{\rm TG} - \langle \hat{\beta}_j^{\rm TG} \rangle_{\rm sim})(\hat{\beta}_{j',i}^{\rm TG} - \langle \hat{\beta}_{j'}^{\rm TG} \rangle_{\rm sim})\,. \tag{6.23b}$$

Here, the index i runs over the simulations. By comparing eq. (6.23) with the theoretical predictions provided in equations (6.15) and (6.16), we can validate our theoretical framework. Additionally, for the validation process, we conduct a significance test to detect the iSW effect and estimate constraints on the density parameter of dark energy, Ω_Λ.

6.4.1 Null-test analysis

The first test we perform to validate the needlet estimator is the null-test, i.e. we check that the estimator is zero when the signal from the map to estimate is zero. To do so, we manually set $C_\ell^{\rm TG} = 0$ and then we generate a set of uncorrelated CMB and galaxies maps. In Fig. 6.5, we present the

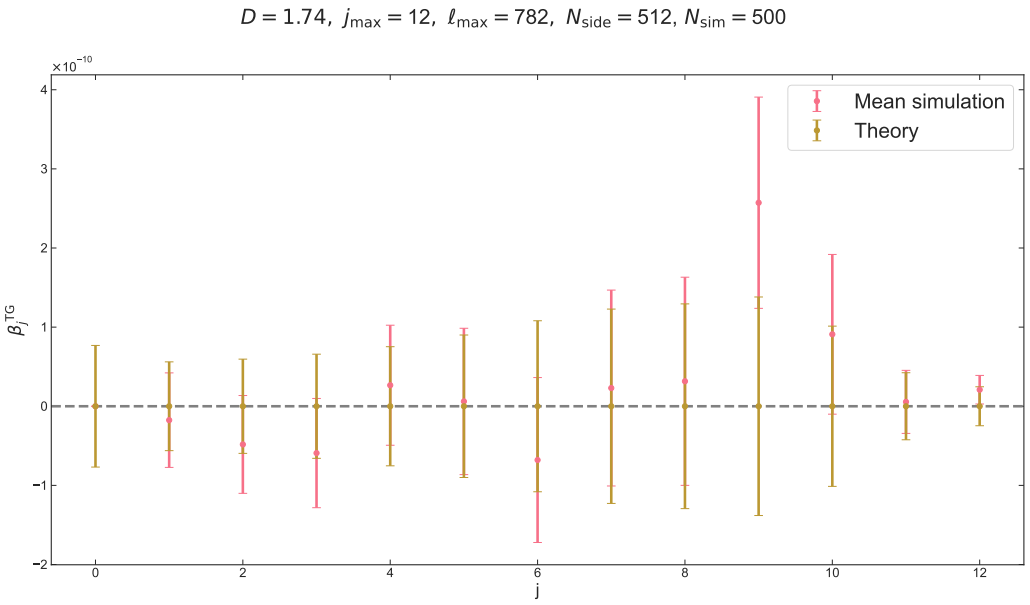


Figure 6.5: Results of the null test. The pink dots are the simulated needlet power spectrum averaged over all the simulations, and the error bars are the variance of the simulated covariance divided by N_{sim}. In yellow is plotted the theoretical β_j^{TG}, which is zero as expected, and the error bars are calculated from the expected variance (eq. (6.16)) divided by N_{sim}.

mean of the recovered needlets spectra from these simulations in pink, with the theoretical β_j^{TG} shown in yellow, which, as anticipated, is expected to be zero. The error bars on the mean are determined by the diagonal elements of the simulated covariance matrix divided by N_{sim}, while the error of the theoretical needlets spectrum is calculated using eq. (6.16). From the plot, it is evident that the estimated power spectrum is consistent with the theoretical one within the error. Additionally, we conduct a χ^2-test under the null hypothesis that the iSW signal is present, defined as:

$$\tilde{\chi}^2 = \sum_j \frac{(\hat{\beta}_{j,\,\mathrm{one-sim}}^{TG} - \langle \hat{\beta}_j^{TG} \rangle_{\mathrm{sim,\,signal}})^2}{(\Delta\beta_j^{TG})^2}\,. \tag{6.24}$$

Here, $\hat{\beta}_{j,\,\mathrm{one-sim}}^{TG}$ represents the needlet power spectrum from one realization of uncorrelated fields, while $\langle \hat{\beta}_j^{TG} \rangle_{\mathrm{sim,\,signal}}$ is the mean obtained from a set of simulated maps involving correlated fields.
The variance $(\Delta\beta_j^{TG})^2$ represents the error associated with the theoretical spectrum β_j^{TG} calculated with $C_\ell^{TG} = 0$. we also verify that the variable $\tilde{\chi}^2$ follows a χ^2 distribution by constructing its probability distribution function from the simulations. In Fig. 6.6, I present the histograms for the $\tilde{\chi}^2$ distribution. These histograms are generated under the null hypothesis

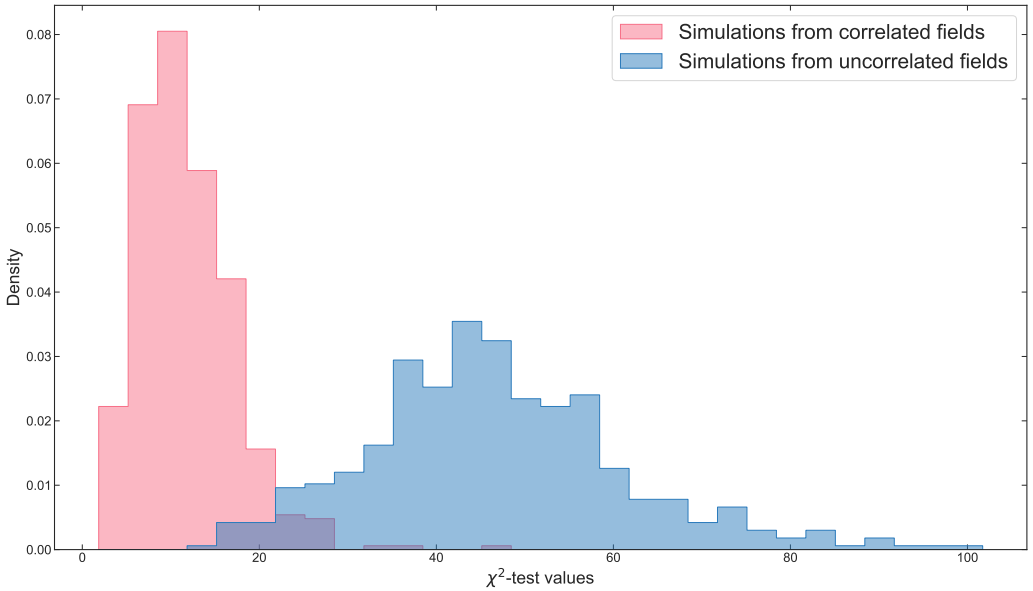


Figure 6.6: The probability distribution function of the $\tilde{\chi}^2$ variables, assuming the null hypothesis that there is no iSW signal, i.e. that the fields are uncorrelated. The histogram in blue is filled from the simulations generated from uncorrelated fields, the pink one is filled from simulations generated from correlated fields.

mentioned earlier (i.e., the presence of the iSW signal). One histogram is built using simulations derived from uncorrelated fields (in blue), while the other uses simulations from correlated fields (in pink). The distributions represent two χ^2 distributions, each with $j_{\rm max}$ degrees of freedom. The two sets of variables calculated from correlated and uncorrelated fields are incompatible under the same hypothesis.
For one realization generated from uncorrelated fields, we obtained $\tilde{\chi}^2 = 68$. Assuming that it follows a χ^2 distribution with $j_{\rm max}$ degrees of freedom, we can reject the null hypothesis with a 99.9% level of significance.

6.4.2 Validation with Planck cosmology with full sky observations

After the null test, we validate the needlet estimator by detecting the iSW signal and estimating the constraints on Ω_Λ using a set of correlated CMB and galaxies HEALPix maps. The cosmology is fixed to ΛCDM with the Planck 2018 best-fit values. As a first step, we consider only the case of a full-sky observation without shot noise. The recovered needlet power spectrum, $\langle \hat{\beta}_j^{\rm TG} \rangle_{\rm sim}$, and the theoretical spectrum are plotted in Fig. 6.7. The errors on the simulation are obtained from the diagonal of the covariance matrix computed from the simulation, while the error bars on the mean

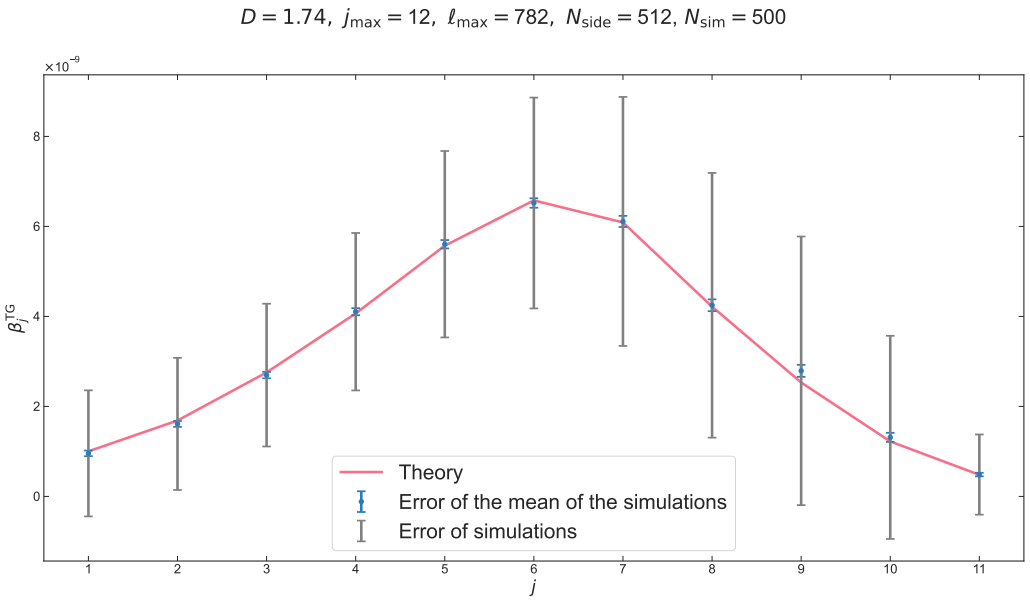


Figure 6.7: The plot displays the theoretical and simulated needlet power spectra. The pink line represents the theoretical power spectrum, the blue curve shows the mean of all simulations along with their respective errors, and the grey shading represents the errors derived from the simulations.

are the same but divided by $N_{\rm sim}$. In Fig. 6.8, it is plotted the relative difference between the mean of the simulated needlet power spectrum and the theoretical power spectrum. The relative differences are less than approximately 10%, and the error bars are consistent. In Fig. 6.9, we also plot the difference between $\langle\hat{\beta}_j^{\rm TG}\rangle_{\rm sim}$ and $\beta_j^{\rm TG}$ in units of the theoretical variance. The recovered power spectrum deviates from the expected value by at most 2σ.

The full covariance matrix computed from the simulations is shown in Fig. 6.10 (notice that the main diagonal is not displayed here). As anticipated in the discussions presented in section 6.2.1, the correlation is stronger for j' values that are close to $j \pm 1$, and then it gradually decreases.

After validating the theoretical variance, we compute the significance of the iSW signal. To begin, we compute the signal-to-noise ratio (S/N), analogous to the equation (4.35), as follows:

$$\left(\frac{\rm S}{\rm N}\right)_j^2 \equiv \frac{(\beta_j^{\rm TG})^2}{(\Delta\beta_j^{\rm TG})^2}\,. \tag{6.25}$$

We then compare the expected value, calculated with the theoretical $\beta_j^{\rm TG}$, with the value computed using the recovered power spectrum $\langle\hat{\beta}_j^{\rm TG}\rangle_{\rm sim}$, as shown in Fig. 6.11. From the S/N analysis, the signal is detected within

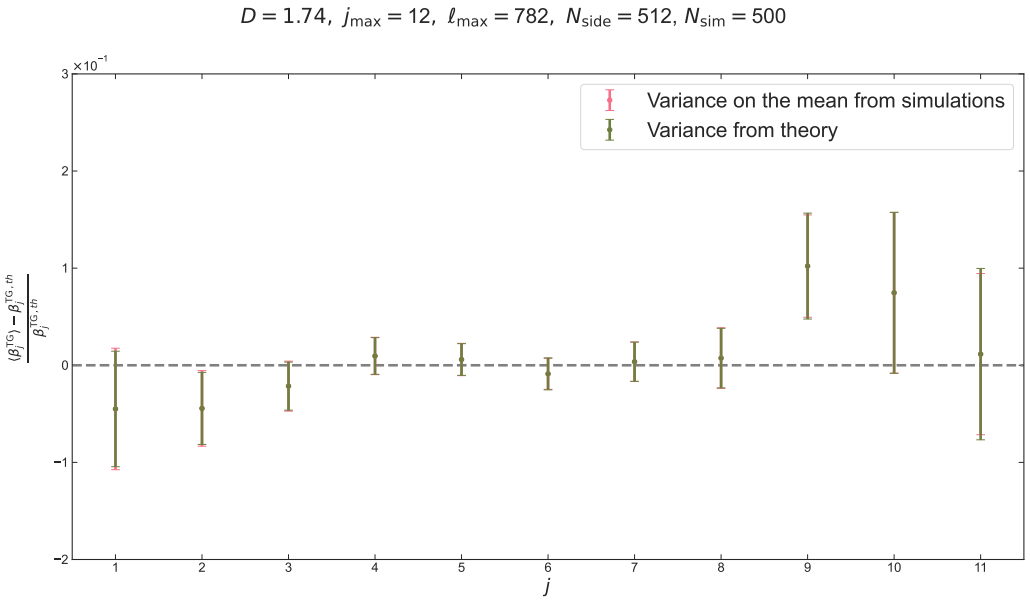


Figure 6.8: This plot illustrates the relative difference between the mean of the simulated $\hat{\beta}_j^{\rm TG}$ and the theoretical value $\beta_j^{\rm TG}$. The green error bars represent the errors on the mean of the simulations, calculated using the diagonal of the simulated covariance matrix (divided by $N_{\rm sim}$). The pink bars indicate the errors computed from the theoretical variance according to equation (6.16).

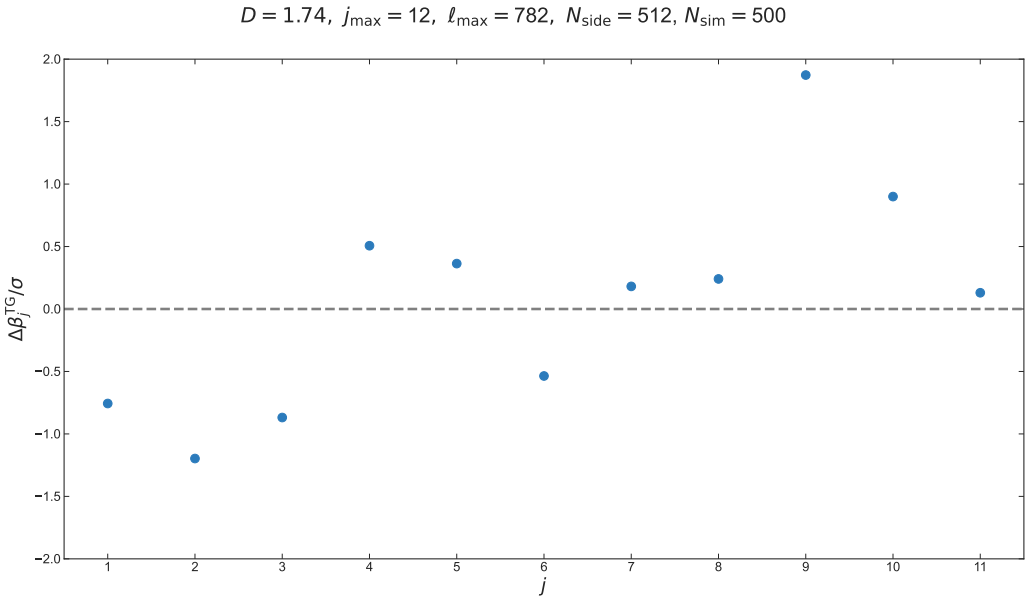


Figure 6.9: The difference between the mean of the simulated power spectra and the theoretical spectrum, normalized by the theoretical variance $(\Delta\beta_j^{\rm TG})^2$.

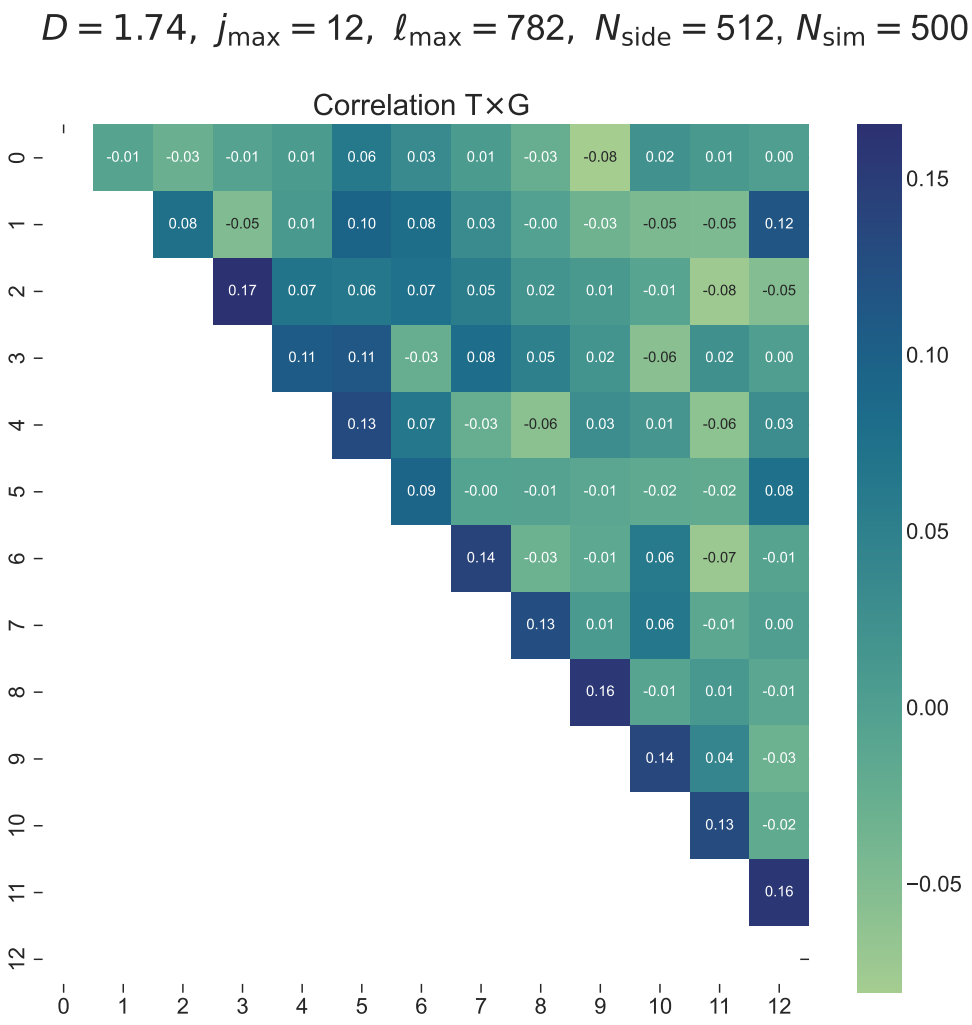


Figure 6.10: The covariance matrix and the correlations between different values of j, computed from the simulations of CMB and galaxies maps, as from the second expression in eq. (6.23). The main diagonal is not displayed here.

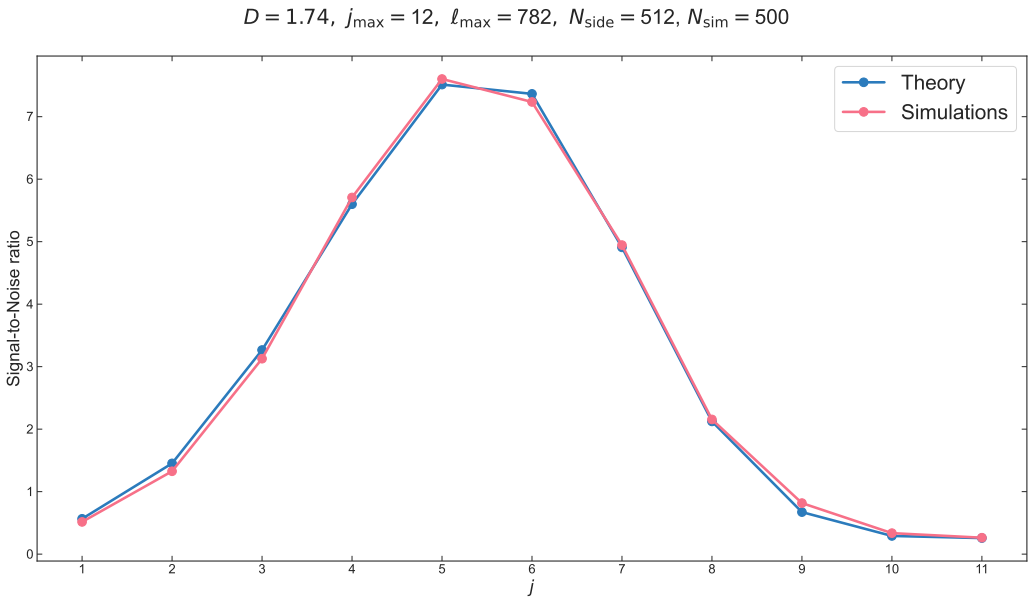


Figure 6.11: The Signal-to-Noise ratio (S/N) as computed from eq. (6.25). The blue line is the expected S/N computed from the theoretical $\beta_j^{\rm TG}$; in pink is shown the S/N computed from the simulations.

the range of approximately $1 \lesssim j \lesssim 10$. Using the relationship between the parameter D, j, and $\ell_{\rm max}$, this corresponds to scales where $\ell \lesssim 200$. This result aligns with the expectations from the harmonic analysis of the S/N, as shown in Fig. 4.15. Calculating the cumulative S/N as:

$$\frac{S}{N} = \sqrt{\sum_{j=0}^{j_{\rm max}} \left(\frac{\rm S}{\rm N}\right)_j^2}, \tag{6.26}$$

the overall value is 5.8. When calculating the S/N with the full covariance from the simulations, i.e. not with the analytical expression for the variance, the cumulative S/N decreases up to 5. This is because the analytical approximation of the variance fails with the off-diagonal term and we are limited by the number of simulations available. This point is under investigation.
It's important to note that this analysis is conducted under the simplest conditions, assuming full-sky observations and zero noise. Therefore, this number may not be particularly significant on its own, but it will be useful for comparison with the case of masked sky and shot noise.
To assess the statistical significance of the iSW detection, we conduct a χ^2 test, similar to Eq. 6.24, but under the null hypothesis that no detection exists. In this test, we extract one estimated needlet power spectrum from the set of estimations derived from the simulations of correlated fields, and then we compare it with the mean of the simulations generated from the uncorrelated maps.

The result of this test yields a χ^2 value of 23.75. Thus it is excluded that the $\hat{\beta}_j^{\rm TG}$ were generated under the null hypothesis with a confidence level of 96.65%.

Parameter estimation We recover the fiducial values of $A_{\rm iSW}$ and Ω_Λ parameters from the simulations. For the $A_{\rm iSW}$ parameter, we create a grid of 100 values spanning the range $A_{\rm iSW} \in [0, 2]$. We then compute a Gaussian likelihood function given by:

$$\begin{aligned}\chi^2 &\equiv -2\log\mathcal{L} \\ &= \sum_{jj'} \left(\hat{\beta}_j^{\rm TG} - A_{\rm iSW}\beta_j^{\rm TG}\right) \left[{\rm Cov}_{jj'}\right]^{-1} \left(\hat{\beta}_{j'}^{\rm TG} - A_{\rm iSW} \times \beta_{j'}^{\rm TG}\right) ,\end{aligned} \tag{6.27}$$

where $\hat{\beta}_j^{\rm TG}$ represents a single realization extracted from the simulations, $\beta_j^{\rm TG}$ is the theoretical needlet power spectrum, and ${\rm Cov}_{jj'}$ is the covariance matrix computed from simulations. The likelihood function is shown in Fig. 6.12, along with the best-fit value and the Confidence Interval at 68% Confidence Level. This plot shows an example of what one can expect to observe from one single realization of the observed sky. We also verify that the results are not sensitive to the choice of realization. To do this, we calculate the likelihood in eq. (6.27) for each simulation and create a histogram of the best-fit values. The result is in Fig. 6.13 and we find the same distribution as the one in Fig. 6.12. However, it's worth noting that the statistical power is limited by the the number of simulations available.

The constraining power for the detection of the iSW effect with needlets is $A_{\rm iSW}/\sigma_{\rm best-fit} = 5$. we apply the same procedure to the DE density parameter Ω_Λ. we produce a grid of 30 theoretical needlet power spectra $\beta_j^{TG,\Omega_\Lambda}$, each computed with a different value of $\Omega_\Lambda \in [0., 0.95]$. As before, we calculate a Gaussian likelihood as

$$\begin{aligned}\chi^2 &\equiv -2\log\mathcal{L} \\ &= \sum_{jj'} \left(\hat{\beta}_j^{\rm TG} - \beta_j^{TG,\Omega_\Lambda}\right) \left[{\rm Cov}_{jj'}\right]^{-1} \left(\hat{\beta}_{j'}^{\rm TG} - \beta_{j'}^{TG,\Omega_\Lambda}\right) ,\end{aligned} \tag{6.28}$$

where $\hat{\beta}_j^{\rm TG}$ is the power spectrum from one simulation. The probability distribution function for Ω_Λ is shown in Fig. 6.14. We also plot the distribution of the best-fit values obtained from all the simulations, as shown in Fig. 6.15. We find that the fiducial value is recovered within the 1σ confidence interval.

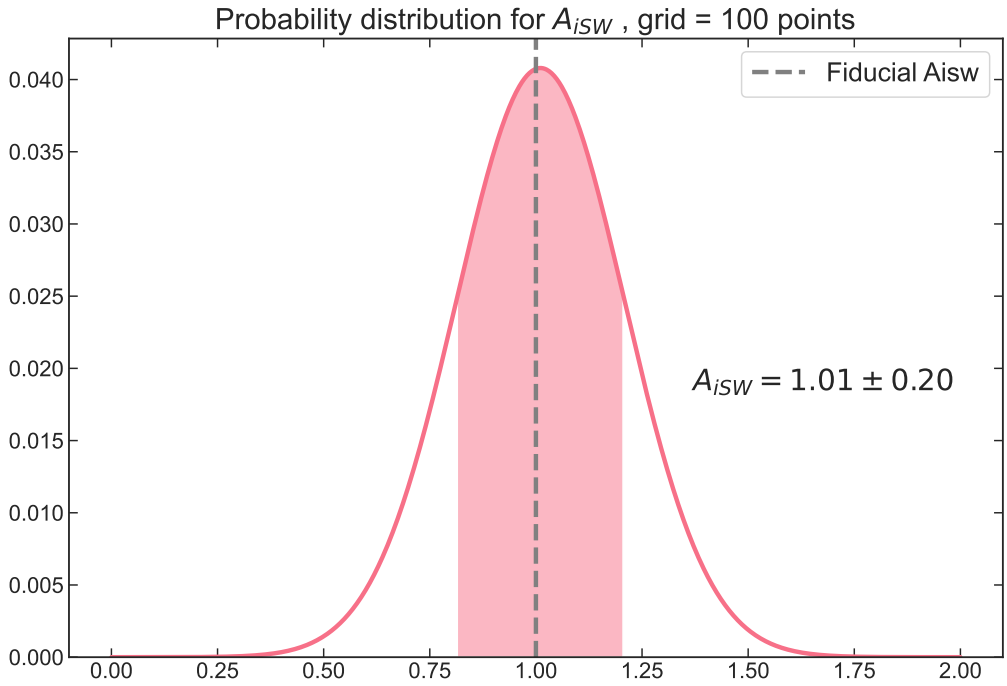


Figure 6.12: The likelihood function for the $A_{\rm iSW}$ parameter. The cosmological parameters are held fixed to the Planck 2018 cosmology. The shaded area corresponds to the 68% confidence interval. The best fit value and the Confidence Interval are indicated in the text.

6.4.3 Validation with Planck cosmology, with masked sky and shot noise

We repeat the same analysis implementing a mask and accounting for shot noise. We apply the Planck mask, which covers 78% of the sky (as shown in Fig. 6.16), to both the simulated CMB and galaxy overdensity maps. The galaxy number count is approximately 5.76×10^5 galaxies per steradian, from the Herschel ATLAS survey [81]. This value for the shot noise is chosen to account for the worst-case scenario.
From the masked maps, we extract unbiased estimates, approximating the recovered needlet power spectra as $\hat{\beta}_j^{\rm TG} \approx \tilde{\beta}_j^{\rm TG}/f{\rm sky}$, where $\tilde{\beta}_j^{\rm TG}$ represents the raw spectra. Given the high sky coverage fraction, around 0.78, this approximation is expected to hold. As in the previous section, Fig. 6.17 shows the recovered power spectrum $\hat{\beta}_j^{\rm TG}$, together with the theoretical power spectrum.

The relative difference between the recovered and theoretical power spectra is shown in Fig. 6.18. This difference becomes more evident as the value of j increases, due to the presence of shot noise. This is evident in Fig. 6.19, where the same difference is plotted with and without the shot noise term.

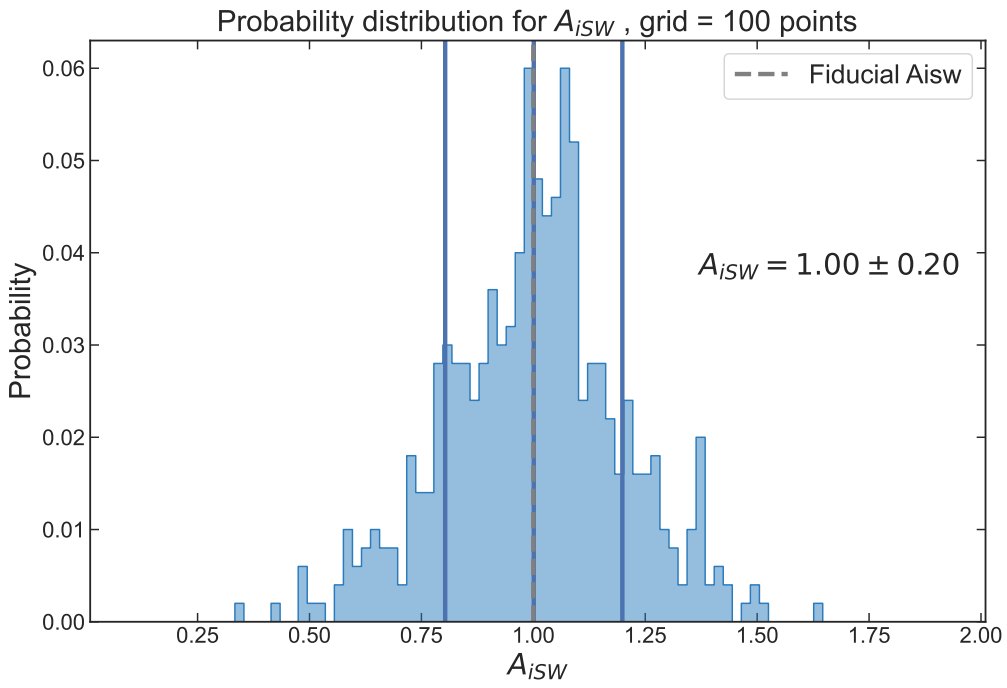


Figure 6.13: Histogram of the best-fit values of $A_{\rm iSW}$ computed from a Gaussian likelihood for each simulation. The dotted line represents the fiducial value, while the blue lines correspond to the percentiles at the 68% confidence level and the mean of the best-fit values.

The shot noise induced a difference in the estimation of the power spectrum of the order of $\sim 1.73 - 2.3\sigma$ from the estimation without shot noise in the last two frequency bins. This is expected since the shot noise affects the smaller scales of the power spectrum. Moreover, the presence of shot noise affects also the scales in the first frequency bin, inducing a difference of $\sim 1.2\sigma$.
As in the previous validation, we compute the difference between $\hat{\beta}_j^{\rm TG}$ and $\beta_j^{\rm TG}$ in order of σ, as shown in Fig. 6.20, where σ is the theoretical variance as defined in eq. (6.16), including the shot noise term. The estimate of the needlet spectrum from the simulations deviates from the theoretical prediction by up to $\sim 3\sigma$. This deviation is more pronounced compared to the full-sky analysis (as shown in Fig. 6.9), which is expected because both the shot noise and the mask contribute to increased uncertainties in the measurements. It's worth noting that our conclusions are also influenced by the limited number of simulations available, which contributes to the observed discrepancy.
The analytical approximation of the variance in eq. (6.16) is consistent with the variance computed from the simulations, with a difference up to 10%, as illustrated in Fig. 6.21. The total covariance from the simulations carries

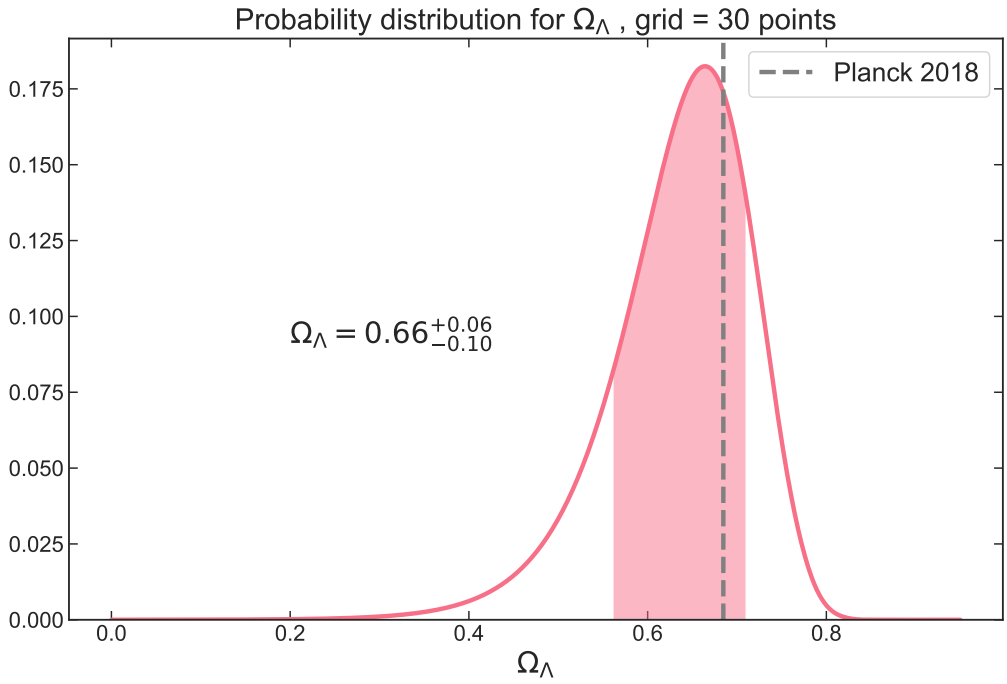


Figure 6.14: The probability distribution function of Ω_Λ drawn from a Gaussian likelihood computed over a grid of 30 needlet power spectrum. The covariance is computed from the simulations. The grey dotted line is the Planck fiducial value $\Omega_\Lambda^{\rm Planck} = 0.6847$, while the shaded area rapresents at 68% confidence level. In the plot are also written the best-fit value and the range of the CI. The fiducial cosmology is ΛCDM fixed at the Planck 2018 parameters.

also interesting information about the properties of the needlets, especially when compared with the covariance from full-sky simulations, as shown in Fig. 6.22. The correlation property discussed in section 6.2.1 holds true even for observations with a cut-sky. Specifically, two different frequency bins exhibit lower correlation as they become more separated from each other. This highlights one of the advantages of employing a needlet-based estimator.
The analysis of the S/N is illustrated in Fig. 6.23, where we calculate the S/N for each j using eq. (6.25). In this calculation, we use the theoretical variance $(\Delta\beta_j^{\rm TG})^2$, which includes the shot noise term, and the $\hat{\beta}_j^{\rm TG}$ values estimated from the cut-sky simulations. The results are consistent with the full-sky analysis, but with a smaller cumulative S/N value due to the noise and the cut-sky, yielding a cumulative S/N of 3.3. When calculating the S/N with the full covariance from the simulations, as in the full-sky analysis, the cumulative S/N decreases up to 2.6.

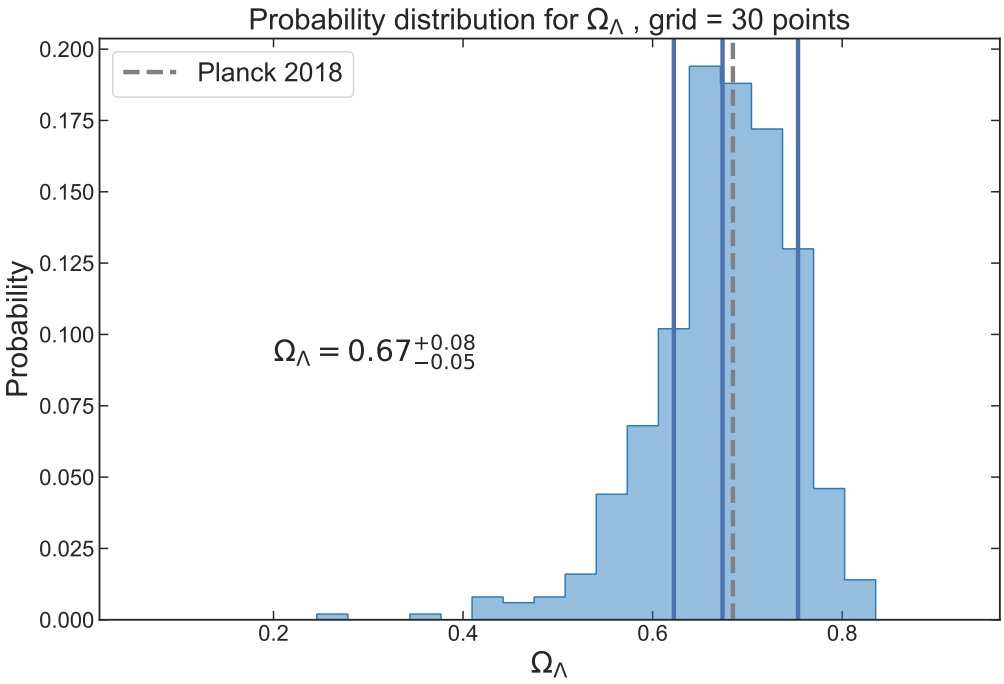


Figure 6.15: Distribution of the best-fit values of Ω_Λ computed from the Gaussian likelihood for each simulation. The blue lines represent the mean of the best-fit values and the percentiles of the distributions at 68%. we recover the fiducial value within the 1σ confidence interval. The fiducial cosmology is ΛCDM fixed at the Planck 2018 parameters.

Parameter estimation As in the case of the full-sky analysis, the needlets estimator is validated at the parameter estimation level. We employ the same methodology outlined in paragraph 6.4.2 for the $A_{\rm iSW}$ and Ω_Λ parameters. Fig. 6.24 shows the probability distribution for $A_{\rm iSW}$ obtained from a single simulation, as an example of what one can expect to observe from one random realization of the observed sky. Fig. 6.25 displays the distribution of best-fit values computed from all the simulations. From the distribution of the best-fit values, the expected fiducial value of $A_{\rm iSW} = 1$ is recovered within 1σ. However, when compared to the results obtained from the full-sky analysis, we observe larger uncertainties on the $A_{\rm iSW}$ parameter, resulting in a constraining power of $A_{\rm iSW,\,best-fit}/\sigma = 2.7$.
The same results hold for the estimation of the Ω_Λ parameter, as Fig. 6.26 and Fig. 6.27 show.

6.5 Analysis with Euclid simulations

We apply the needlets analysis on a set of simulations provided by the CMBXC SWG of the *Euclid* collaboration. These simulations are 1000 HEALPix

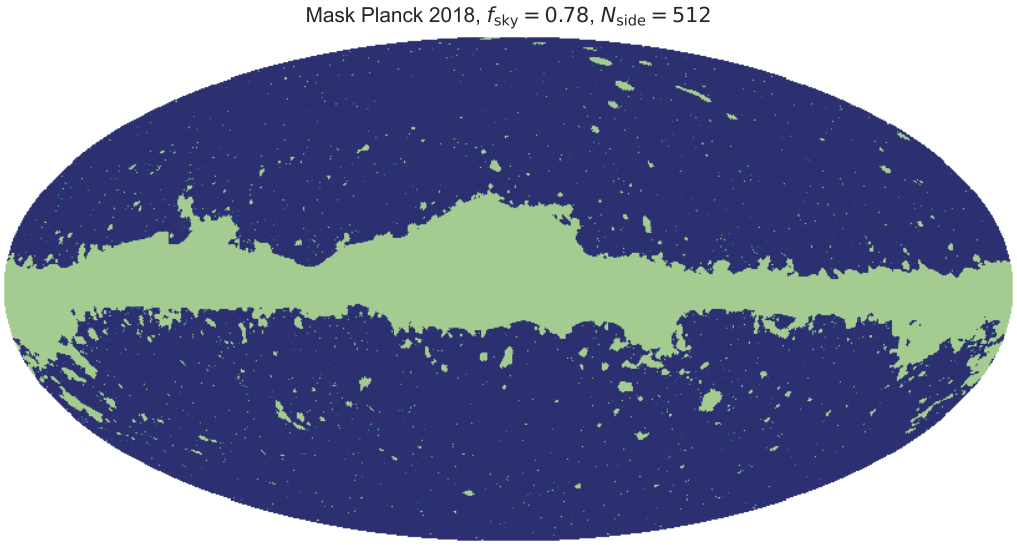


Figure 6.16: Planck mask of the Galactic plane at the 78% sky coverage, from the 2018 release `http://pla.esac.esa.int/pla/aio/product-action?MAP.MAP_ID=COM_Mask_CMB-common-Mask-Int_2048_R3.00.fits`. The resolution of the map is $N_{\text{side}} = 512$.

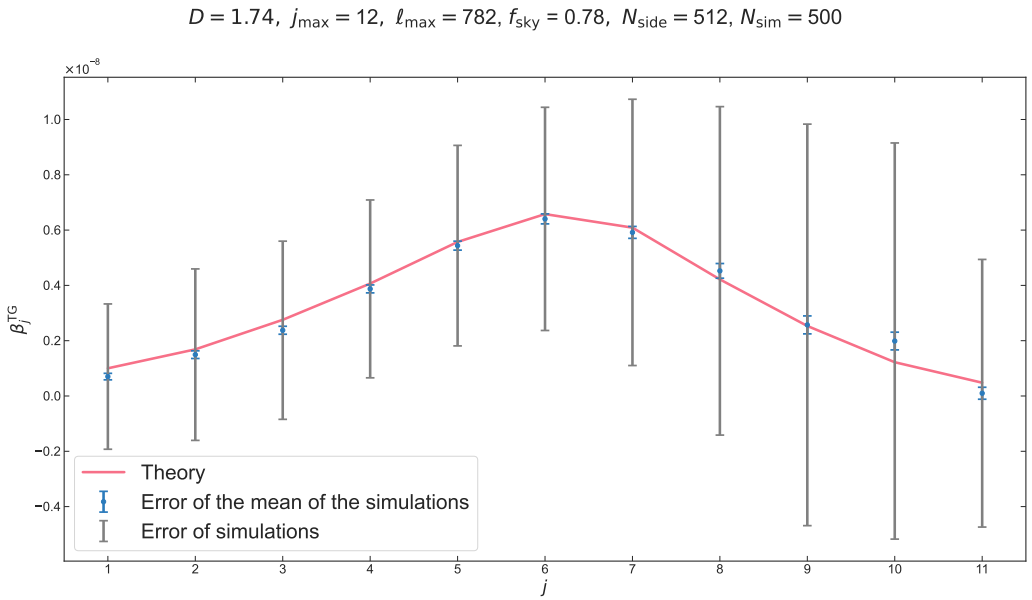


Figure 6.17: This figure displays both the theoretical and simulated needlet power spectra. The theoretical power spectrum is depicted by the pink line, while the blue line represents the mean of all simulations and their associated errors. The grey shading represents the errors derived from the simulations.

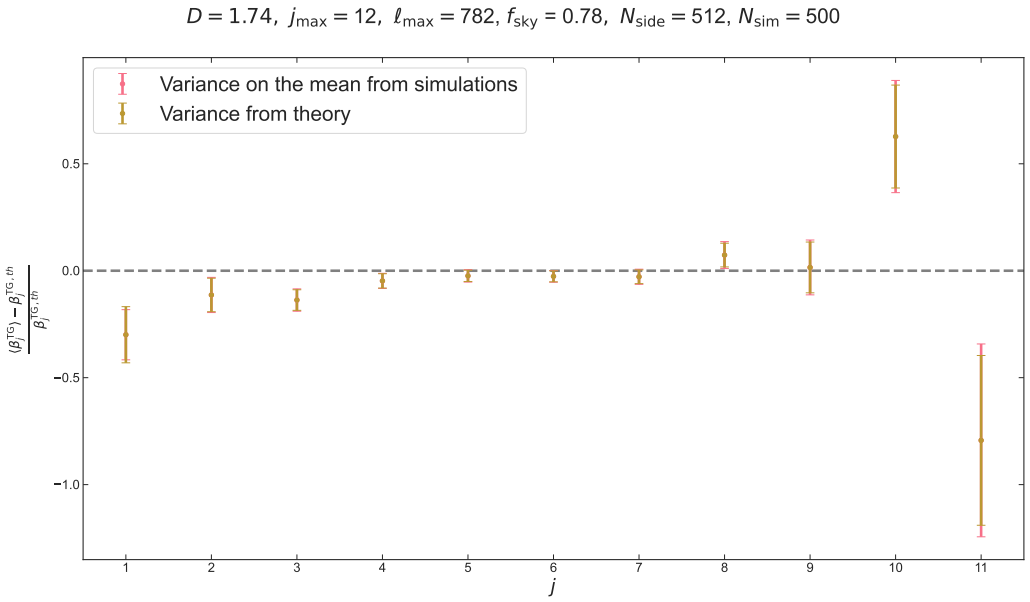


Figure 6.18: The relative difference between the mean of the simulated $\hat{\beta}_j^{\rm TG}$ and the theoretical value $\beta_j^{\rm TG}$. The pink error bars represent the errors on the mean of the simulations, calculated using the diagonal of the simulated covariance matrix (divided by $N_{\rm sim}$), while the pink bars include errors calculated from the theoretical variance in equation (6.16), accounting for the shot noise term.

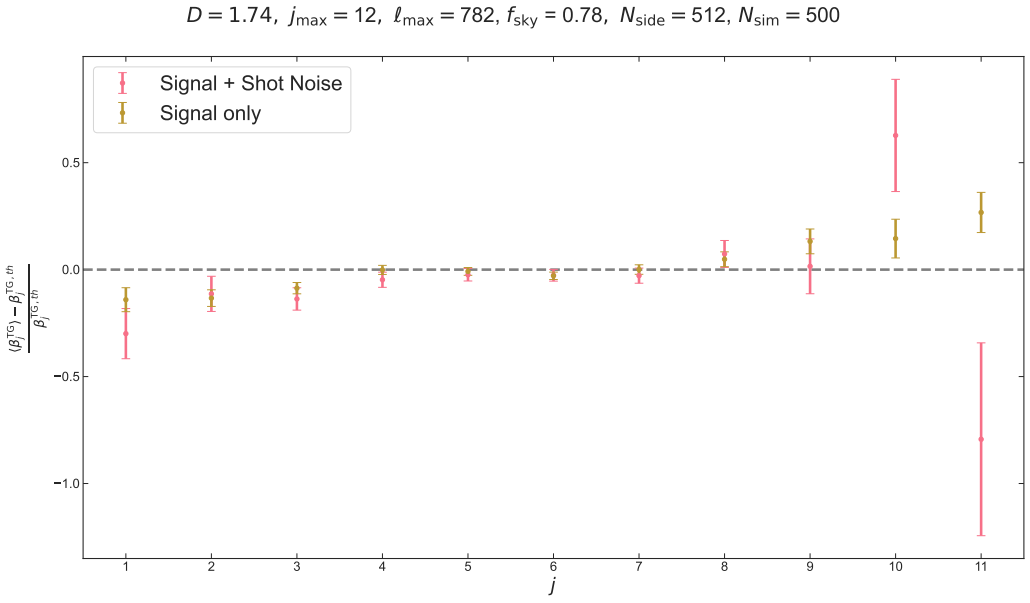


Figure 6.19: The relative difference between the mean of the simulated $\hat{\beta}_j^{\rm TG}$ and the theoretical value $\beta_j^{\rm TG}$ with and without the shot noise. The error bars are calculated from the covariance from the simulations.

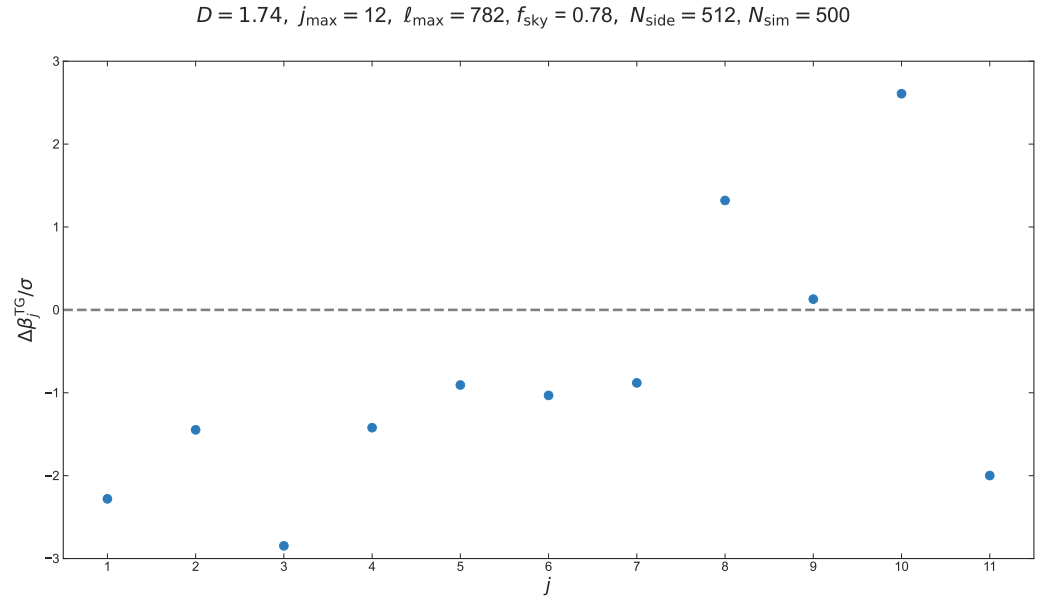


Figure 6.20: The difference between the mean of the simulated power spectra and the theoretical spectrum, divided by the theoretical variance $(\Delta\beta_j^{\mathrm{TG}})^2$.

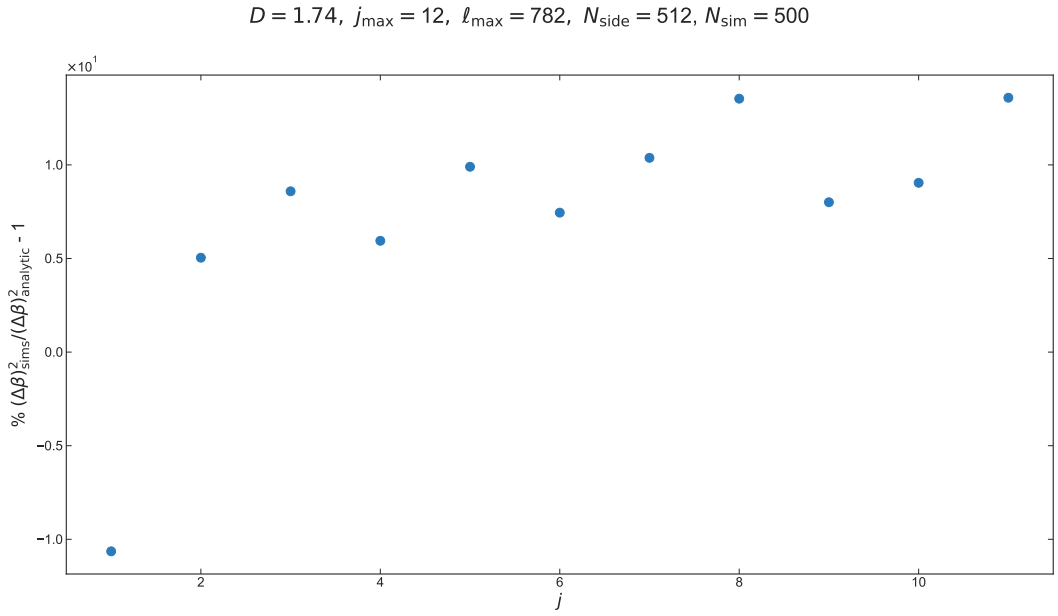


Figure 6.21: The relative difference between the variance computed from the simulations and the theoretical variance $(\Delta\beta_j^{\mathrm{TG}})^2$

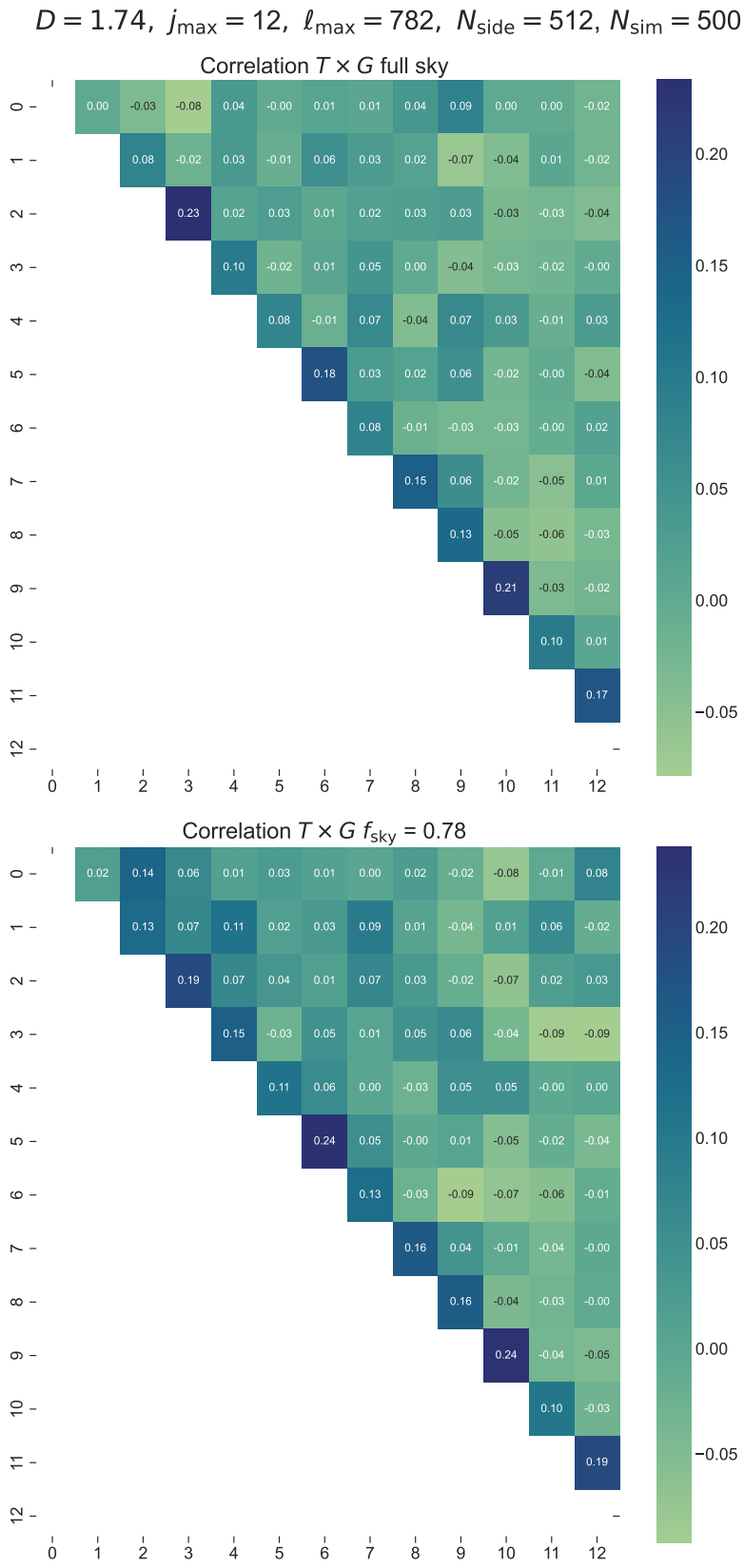


Figure 6.22: The covariance matrix and the correlations between different values of j, computed from the simulations of CMB and galaxies maps, as the second expression in eq. (6.23). The simulations include the shot noise. The matrix on the left is from full-sky maps, whereas on the right it is computed from masked-sky maps. The main diagonal is not displayed here.

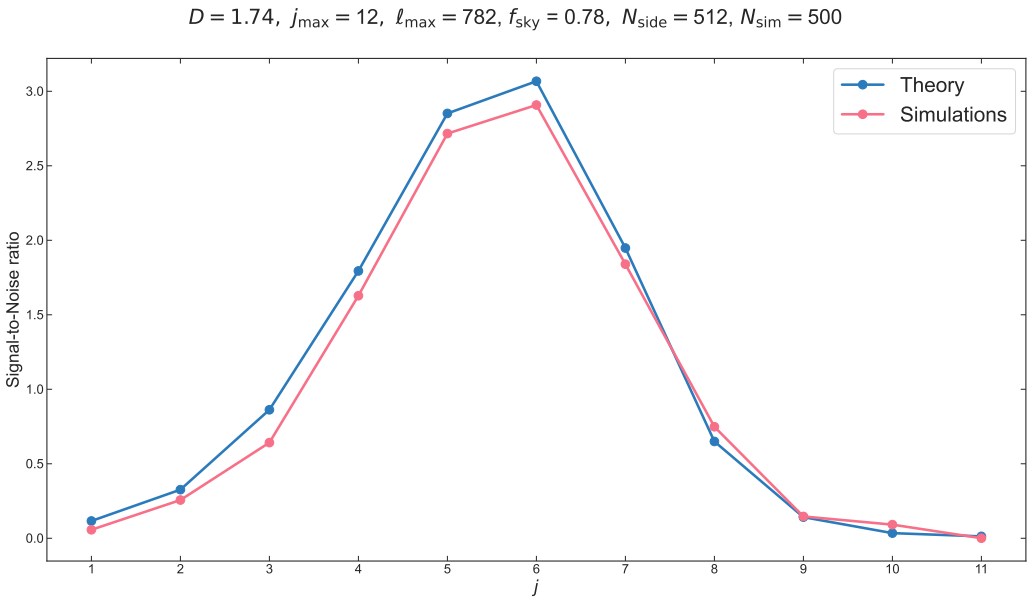


Figure 6.23: The Signal-to-Noise ratio (S/N) as computed from eq. (6.25), in the case of cut-sky and shot noise. The blue line is the expected SNR computed from the theoretical $\beta_j^{\rm TG}$; in pink is shown the SNR computed from the simulations.

maps of the correlated CMB and galaxy overdensity. They were generated following the *Euclid* forecast's prescription for the cosmological parameters and the photometric survey parameters, as detailed in tables 5.1 and 5.2. For our analysis, we consider the non-tomographic case (i.e., $n_{\rm bins} = 1$), but we intend to expand it to include $n_{\rm bins} = 10$ soon. The maps also include *Euclid*'s galaxy shot noise, which corresponds to $N^{\rm G} = 30$ galaxies per square arcminute.

We apply the combination of the Planck mask (as shown in Fig. 6.16) and the mask provided by the *Euclid* collaboration, which simulates the Euclid observational strategy with $f_{\rm sky} = 0.36$, to both the CMB and galaxy maps. Fig. 6.28 shows the *Euclid* mask in the upper panel and the combination with the Planck mask on the right. The total fraction of the observed sky is $f_{\rm sky} = 0.35$ in this analysis. The maps are analyzed up to $\ell_{\rm max} = 256$ and $j_{\rm max} = 12$, which fixes $D = 1.59$.

Initially, we compare the corrected estimated needlet spectra $\beta_j^{\rm TG} \approx \tilde{\beta}'^{\rm TG} j / f_{\rm sky}$ with the expected values $\beta_j^{\rm TG}$. However, we find that the estimates are biased. This bias is clearly visible in Fig. 6.29, which shows both the recovered power spectrum and the theoretical one. Additionally, Fig. 6.30 shows the relative difference between the recovered and theoretical power spectra.

We correct the bias by comparing the raw measurements, which are affected by the mask, with the $\Gamma_j'^{\rm TG}$ estimator described in eq. (6.19). As shown in

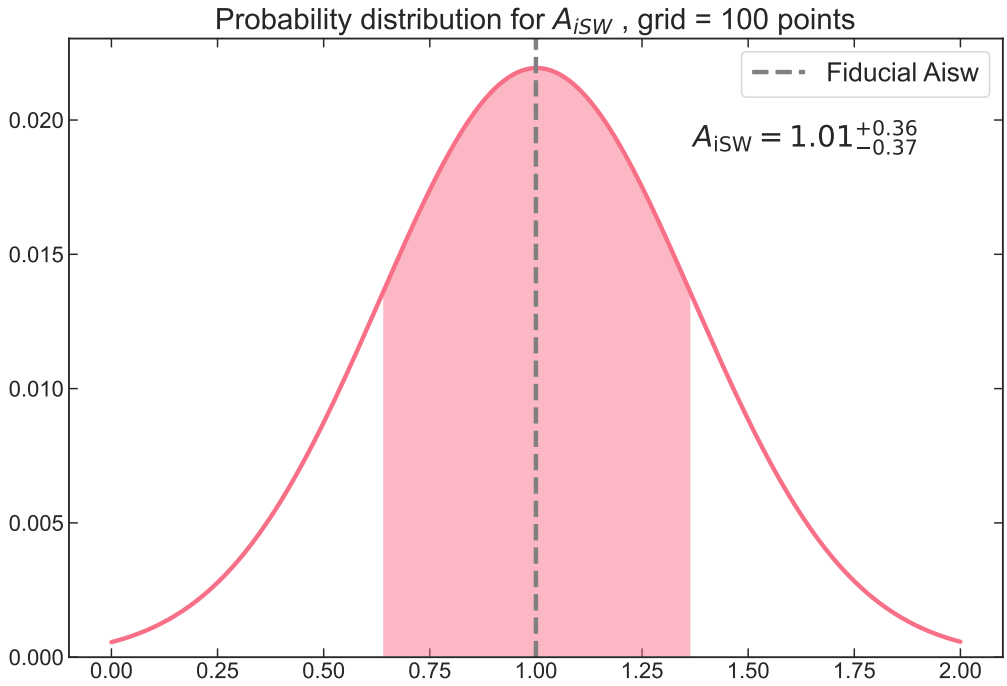


Figure 6.24: The probability distribution function for the parameter $A_{\rm iSW}$, drawn from a Gaussian likelihood computed over a grid of 30 values. The shaded area represents the 68% confidence level, while the dotted line represents the fiducial value for $A_{\rm iSW}$. In the figure are also written he best-fit value and the range of the CI. These results are derived from masked-sky simulations with shot noise included.

Fig. 6.31, this estimator is unbiased, and r the theoretical power spectrum is recovered from the simulations. The estimated power spectrum, $\hat{\Gamma}_j^{\rm TG}$, is computed as the mean from all the simulations and exhibits a difference from the theoretical power spectrum of approximately 1σ. This is shown in Fig. 6.32, where σ represents the theoretical variance as defined in eq. (6.20). It's worth noting that in this analysis with *Euclid* simulations, the number of simulations is larger, which results in increased statistical power compared to the Planck analysis.

We also validate the covariance associated with the pseudo-needlet estimator as defined in eq. (6.20) by comparing it to the simulated covariance, as shown in Fig. 6.33. The discrepancy between the simulated and analytical variances is approximately $40-30\%$ for the first two frequency bins, and then it decreases to around 10% for the other bins, as shown in Fig. 6.34. We are currently investigating the larger difference in the first two j's, and it is likely related to the morphology of the *Euclid* mask, which has sharp edges. Further testing is needed to determine the best way to model the mask in our theoretical prescription in order to reduce this error.

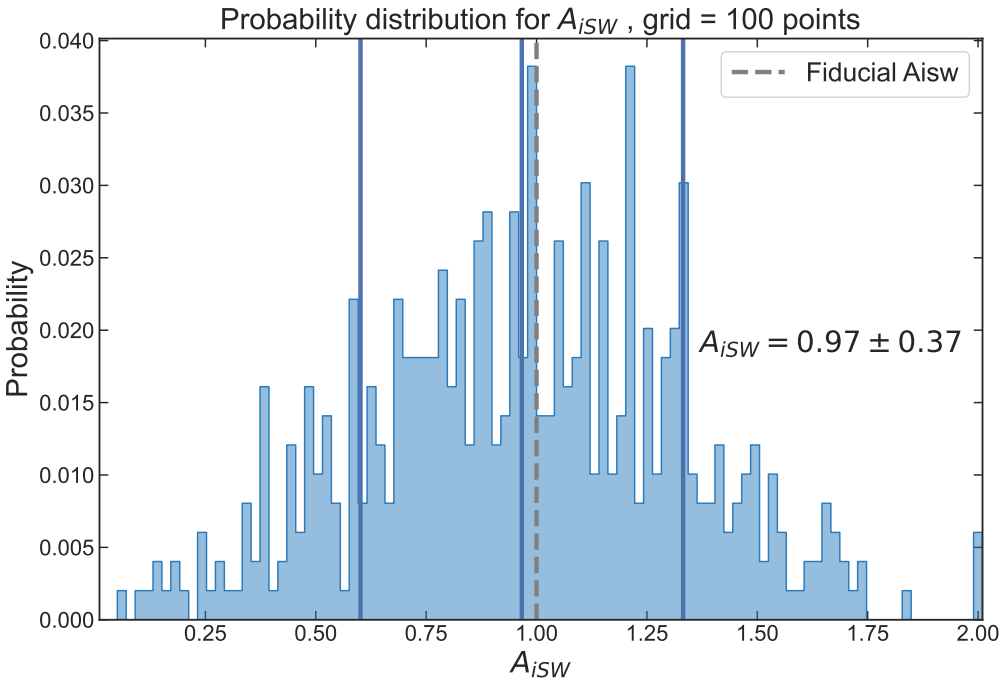


Figure 6.25: Histogram of the best-fit values of $A_{\rm iSW}$ computed from a Gaussian likelihood for each simulation. The dotted line is the fiducial value, while the blue lines correspond to the percentiles at the 68% and the mean of the best-fit values. Here we analyze masked-sky simulations with the shot noise.

The S/N ratio is displayed in Fig. 6.35, and it is compared to the theoretical S/N calculated from $\beta_j^{\rm TG}$ and the theoretical variance $(\Delta\beta_j^{\rm TG})^2$, with the shot noise term. The S/N obtained from the simulations is smaller due to the mask's effect.
We also compare this result with the cumulative S/N calculated from the estimated pseudo-$C_\ell^{\rm TG}$ power spectrum (see (4.27)). Fig. 6.36 shows the cumulative S/N for the needlets and the Pseudo-$C_\ell^{\rm TG}$. When using the Pseudo-$C_\ell^{\rm TG}$ estimator, the resulting S/N is slightly higher than that obtained with the pseudo-needlets. The values are $S/N = 3.73$ (corresponding to the cumulative S/N at $\ell_{\rm max} = 256$) for the Pseudo-$C_\ell^{\rm TG}$ and $S/N = 3.69$ for the needlets estimator, showing that they are comparable.

Parameter estimation As in the analysis of simulations with the Planck 2018 cosmology, We draw the probability distributions for the parameters $A_{\rm iSW}$ and Ω_Λ using the *Euclid* simulations. The distribution for $A_{\rm iSW}$ is centered around the fiducial value with approximately 30% uncertainty, as demonstrated by the results from all the simulations, which are shown in Fig. 6.38. Fig. 6.37 shows the distribution drawn from a single simulation,

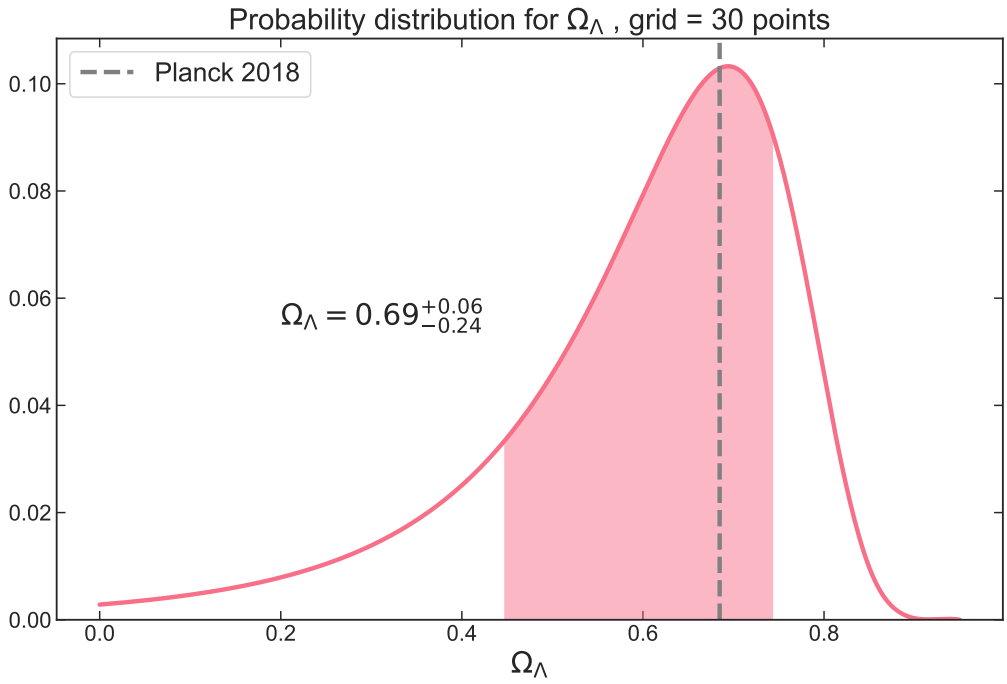


Figure 6.26: The probability distribution function of Ω_Λ drawn from a Gaussian likelihood computed over a grid of 30 needlet power spectrum. The covariance is derived from the simulations. The grey dotted line represents the Planck fiducial value $\Omega_\Lambda^{\text{Planck}} = 0.6847$, while the shaded area represents the 68% Confidence Interval. In the figure are also written he best-fit value and the range of the CI. The fiducial cosmology is based on the Planck 2018 parameters.

showing what we can expect to observe from one random realization of the observed sky. The results for Ω_Λ are shown in Fig. 6.40 and Fig. 6.39. The former exhibits the distribution of best-fit values derived from all the simulations, centered around the fiducial value with an error of approximately of 18%. The latter presents a distribution drawn from a single simulation, as an example of what one can expect to observe from one single realization of the observed sky. We find that the fiducial value of the parameter from the simulations is recovered. Although the combined *Euclid* and Planck mask is larger compared to the Planck mask used in simulations with the Planck cosmology, it is important to note that the *Euclid* shot noise is lower, and the number of simulations is larger. Consequently, this results in more tightly constrained parameter distributions.

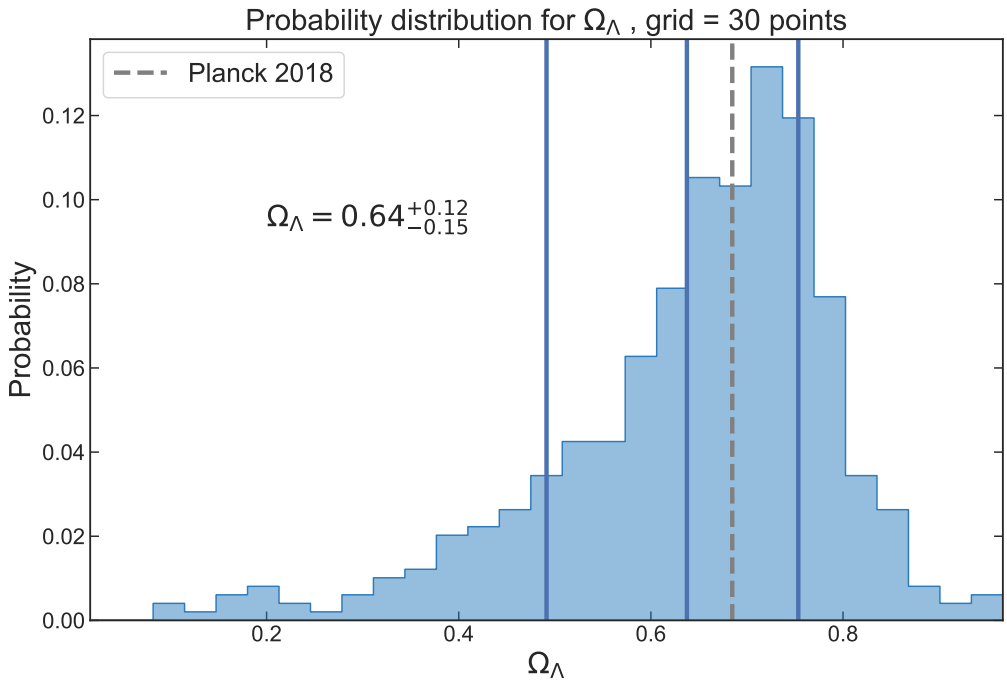
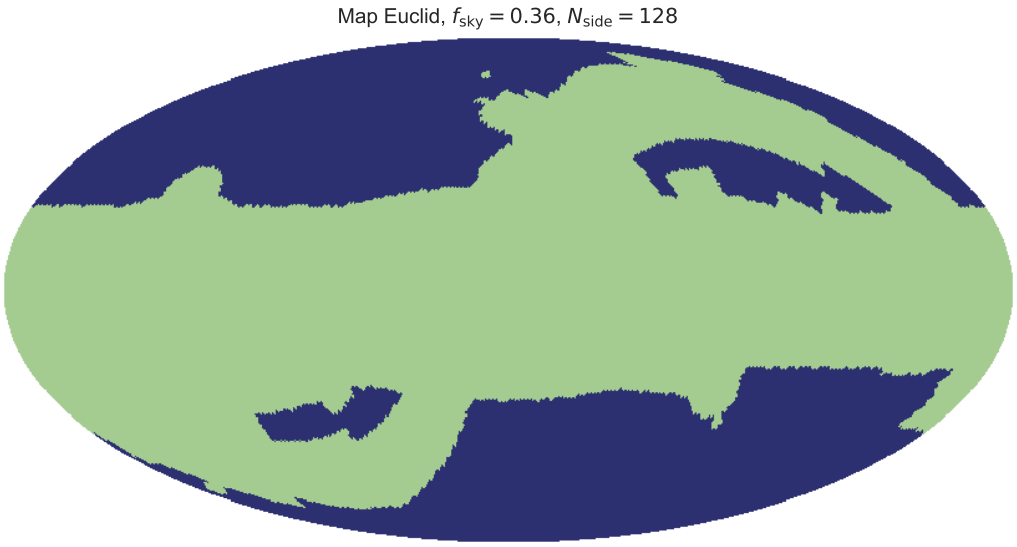
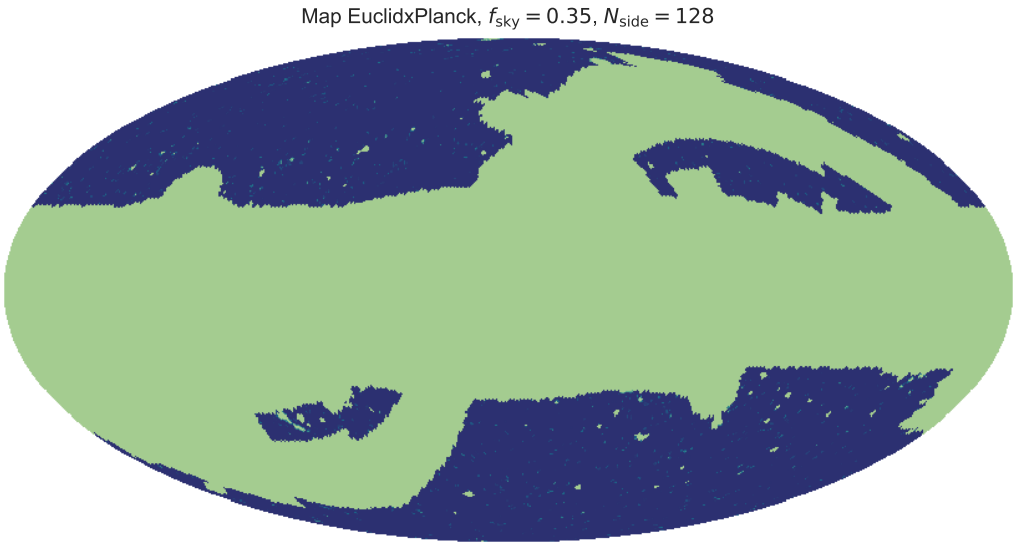


Figure 6.27: Distribution of the best-fit values of Ω_Λ computed from the Gaussian likelihood for each simulation. The blue lines represent the mean of the best-fit values and the percentiles of the distributions at the 68% confidence level. we observe that the fiducial value falls within the 1σ interval. The fiducial cosmology is based on the Planck 2018 parameters.



(a) *Euclid* mask map with $f_{\rm sky} = 0.36$.



(b) Map of the combination of the Planck and *Euclid* masks, with $f_{\rm sky} = 0.35$.

Figure 6.28: Maps of the *Euclid* mask, on the upper panel, and of the combination with the Planck mask (same as Fig. 6.16), on the lower panel. For the analysis, we apply the combined mask to both the CMB and galaxy field.

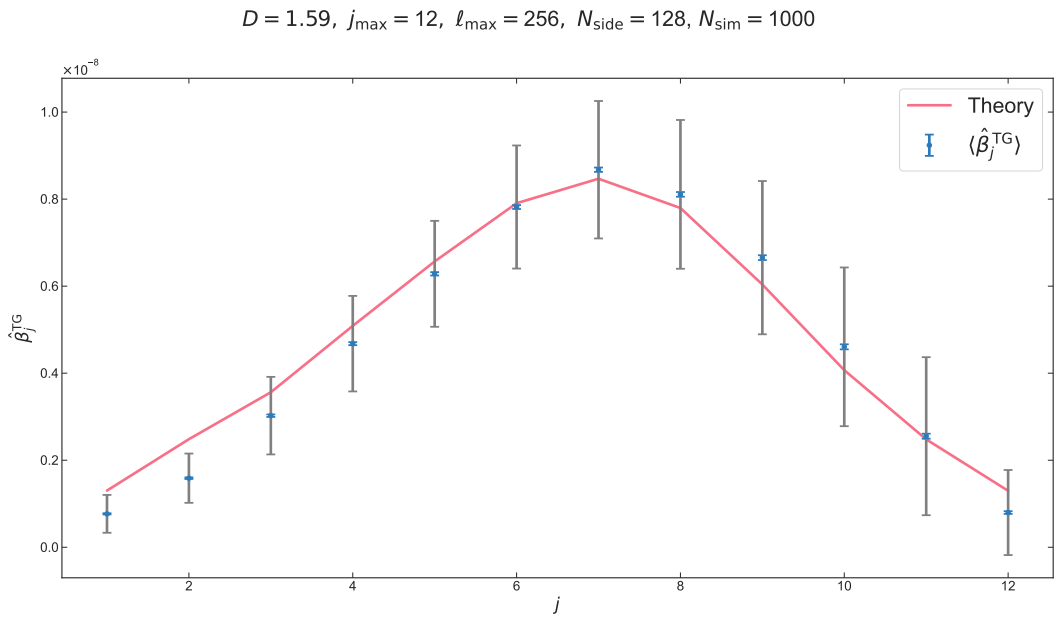


Figure 6.29: The theoretical needlet power spectrum (the blue error bars represent the errors of the mean, while the grey error bars correspond to the errors of individual simulations. It becomes evident that the $f_{\rm sky}$ approximation is inadequate for the analyzed sky fraction, resulting in a bias in the power spectrum estimate.

[!h]

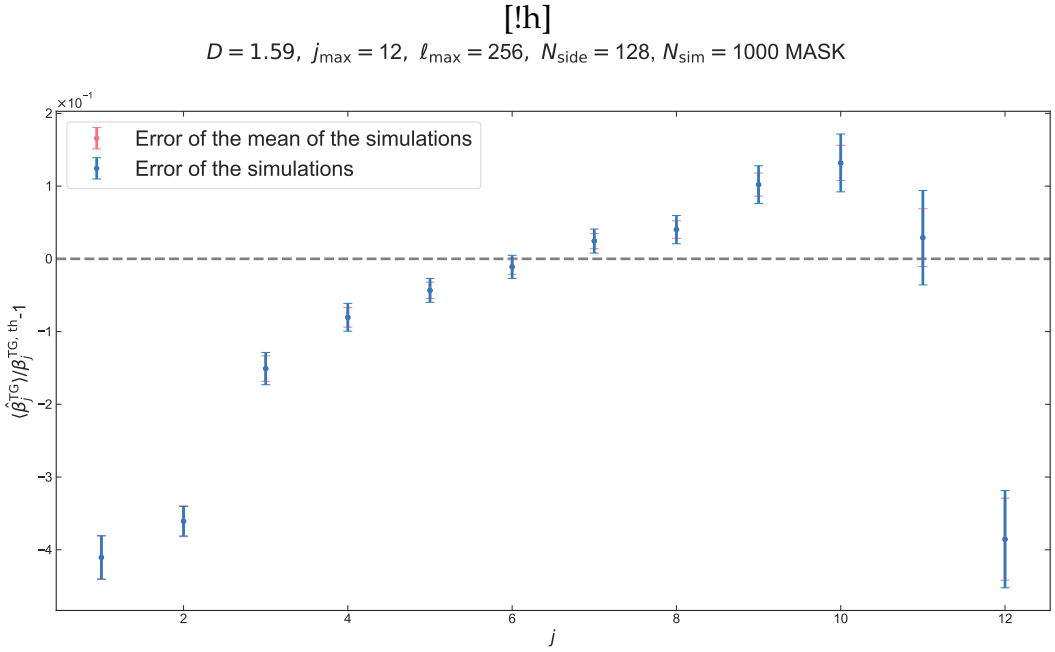


Figure 6.30: The relative difference between the theoretical and the recovered needlet power spectrum. The pink error bars represent the theoretical variance, while the blue error bars indicate the error of the mean estimated power spectrum from the simulations. It is evident from this plot that the estimator is biased for this particular sky fraction value.

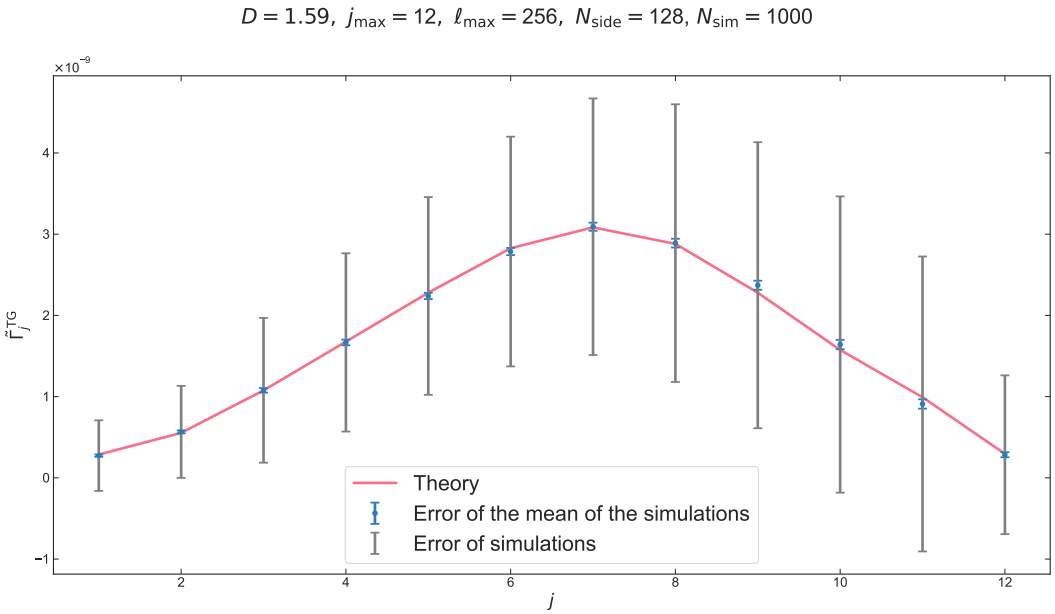


Figure 6.31: Same as Fig.6.29, but using the pseudo-needlet power spectrum, $\tilde{\Gamma}_j^{TG}$, in pink. we find that this estimator is unbiased.

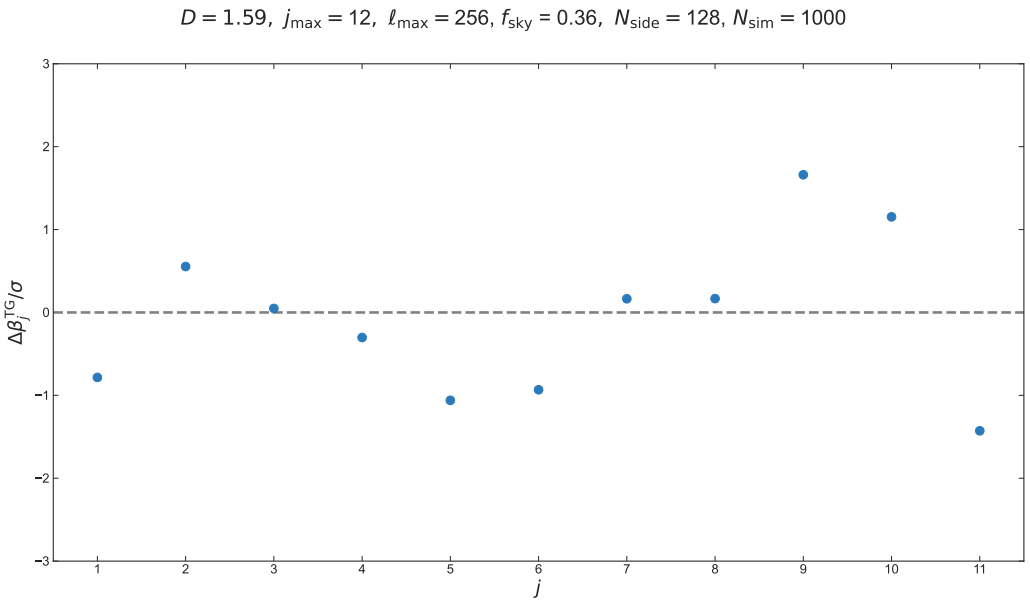


Figure 6.32: The difference between the mean of the simulated power spectra and biased expected spectrum $\tilde{\Gamma}_j^{TG}$, divided by the theoretical variance $(\Delta\Gamma_j^{TG})^2$.

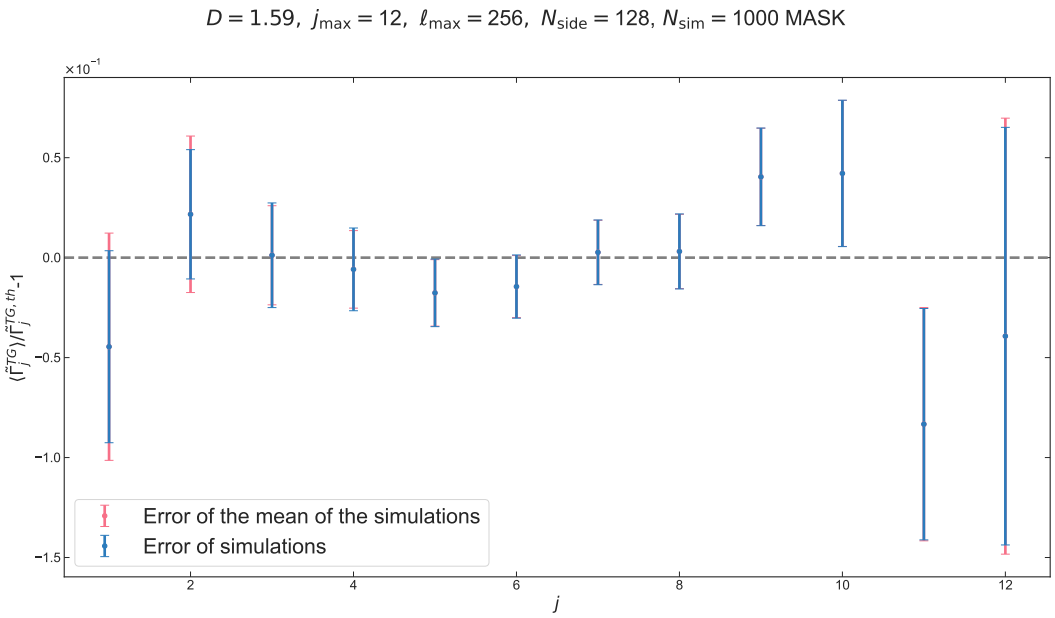


Figure 6.33: Same as Fig 6.30 but with the pseudo-needlet power spectrum. we find that the estimate is unbiased and that the theoretical variance is consistent with the simulated one.

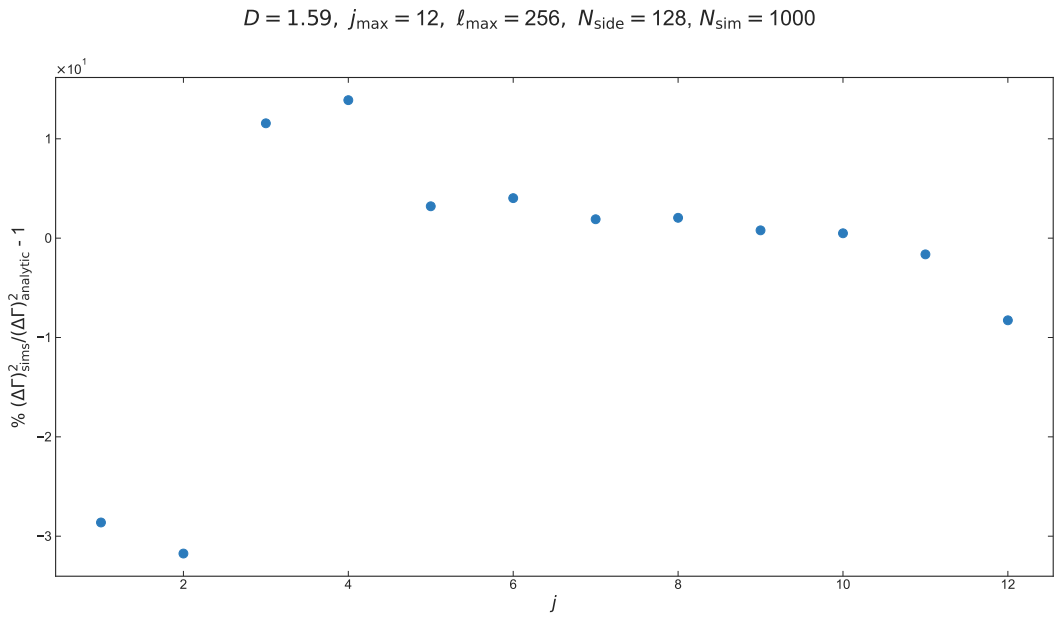


Figure 6.34: The relative difference between the variance computed from the simulations and the theoretical variance $\Delta\Gamma_j^{TG})^2$

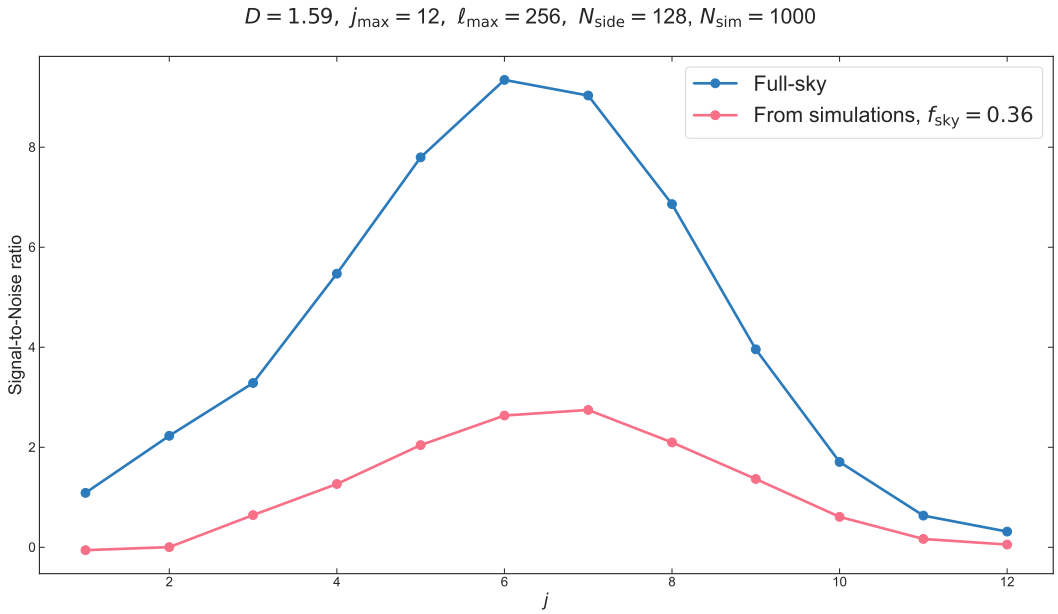


Figure 6.35: The S/N from the *Euclid* simulations. The blue line represents the expected ratio, calculated using equation (6.25), which includes the theoretical β_j^{TG} in the case of full sky and the theoretical variance $(\Delta\beta_j^{TG})^2$, accounting for shot noise. The S/N obtained from the simulations is shown in pink, and it is smaller due to the mask.

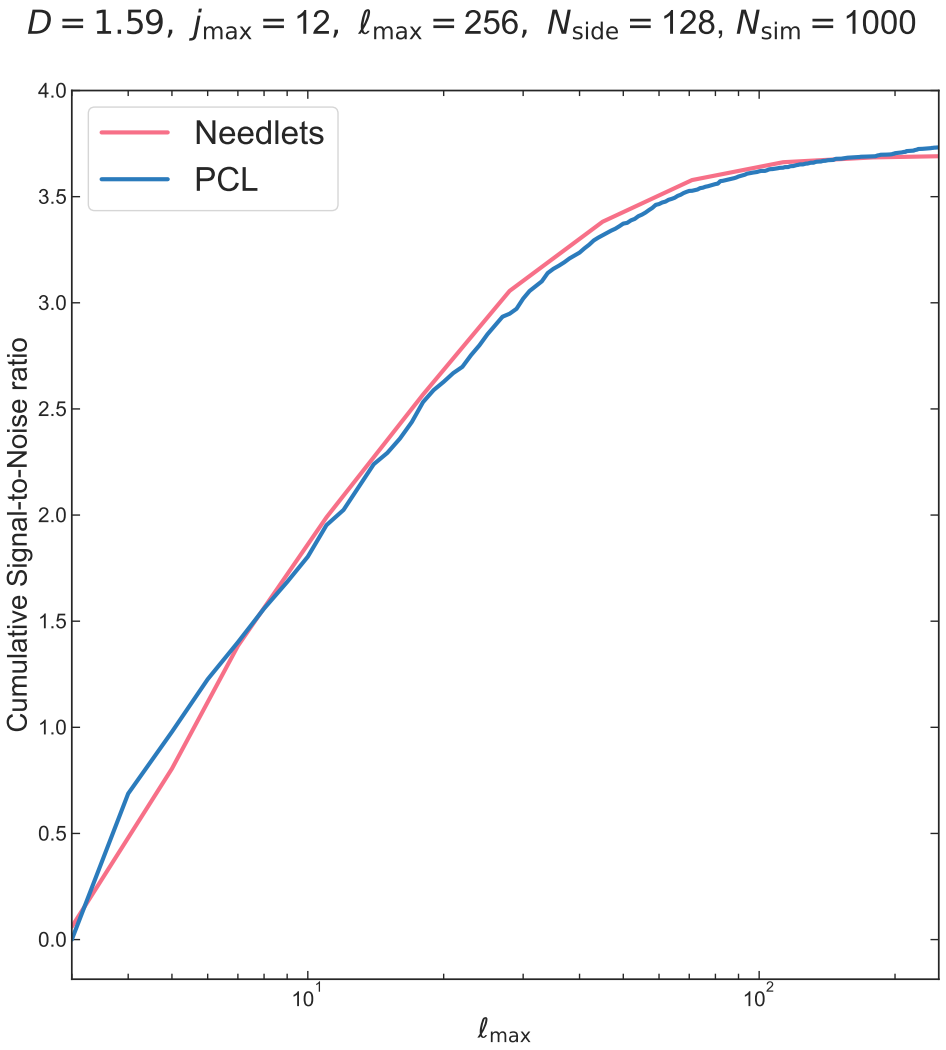


Figure 6.36: The cumulative SNR from the *Euclid* simulations.The blue line represents the cumulative SNR estimated using the Pseudo-$C_\ell^{\rm TG}$ estimator (as defined in equation (4.27)), while the pink line represents the S/N estimated with the pseudo-needlets estimators. These two lines are comparable, although the Pseudo-$C_\ell^{\rm TG}$ estimator can recover a slightly higher value, approximately $S/N \sim 3.73$, compared to the pseudo-needlets, which yield $S/N = 3.69$.

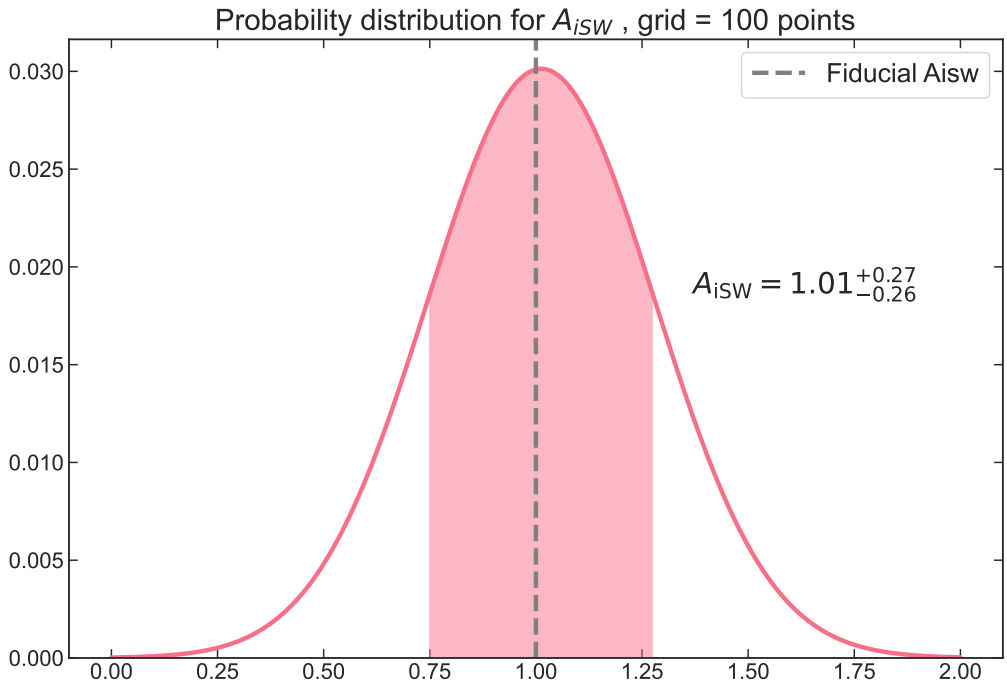


Figure 6.37: The probability distribution function for the parameter $A_{\rm iSW}$, drawn from a Gaussian likelihood computed over a grid of 30 values. The cosmological parameters are fixed according to the *Euclid* forecast prescription in table 5.1. The shaded area corresponds to the 68% CI, while the dotted line represents the fiducial value for $A_{\rm iSW}$. In the plot are also written the best fit values and the range of the CI. These results are based on the analysis of simulations for the *Euclid* survey, with the applied mask shown in Fig. 6.28b.

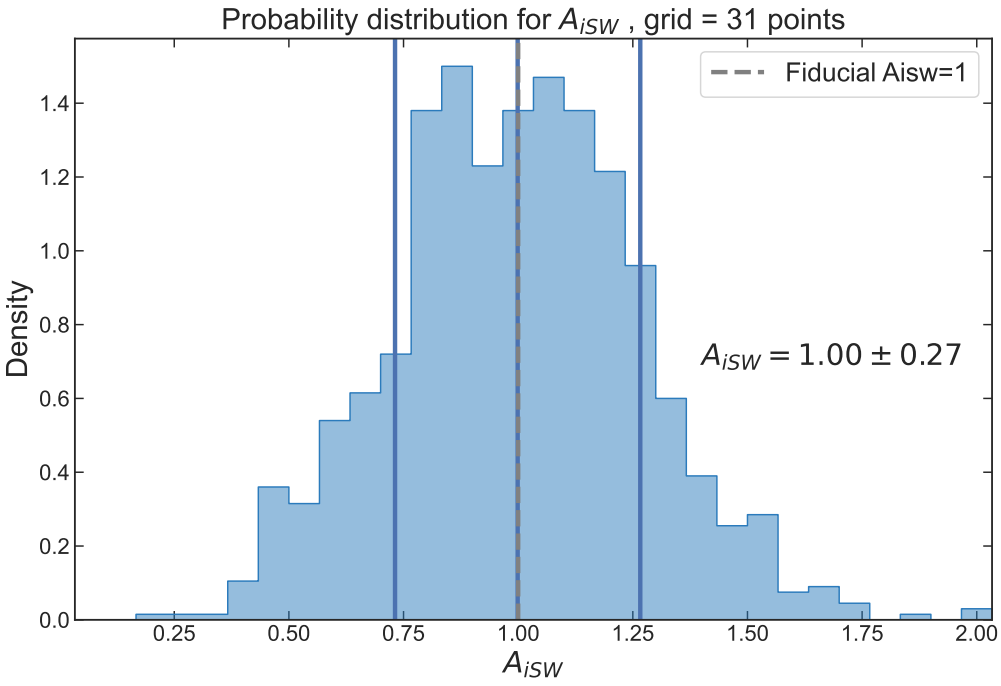


Figure 6.38: Histogram of the best-fit values of $A_{\rm iSW}$ computed from a Gaussian likelihood for each simulation. The fiducial value is indicated by the dotted line, while the blue lines represent the percentiles at 68% and the mean of the best-fit values. These results are based on the analysis of simulations for the *Euclid* survey, with the mask depicted in Fig. 6.28b, while the cosmological parameters are fixed to the *Euclid* forecast prescription outlined in table 5.1.

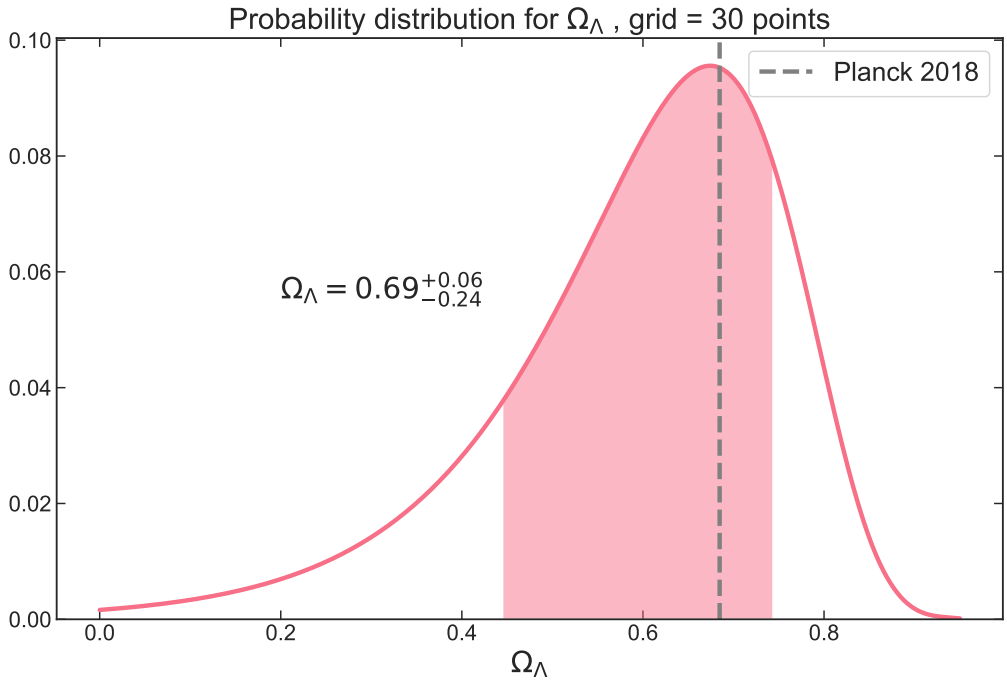


Figure 6.39: The probability distribution function of Ω_Λ drawn from a Gaussian likelihood computed over a grid of 30 needlet power spectrum. The covariance is determined from the simulations. The grey dotted line represents the fiducial value $\Omega_\Lambda^{\text{fiducial}} = 0.68$, while the shaded area corresponds to the 68% CI. In the plots are also written the best-fit value and the range of the CI. These results are based on the cosmological parameters specified in the *Euclid* forecast outlined in table 5.1. These results are based on the analysis of simulations for the *Euclid* survey, with the mask applied as shown in Fig. 6.28b.

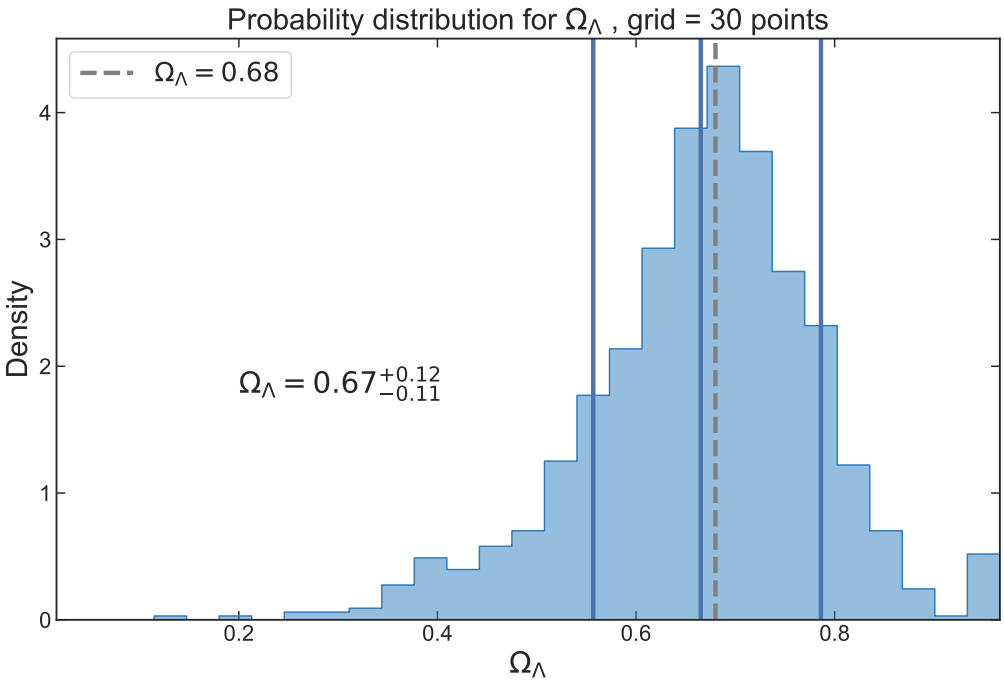


Figure 6.40: Distribution of the best-fit values of Ω_Λ computed from the Gaussian likelihood for each simulation. The blue lines represent the mean values and the percentiles at the 68% confidence level. Within this interval, we successfully recovered the fiducial value. The cosmological parameters used for this analysis are fixed to the *Euclid* forecast prescription listed in table 5.1. These results are based on the analysis of simulations for the *Euclid* survey with the mask shown in Fig. 6.28b.

Chapter 7

Conclusions

In recent decades, cosmology has made remarkable achievements thanks to observations. The ΛCDM seems to be the prominent model that describes the physics of the Universe, according to the data. Nevertheless, there are still a plethora of open questions to be answered and tensions between data to be solved: first of all, about the nature of the dark energy.
However, the future looks bright, thanks to the high-precision experiments that will take place in the next years. Novel methods to exploit the amount of data provided by the upcoming CMB and LSS surveys are under development and validation, in order to extract cosmological and astrophysical information. In this direction, one of the prominent analyses is the cross-correlation of different probes. In this thesis, I have investigated the advantages of the cross-correlation between CMB data and the distribution of galaxies provided by the Euclid survey in the next years.
In chapter 4 I have first discussed how the cross-correlation analysis could improve the detection of the iSW effect, reviewing the state-of-the-art and the benchmark that the Euclid survey aims to reach. I show that the tomography analysis improves the S/N ratio by a factor of 6.8%, as shown in Fig. 4.17, and that the uncertainties on the detection are dominated by the cosmic variance.
Chapter 5 showed the work developed by myself and the CMBXC SWG within the Euclid collaboration. We have tested and validated the Gaussian likelihood approximation for the cross-correlation between the distribution of galaxies and the CMB temperature that will be implemented in the official Euclid code CLOE. This approximation is based on the assumption that the "fiducial" model is smooth and closely aligned with the underlying physical model. In our case, the "fiducial" model is ΛCDM with the cosmological parameters fixed at the forecast prescriptions for *Euclid*. To validate this approximation, we have considered different cosmological models, ΛCDM and minimal extensions to it, and we have inferred the probability distribution for one or two cosmological parameters for each model. The analysis is limited to a short number of inferred parameters since we do not expect

high constraining power from the cross-correlation signal over all the cosmological parameters. In particular, we have tested the likelihood for the iSW signal amplitude ($A_{\rm iSW}$), fixing the cosmological parameters to ΛCDM; we have tested the ΛCDM inferring the joint probabilities of the primordial amplitude (A_s) and cold dark matter energy density ($\Omega_c h^2$). We analyzed the $w_0 w_a$CDM model, inferring the joint probabilities of the parameters for the dynamical dark energy equation of state (w_0 and w_a, see eq. (1.21)). Finally, we have tested the extended model w_0CDM model calculating the probability for w_0 alone.

To achieve this, we have varied the parameters of interest across a grid of values (while keeping others fixed at fiducial values), computed a smooth angular power spectrum at each point, and then computed the likelihood. We performed similar analyses by changing the fiducial parameters of the *Euclid* photometric surveys, as detailed in Table 5.2, to assess how much the probability distributions of the parameters are affected by the features of the survey.

In our results, we have observed that the fiducial values were recovered as best-fit values for all parameters under study, except for w_a, which remains unconstrained by the cross-correlation signal. Additionally, we have found that the probability distributions are influenced by $\ell_{\rm min}$, the minimum multipole of the analysis. Since the iSW signal peaks at very large scales, it is highly sensitive to this parameter. Furthermore, the sky fraction observed during the survey ($f_{\rm sky}$) affects the probability distributions. Smaller values of $f_{\rm sky}$ reduced the iSW's constraining power due to information loss from the largest scales. Conversely, larger $f_{\rm sky}$ values tightened parameter constraints. Finally, tomography improved parameter constraints compared to non-tomographic analysis. Other photometric survey parameters such as photometric error and the fraction of catastrophic outliers have negligible impact on parameter distributions. Comprehensive results from this analysis are summarized in tables 5.5, 5.8, 5.11, and 5.14.

The shot noise's impact is negligible in all the analyses that we have performed, indicating that cosmic variance dominates uncertainties.

Additionally, we have tested the influence on the cosmological parameter probability distribution when computing the likelihood using a covariance evaluated at a "wrong" fiducial cosmology. In the frame of ΛCDM, we have tested the impact on the $\Omega_c h^2$ distribution when computing the covariance at a "wrong" value of $\Omega_c h^2$, leaving the other parameters untouched. We found that the best-fit values are the same as it is computed from the covariance at the "exact" fiducials, both in the tomographic and in the non-tomographic analysis. The impact on the CI is of the order of a few percent. We repeat the test for the w_0CDM model, computing the distribution for w_0. The result was that the constraints on w_0 are unaffected by the "wrong" fiducials in the covariance, both in the tomographic and non-tomographic analysis. Thus, this test showed that computing the covariance matrix at a "wrong" fiducial has a negligible effect on the distribution of the parameters.

We have tested the impact of the Limber approximation using a χ^2 test. This analysis demonstrated that the error introduced by the Limber approximation is negligible, given the high level of uncertainty from cosmic variance.
Chapter 6 is dedicated to introducing and validating a promising estimator for the CMB-GC cross-correlation observable. This estimator is based on a specific type of spherical wavelets known as needlets, which possess unique properties suitable for cosmological data analysis. As discussed in this chapter, the localization and the correlation properties between different frequency bins make the needlets a suitable tool for analyzing signals that are affected by unknown systematics.
As described in this thesis, during my PhD I have built and validated this estimator for the cross-correlated observable using simulated temperature and sky maps generated from theoretical angular power spectra. The assumed cosmological model is ΛCDM with cosmological parameters fixed at the Planck 2018 best-fit values.
We have estimated the needlets power spectrum from full-sky simulated maps and detected the iSW signal with a significance of 96.65%, recovering the expected power spectrum within a 2σ range of the theoretical variance. As a final step, we have derived the probability distributions for the $A_{\rm iSW}$ and Ω_Λ parameters. We have found that the fiducial values were recovered within a 1σ interval for both parameters.
We have repeated this validation for the masked sky with shot noise. We have obtained the theoretical power spectrum within 3σ of the theoretical variance.
We have confirmed that the presence of the mask did not affect correlation properties between different frequency bins in the full covariance, aligning with expectations from needlets' properties. Similar to the full-sky analysis, We have retrieved the probability distributions for the $A_{\rm iSW}$ and Ω_Λ parameters, recovering fiducial values within 1σ intervals.
Finally, we have tested the needlets estimator using simulated maps from the *Euclid* survey provided by the CMBxC SWG collaborators. These maps included the shot noise and the mask predicted for *Euclid*. We have found that the mean of the estimated power spectra from the maps is in agreement with the theoretical expectation value of the needlets power spectrum. We have found a discrepancy at the very large scales between the variance calculated from the simulations and the analytical prescription that I implemented. This is probably caused by the fact that the morphology of the implemented mask is complex and it might require a different modelling. This work is currently under further analysis. We have compared the S/N computed by means of the needlets estimator with the one computed by mean of the harmonic estimator. We have found the two ratios in agreement and compatible with the theoretical prediction of the S/N discussed in section 4.5.
To conclude the validation, the needlets estimator effectively recovered fiducial values for the $A_{\rm iSW}$ and Ω_Λ parameters within a 1σ interval from

probability distributions drawn from the simulations. Further refinements and validation efforts are ongoing to address the variance discrepancy and improve accuracy in future analyses. As the next step, I am working on the implementation of the tomographic analysis with the needlets estimator.

To conclude, the aim of this thesis was to test the robustness of the analysis of the cross-correlation signal between the large-scale CMB temperature anisotropies and the *Euclid*'s galaxy distribution in preparation of the first data. On one hand, the likelihood function that we have implemented is validated and ready to be implemented in the official *Euclid* likelihood code for cosmological parameter estimations. On the other hand, the needlets estimator can provide complementary information with respect to the harmonic estimator, especially when applied to the analysis of real data with possible uncorrected systematic effects.

Bibliography

[1] D. Pietrobon, A. Balbi, and D. Marinucci, "Integrated Sachs-Wolfe effect from the cross-correlation of WMAP 3 year and NVSS: new results and constraints on dark energy", *Phys. Rev. D*, vol. 74, p. 043 524, 2006. DOI: 10.1103/PhysRevD.74.043524. arXiv: astro-ph/0606475.

[2] F. Bianchini, A. Renzi, and D. Marinucci, "Needlet estimation of cross-correlation between CMB lensing maps and LSS", *JCAP*, vol. 11, p. 050, 2016. DOI: 10.1088/1475-7516/2016/11/050. arXiv: 1607.05223 [astro-ph.CO].

[3] S. Weinberg, *Cosmology*. 2008, ISBN: 978-0-19-852682-7.

[4] R. Durrer, *The Cosmic Microwave Background*. Cambridge: Cambridge University Press, 2008, ISBN: 978-0-511-81720-5. DOI: 10.1017/CBO9780511817205.

[5] N. Aghanim *et al.*, "Planck 2018 results. VI. Cosmological parameters", *Astron. Astrophys.*, vol. 641, A6, 2020, [Erratum: Astron.Astrophys. 652, C4 (2021)]. DOI: 10.1051/0004-6361/201833910. arXiv: 1807.06209 [astro-ph.CO].

[6] E. Hubble, "A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae", *Proceedings of the National Academy of Science*, vol. 15, no. 3, pp. 168–173, Mar. 1929. DOI: 10.1073/pnas.15.3.168.

[7] A. Goobar, S. Perlmutter, G. Goldhaber, R. Knop, P. Nugent, P. Castro, S. Deustua, S. Fabbro, D. Groom, I. Hook, *et al.*, "The acceleration of the universe: Measurements of cosmological parameters from type ia supernovae", *Physica Scripta*, vol. 2000, no. T85, p. 47, 2000.

[8] E. V. Linder, "Exploring the expansion history of the universe", *Phys. Rev. Lett.*, vol. 90, p. 091 301, 9 2003. DOI: 10.1103/PhysRevLett.90.091301. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.90.091301.

[9] M. Chevallier and D. Polarski, "Accelerating universes with scaling dark matter", *Int. J. Mod. Phys. D*, vol. 10, pp. 213–224, 2001. DOI: 10.1142/S0218271801000822. arXiv: gr-qc/0009008.

[10] J. M. Bardeen, "Gauge-invariant cosmological perturbations", *Phys. Rev. D*, vol. 22, pp. 1882–1905, 8 1980. DOI: 10.1103/PhysRevD.22.1882. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.22.1882.

[11] C.-P. Ma and E. Bertschinger, "Cosmological perturbation theory in the synchronous and conformal Newtonian gauges", *Astrophys. J.*, vol. 455, pp. 7–25, 1995. DOI: 10.1086/176550. arXiv: astro-ph/9506072.

[12] A. H. Guth, "Inflationary universe: A possible solution to the horizon and flatness problems", *Phys. Rev. D*, vol. 23, pp. 347–356, 2 1981. DOI: 10.1103/PhysRevD.23.347. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.23.347.

[13] W. H. Kinney, A. Melchiorri, and A. Riotto, "New constraints on inflation from the cosmic microwave background", *Phys. Rev. D*, vol. 63, p. 023 505, 2001. DOI: 10.1103/PhysRevD.63.023505. arXiv: astro-ph/0007375.

[14] D. H. Lyth, "What would we learn by detecting a gravitational wave signal in the cosmic microwave background anisotropy?", *Phys. Rev. Lett.*, vol. 78, pp. 1861–1863, 1997. DOI: 10.1103/PhysRevLett.78.1861. arXiv: hep-ph/9606387.

[15] Dodelson, Scott and Schmidt, Fabian, *Modern Cosmology*. Academic Press, Elsevier Science, 2020, ISBN: 978-0-12-815948-4. DOI: 10.1016/C2017-0-01943-2.

[16] A. Lewis, A. Challinor, and A. Lasenby, "Efficient computation of cosmic microwave background anisotropies in closed friedmann-robertson-walker models", *The Astrophysical Journal*, vol. 538, no. 2, p. 473, 2000. DOI: 10.1086/309179. [Online]. Available: https://dx.doi.org/10.1086/309179.

[17] J. Lesgourgues, "The cosmic linear anisotropy solving system (class) i: Overview", *arXiv preprint arXiv:1104.2932*, 2011.

[18] W. T. Hu, "Wandering in the Background: A CMB Explorer", Aug. 1995. arXiv: astro-ph/9508126.

[19] W. Hu and N. Sugiyama, "Small scale cosmological perturbations: An Analytic approach", *Astrophys. J.*, vol. 471, pp. 542–570, 1996. DOI: 10.1086/177989. arXiv: astro-ph/9510117.

[20] J. Silk, "Cosmic black body radiation and galaxy formation", *Astrophys. J.*, vol. 151, pp. 459–471, 1968. DOI: 10.1086/149449.

[21] D. J. Fixsen, "The Temperature of the Cosmic Microwave Background", *Astrophys. J.*, vol. 707, no. 2, pp. 916–920, 2009. DOI: 10.1088/0004-637X/707/2/916.

[22] M. Gerbino, M. Lattanzi, M. Migliaccio, L. Pagano, L. Salvati, L. Colombo, A. Gruppuso, P. Natoli, and G. Polenta, "Likelihood methods for CMB experiments", *Front. in Phys.*, vol. 8, p. 15, 2020. DOI: 10.3389/fphy.2020.00015. arXiv: 1909.09375 [astro-ph.CO].

[23] U. Seljak and M. Zaldarriaga, "A Line of sight integration approach to cosmic microwave background anisotropies", *Astrophys. J.*, vol. 469, pp. 437–444, 1996. DOI: 10.1086/177793. arXiv: astro-ph/9603033.

[24] N. Aghanim, S. Majumdar, and J. Silk, "Secondary anisotropies of the cmb", *Reports on Progress in Physics*, vol. 71, no. 6, p. 066 902, 2008. DOI: 10.1088/0034-4885/71/6/066902. [Online]. Available: https://dx.doi.org/10.1088/0034-4885/71/6/066902.

[25] N. Kaiser, "Clustering in real space and in redshift space", *Mon. Not. Roy. Astron. Soc.*, vol. 227, pp. 1–27, 1987.

[26] A. Almeida *et al.*, "The Eighteenth Data Release of the Sloan Digital Sky Surveys: Targeting and First Spectra from SDSS-V", *The Astrophysical Journal*, vol. 267, no. 2, 44, p. 44, 2023. DOI: 10.3847/1538-4365/acda98. arXiv: 2301.07688 [astro-ph.GA].

[27] E. S. Rykoff *et al.*, "The Dark Energy Survey Six-Year Calibration Star Catalog", *arXiv e-prints*, arXiv:2305.01695, arXiv:2305.01695, 2023. DOI: 10.48550/arXiv.2305.01695. arXiv: 2305.01695 [astro-ph.IM].

[28] A. Lue, R. Scoccimarro, and G. Starkman, "Differentiating between modified gravity and dark energy", *Phys. Rev. D*, vol. 69, p. 044 005, 2004. DOI: 10.1103/PhysRevD.69.044005. arXiv: astro-ph/0307034.

[29] L. Guzzo *et al.*, "A test of the nature of cosmic acceleration using galaxy redshift distortions", *Nature*, vol. 451, pp. 541–545, 2008. DOI: 10.1038/nature06555. arXiv: 0802.1944 [astro-ph].

[30] R. Laureijs *et al.*, "Euclid Definition Study Report", *arXiv e-prints*, arXiv:1110.3193, arXiv:1110.3193, 2011. DOI: 10.48550/arXiv.1110.3193. arXiv: 1110.3193 [astro-ph.CO].

[31] W. Hu, "Power spectrum tomography with weak lensing", *Astrophys. J. Lett.*, vol. 522, pp. L21–L24, 1999. DOI: 10.1086/312210. arXiv: astro-ph/9904153.

[32] Dark Energy Survey Collaboration, "The Dark Energy Survey: more than dark energy - an overview", *Monthly Notices of the Royal Astronomical Society*, vol. 460, no. 2, pp. 1270–1299, 2016. DOI: 10.1093/mnras/stw641. arXiv: 1601.00329 [astro-ph.CO].

[33] Ž. Ivezić *et al.*, "LSST: From Science Drivers to Reference Design and Anticipated Data Products", *The Astrophical Journey*, vol. 873, no. 2, 111, p. 111, 2019. DOI: 10.3847/1538-4357/ab042c. arXiv: 0805.2366 [astro-ph].

[34] O. Ilbert *et al.*, "Euclid preparation: XI. Mean redshift determination from galaxy redshift probabilities for cosmic shear tomography", *Astron. Astrophys.*, vol. 647, A117, 2021. DOI: 10.1051/0004-6361/202040237. arXiv: 2101.02228 [astro-ph.CO].

[35] A. Troja, I. Tutusaus, and J. G. Sorce, "Euclid in a nutshell", *PoS*, vol. ICHEP2022, p. 094, Nov. 2022. DOI: 10.22323/1.414.0094. arXiv: 2211.09668 [astro-ph.IM].

[36] R. Scaramella *et al.*, "Euclid preparation - I. The Euclid Wide Survey", *Astron. Astrophys.*, vol. 662, A112, 2022. DOI: 10.1051/0004-6361/202141938. arXiv: 2108.01201 [astro-ph.CO].

[37] M. Cropper *et al.*, "Vis: The visible imager for euclid", in *Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave*, SPIE, vol. 9904, 2016, pp. 269–284.

[38] K. Bogna *et al.*, "A new signal estimator from the nir detectors of the euclid mission", *Publications of the Astronomical Society of the Pacific*, vol. 128, no. 968, p. 104 504, 2016. DOI: 10.1088/1538-3873/128/968/104504. [Online]. Available: https://dx.doi.org/10.1088/1538-3873/128/968/104504.

[39] J. W. Beletic *et al.*, "Teledyne Imaging Sensors: infrared imaging technologies for astronomy and civil space", in *High Energy, Optical, and Infrared Detectors for Astronomy III*, D. A. Dorn and A. D. Holland, Eds., International Society for Optics and Photonics, vol. 7021, SPIE, 2008, 70210H. DOI: 10.1117/12.790382. [Online]. Available: https://doi.org/10.1117/12.790382.

[40] T. Maciaszek *et al.*, "Euclid near infrared spectrometer and photometer instrument concept and first test results obtained for different breadboards models at the end of phase c", in *Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave*, SPIE, vol. 9904, 2016, pp. 298–315.

[41] P. Monaco, *Mapping the universe with slitless spectroscopy*, 2023. arXiv: 2301.05669 [astro-ph.CO].

[42] E. Ayçoberry, V. Ajani, A. Guinot, M. Kilbinger, V. Pettorino, S. Farrens, J.-L. Starck, R. Gavazzi, and M. J. Hudson, "UNIONS: The impact of systematic errors on weak-lensing peak counts", *Astronomy & Astrophysics*, vol. 671, A17, A17, 2023. DOI: 10.1051/0004-6361/202243899. arXiv: 2204.06280 [astro-ph.CO].

[43] M. Salvato, O. Ilbert, and B. Hoyle, "The many flavours of photometric redshifts", *Nature Astronomy*, vol. 3, pp. 212–222, Jun. 2019. DOI: 10.1038/s41550-018-0478-0. arXiv: 1805.12574 [astro-ph.GA].

[44] A. Blanchard *et al.*, "Euclid preparation: VII. Forecast validation for Euclid cosmological probes", *Astron. Astrophys.*, vol. 642, A191, 2020. DOI: 10.1051/0004-6361/202038071. arXiv: 1910.09273 [astro-ph.CO].

[45] A. Pocino *et al.*, "Euclid preparation - XII. Optimizing the photometric sample of the Euclid survey for galaxy clustering and galaxy-galaxy lensing analyses", *Astron. Astrophys.*, vol. 655, A44, 2021. DOI: 10.1051/0004-6361/202141061. arXiv: 2104.05698 [astro-ph.CO].

[46] A. Nicola, A. Refregier, and A. Amara, "Integrated approach to cosmology: Combining CMB, large-scale structure and weak lensing", *Phys. Rev. D*, vol. 94, no. 8, p. 083 517, 2016. DOI: 10.1103/PhysRevD.94.083517. arXiv: 1607.01014 [astro-ph.CO].

[47] D. J. Eisenstein, W. Hu, and M. Tegmark, "Cosmic complementarity: Joint parameter estimation from cosmic microwave background experiments and redshift surveys", *The Astrophysical Journal*, vol. 518, no. 1, p. 2, 1999. DOI: 10.1086/307261. [Online]. Available: https://dx.doi.org/10.1086/307261.

[48] A. Cooray, "Integrated sachs-Wolfe effect: Large scale structure correlation", *Phys. Rev. D*, vol. 65, p. 103 510, 2002. DOI: 10.1103/PhysRevD.65.103510. arXiv: astro-ph/0112408.

[49] J. J. Condon, W. D. Cotton, *et al.*, "The NRAO VLA Sky Survey", *The Astronomical Journal*, vol. 115, no. 5, pp. 1693–1716, 1998. DOI: 10.1086/300337.

[50] M. F. Skrutskie, R. M. Cutri, *et al.*, "The Two Micron All Sky Survey (2MASS)", *The Astronomical Journal*, vol. 131, no. 2, pp. 1163–1183, 2006. DOI: 10.1086/498708.

[51] P. A. R. Ade *et al.*, "Planck 2015 results. XXI. The integrated Sachs-Wolfe effect", *Astron. Astrophys.*, vol. 594, A21, 2016. DOI: 10.1051/0004-6361/201525831. arXiv: 1502.01595 [astro-ph.CO].

[52] M. Ballardini, D. Paoletti, F. Finelli, L. Moscardini, B. Sartoris, and L. Valenziano, "On the ISW-cluster cross-correlation in future surveys", *Mon. Not. Roy. Astron. Soc.*, vol. 482, no. 2, pp. 2670–2680, 2019. DOI: 10.1093/mnras/sty2795. arXiv: 1712.02380 [astro-ph.CO].

[53] S. Ilić *et al.*, "Euclid preparation - XV. Forecasting cosmological constraints for the Euclid and CMB joint analysis", *Astron. Astrophys.*, vol. 657, A91, 2022. DOI: 10.1051/0004-6361/202141556. arXiv: 2106.08346 [astro-ph.CO].

[54] J. R. Bermejo-Climent, M. Ballardini, *et al.*, "Cosmological parameter forecasts by a joint 2D tomographic approach to CMB and galaxy clustering", *Phys. Rev. D*, vol. 103, no. 10, p. 103 502, 2021. DOI: 10.1103/PhysRevD.103.103502. arXiv: 2106.05267 [astro-ph.CO].

[55] M. Hazumi, P. Ade, *et al.*, "Litebird: A satellite for the studies of b-mode polarization and inflation from cosmic background radiation detection", English, *Journal of Low Temperature Physics*, vol. 194, no. 5-6, pp. 443–452, 2019, ISSN: 0022-2291. DOI: 10.1007/s10909-019-02150-5.

[56] P. Ade *et al.*, "The Simons Observatory: Science goals and forecasts", *JCAP*, vol. 02, p. 056, 2019. DOI: 10.1088/1475-7516/2019/02/056. arXiv: 1808.07445 [astro-ph.CO].

[57] K. N. Abazajian *et al.*, "CMB-S4 Science Book, First Edition", Oct. 2016. arXiv: 1610.02743 [astro-ph.CO].

[58] A. Challinor and A. Lewis, "Linear power spectrum of observed source number counts", *Phys. Rev. D*, vol. 84, p. 043 516, 4 2011. DOI: 10.1103/PhysRevD.84.043516. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.84.043516.

[59] D. N. Limber, "The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field. II", *Astrophys. J.*, vol. 119, p. 655, 1954. DOI: 10.1086/145870.

[60] M. LoVerde and N. Afshordi, "Extended Limber Approximation", *Phys. Rev. D*, vol. 78, p. 123 506, 2008. DOI: 10.1103/PhysRevD.78.123506. arXiv: 0809.5112 [astro-ph].

[61] B. Stölzner, A. Cuoco, J. Lesgourgues, and M. Bilicki, "Updated tomographic analysis of the integrated Sachs-Wolfe effect and implications for dark energy", *Phys. Rev. D*, vol. 97, no. 6, p. 063 506, 2018. DOI: 10.1103/PhysRevD.97.063506. arXiv: 1710.03238 [astro-ph.CO].

[62] B. D. Wandelt, E. Hivon, and K. M. Gorski, "The pseudo-c_l method: cosmic microwave background anisotropy power spectrum statistics for high precision cosmology", *Phys. Rev. D*, vol. 64, p. 083 003, 2001. DOI: 10.1103/PhysRevD.64.083003. arXiv: astro-ph/0008111.

[63] 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, 2002. DOI: 10.1086/338126. [Online]. Available: https://dx.doi.org/10.1086/338126.

[64] G. Efstathiou, "Myths and truths concerning estimation of power spectra", *Mon. Not. Roy. Astron. Soc.*, vol. 349, p. 603, 2004. DOI: 10.1111/j.1365-2966.2004.07530.x. arXiv: astro-ph/0307515.

[65] E. Komatsu, B. D. Wandelt, D. N. Spergel, A. J. Banday, and K. M. Górski, "Measurement of the cosmic microwave background bispectrum on the cobe dmr sky maps", *The Astrophysical Journal*, vol. 566, no. 1, p. 19, 2002. DOI: 10.1086/337963. [Online]. Available: https://dx.doi.org/10.1086/337963.

[66] M. Tegmark, "How to measure CMB power spectra without losing information", *Phys. Rev. D*, vol. 55, pp. 5895–5907, 1997. DOI: 10.1103/PhysRevD.55.5895. arXiv: astro-ph/9611174.

[67] D. Marinucci, D. Pietrobon, A. Balbi, P. Baldi, P. Cabella, G. Kerkyacharian, P. Natoli, D. Picard, and N. Vittorio, "Spherical Needlets for CMB Data Analysis", *Mon. Not. Roy. Astron. Soc.*, vol. 383, p. 539, 2008. DOI: 10.1111/j.1365-2966.2007.12550.x. arXiv: 0707.0844 [astro-ph].

[68] U. Seljak, N. Hamaus, and V. Desjacques, "How to suppress the shot noise in galaxy surveys", *Physical Review Letters*, vol. 103, no. 9, p. 091 303, 2009.

[69] L. Knox, "Determination of inflationary observables by cosmic microwave background anisotropy experiments", *Physical Review D*, vol. 52, no. 8, p. 4307, 1995.

[70] W. J. Percival and M. L. Brown, "Likelihood methods for the combined analysis of CMB temperature and polarisation power spectra", *Mon. Not. Roy. Astron. Soc.*, vol. 372, pp. 1104–1116, 2006. DOI: 10.1111/j.1365-2966.2006.10910.x. arXiv: astro-ph/0604547.

[71] M. Gerbino, M. Lattanzi, M. Migliaccio, L. Pagano, L. Salvati, L. Colombo, A. Gruppuso, P. Natoli, and G. Polenta, "Likelihood methods for CMB experiments", *Front. in Phys.*, vol. 8, p. 15, 2020. DOI: 10.3389/fphy.2020.00015. arXiv: 1909.09375 [astro-ph.CO].

[72] S. Hamimeche and A. Lewis, "Likelihood Analysis of CMB Temperature and Polarization Power Spectra", *Phys. Rev. D*, vol. 77, p. 103 013, 2008. DOI: 10.1103/PhysRevD.77.103013. arXiv: 0801.0554 [astro-ph].

[73] P. Baldi, G. Kerkyacharian, D. Marinucci, and D. Picard, "Asymptotics for spherical needlets", *The Annals of Statistics*, vol. 37, no. 3, pp. 1150–1171, 2009. [Online]. Available: http://www.jstor.org/stable/30243664 (visited on 08/09/2023).

[74] X. Lan and D. Marinucci, "The needlets bispectrum", *Electronic Journal of Statistics*, vol. 2, no. none, pp. 332–367, 2008. DOI: 10.1214/08-EJS197. [Online]. Available: https://doi.org/10.1214/08-EJS197.

[75] Planck Collaboration, "Planck 2015 results - xvii. constraints on primordial non-gaussianity", *A&A*, vol. 594, A17, 2016. DOI: 10.1051/0004-6361/201525836. [Online]. Available: https://doi.org/10.1051/0004-6361/201525836.

[76] D. Marinucci and G. Peccati, *Random Fields on the Sphere: Representation, Limit Theorems and Cosmological Applications*, ser. London Mathematical Society Lecture Note Series. Cambridge University Press, 2011. DOI: 10.1017/CBO9780511751677.

[77] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. 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, no. 2, p. 759, 2005. DOI: 10.1086/427976. [Online]. Available: https://dx.doi.org/10.1086/427976.

[78] K. P. Drake and G. B. Wright, "A Fast and Accurate Algorithm for Spherical Harmonic Analysis on HEALPix Grids with Applications to the Cosmic Microwave Background Radiation", *J. Comput. Phys.*, vol. 416, p. 109 544, 2020. DOI: 10.1016/j.jcp.2020.109544. arXiv: 1904.10514 [math.NA].

[79] F. J. Narcowich, P. Petrushev, and J. D. Ward, "Localized tight frames on spheres", *SIAM Journal on Mathematical Analysis*, vol. 38, no. 2, pp. 574–594, 2006. DOI: 10.1137/040614359. eprint: https://doi.org/10.1137/040614359. [Online]. Available: https://doi.org/10.1137/040614359.

[80] M. Kamionkowski, A. Kosowsky, and A. Stebbins, "Statistics of cosmic microwave background polarization", *Phys. Rev. D*, vol. 55, pp. 7368–7388, 1997. DOI: 10.1103/PhysRevD.55.7368. arXiv: astro-ph/9611125.

[81] S. Eales, L. Dunne, D. Clements, *et al.*, "The Herschel ATLAS", *Publications of the Astronomical Society of the Pacific*, vol. 122, no. 891, p. 499, 2010. DOI: 10.1086/653086. arXiv: 0910.4279 [astro-ph.CO].