Alessio Spurio Mancini

Testing modified gravity theories with weak gravitational lensing

Dissertation

submitted to the
Combined Faculties of the Natural Sciences and Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Put forward by
Alessio Spurio Mancini
born in Ancona
Oral examination: October 17, 2018

In loving memory of Liliana, Fiorella, Salvatore, Vittorio and Tham

Testing modified gravity theories with weak gravitational lensing

---

Referees: Dr. Valeria Pettorino
Prof. Dr. Björn Malte Schäfer

Tests von modifizierten Gravitationstheorien mit dem schwachen kosmischen Gravitationslinseneffekt

Kosmische Scherung ist der schwache Gravitationslinseneffekt, der durch Fluktuationen der gravitativen Gezeitenkräfte der kosmischen großskaligen Struktur hervorgerufen wird und der sich in korrelierten Verzerrungen der beobachteten Formen von Galaxien manifestiert. Der Linseneffekt hängt sowohl von der Geometrie der Raumzeit wie auch vom Wachstum kosmischer Strukturen ab und ist daher ein hervorragendes Werkzeug, um in Durchmusterungen die Eigenschaften der Gravitation zu überprüfen.

In dieser Dissertation untersuche ich die Möglichkeit, mit Hilfe des schwachen Linseneffekts Alternativen zu dem Standardmodell der Kosmologie zu untersuchen, die die beschleunigte Expansion des Universums erklären können. Dabei konzentriere ich mich auf eine spezielle Klasse von Alternativen zu allgemeiner Relativität, nämlich der Horndeski-Klasse von Gravitationstheorien, die die Mehrzahl der universell an dunkle Materie gekoppelten Theorien mit einem zusätzlichen skalaren gravitativen Freiheitsgrad darstellen.

Bei einem vorgegebenen Hintergrund ist die Zeitentwicklung von linearen Störungen in Horndeski-Gravitation durch vier nur von der Zeit abhängenden Funktionen beschrieben. Zuerst bestimme ich die Empfindlichkeit zukünftiger Durchmusterungen wie die des Euclid-Satelliten, diese Funktionen zu bestimmen. Dabei vergleichen wir zwei verschiedene Ansätze: Tomographie, bei der Korrelationen des Linsensignals bei verschiedenen Rotverschiebungsintervallen untersucht wird und eine dreidimensionale Fourier-Bessel-Zerlegung in sphärischen Koordinaten. Die zweite Methode ist in der Lage, genauere Messungen aller kosmologischen Parameter durchzuführen, und liefert Verbesserungen um Faktoren bis zu zwanzig.

Anschließend untersuche ich die Möglichkeit, mit Kreuzkorrelationen zwischen dem schwachen Gravitationslinseneffekt und anderen Beobachtungen Horndeski-Gravitationstheorien stärker einzuschränken. Ich analysiere die Kombination des Gravitationslinseneffeks, dem Clustern von Galaxien und dem Linseneffekt in Galaxienpaaren auf der Basis der Kilo Degree Survey und der Galaxy And Mass Assembly Survey, um Horndeski-Parameter zu messen. Danach erweitere ich diese Untersuchung auf zukünftige Durchmusterungen des Linseneffekts, der Galaxienverteilung und des kosmischen Mikrowellenhintergrunds. Während aktuelle Einschränkungen noch nicht sehr stark sind, ändert sich die Situation in der nächsten Generation der Experimente grundlegend.

Zuletzt bespreche ich im Detail numerische Techniken, die ich für Analysen des schwachen kosmischen Linseneffekts benutzt habe und vergleiche meine Ergebnisse mit einer alternativen, unabhängig entwickelten Methode, wobei ich eine hervorragende Übereinstimmung finde. Ich benutze meine Methode, um dreidimensionale Karten des schwachen kosmischen Linseneffekts aus den entsprechenden Kovarianzmatrizen zu konstruieren. Ich bestimme Minkowski-Funktionale dieser Zufallsfelder und benutze sie, die Genauigkeit der Realisierung von Zufallsfeldern zu überprüfen ebenso wie zur kosmologischen Inferenz einzusetzen.

Testing modified gravity theories with weak gravitational lensing

'Cosmic shear' is the weak gravitational lensing effect generated by fluctuations of the gravitational tidal fields of the large-scale structure that induce correlations in the distortion of observed galaxy shapes. Being sensitive to spacetime geometry and the growth of cosmic structure, cosmic shear is one of the primary probes to test gravity with current and future surveys.

In this thesis we analyse the power of cosmic shear to constrain alternatives to the standard cosmological model that could explain cosmic acceleration. We focus in particular on a large class of alternatives to General Relativity, the Horndeski class, which includes the majority of universally coupled extensions to ΛCDM with one scalar degree of freedom in addition to the metric.

Given a fixed background, the evolution of linear perturbations in Horndeski gravity is described by a set of four functions of time only. First, we forecast the sensitivity to these functions that will be achieved by future cosmic shear surveys like Euclid. We produce our forecasts with two methods to analyse a cosmic shear survey: a tomographic approach, based on correlations of the lensing signal in different redshift bins, and a fully 3D spherical Fourier-Bessel decomposition of the shear field. We show how the latter produces tighter constraints on all cosmological parameters with a sensitivity gain of the order of 20% in particular on the ones that describe Horndeski gravity.

We then consider the possibility of using cross-correlations of cosmic shear with other probes to constrain Horndeski theories of gravity. We analyse a combination of cosmic shear, galaxy-galaxy lensing and galaxy clustering data from the Kilo Degree Survey and Galaxy And Mass Assembly survey and set constraints on the aforementioned Horndeski parameters. We also forecast the expected sensitivity to the same parameters that could be achieved with future cross-correlations of Stage IV cosmic shear, galaxy clustering and CMB experiments. While current constraints are not very tight, our implementation could be used in the future with data coming from Stage IV surveys, which we show to have great constraining power on these theories.

Finally, we present in detail the numerical techniques that we used to produce our 3D cosmic shear forecasts and compare our predictions with an alternative, independent method developed with the same purpose. We find excellent agreement between the two methods and use our simulated 3D cosmic shear covariance matrices within a new algorithm that we develop to generate 3D lensing random fields. We calculate the Minkowski Functionals associated to our random fields and use them to test our field-generation procedure, as well as to demonstrate the possibility of a new approach to cosmological inference leveraging the estimated Minkowski Functionals.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Towards the end of last century supernovae observations (Riess et al., 1998; Perlmutter et al., 1999) confirmed that the Universe has been experiencing an accelerated expansion in recent times. The observed acceleration can be ascribed to a dark energy component accounting for approximately 70% of the energy budget of the Universe. The simplest and best-known candidate for dark energy is the energy of the vacuum, represented in Einstein's equations by the cosmological constant term. This vacuum energy density, unchanging in time and spatially constant, is currently in good agreement with existing data and defines the standard concordance model for cosmology, denoted ΛCDM to highlight its two most important (and least understood) ingredients: the cosmological constant and a Cold Dark Matter component, i.e. a non-relativistic collisionless fluid that interacts only gravitationally with baryonic matter. Together, dark energy (in the form of a cosmological constant in the ΛCDM scenario) and dark matter constitute approximately 95% of the energy budget of the Universe. However, despite the ΛCDM model being firmly established as the 'standard model' of cosmology, as it is often referred to in analogy to particle physics, the nature of its major components is far from understood: on the one hand, dark matter particles have not been detected yet; on the other hand, while identifying the dark energy with a cosmological constant term Λ fits well the observations, it has long been questioned in terms of naturalness and interpretation in terms of energy density of the vacuum (see Martin, 2012, for a recent review).

Alternatives to the cosmological constant can be generally grouped into two main categories: either dark energy is a modification of gravity on the largest scales ("modified gravity" theories), or it is given by a scalar field that effectively behaves as a fluid with negative pressure (usually referred to as proper "dark energy" models). The distinction between these two classes can at times be feeble (see Joyce et al., 2016, for a recent discussion) and the vast amount of proposed theories (see Clifton et al., 2012, for a review) urgently calls for methods to be developed, aiming at distinguishing among the large number of theoretical options with advanced statistical methods and efficient computational effort. This is particularly relevant in light of the unprecedented amount of data that will come from many space- and ground-based experiments, such as Euclid [1](Laureijs et al., 2011), SKA [2](Maartens et al., 2015), LSST [3](LSST Science Collaboration et al., 2009) and WFIRST [4](Spergel et al., 2013), whose launch in the next few years is planned with the goal of unveiling the true nature of the cosmic acceleration.

On the observational side, many different probes have been proposed to investigate dark energy and modified gravity models. These include type Ia Supernovae, Baryon Acoustic Oscillations, galaxy clustering and weak gravitational lensing, to name a few (see e.g. Weinberg et al., 2013, for an exhaustive review of the different probes). One of the most important lessons learned in cosmology in

[1]https://www.euclid-ec.org/
[2]https://www.skatelescope.org/
[3]https://www.lsst.org/
[4]https://wfirst.gsfc.nasa.gov/

recent years, in particular following the scientific results produced by the Planck Collaboration (Planck Collaboration et al., 2016b), is the need to move away from the cosmological background towards observational probes of the perturbations, if we want to distinguish between different models of dark energy and modified gravity. Gravitational lensing is a leading observational tool in this context and the main focus of this thesis.

We concentrate in particular on the weak gravitational lensing caused by the large-scale structure of the Universe, or cosmic shear. Since the first detections in early 2000s (e.g. Bacon et al., 2000; Van Waerbeke et al., 2000; Brown et al., 2003), this field has developed within a well-established theoretical and experimental framework. Cosmic shear is particularly appealing as one of the most promising probes of dark energy (Jain & Taylor, 2003; Bernstein & Jain, 2004; Hannestad et al., 2006; Amendola et al., 2008; Huterer, 2010): the differential deflection in light bundles from distant galaxies caused by variations of the gravitational fields of the large-scale structure result in a coherent distortion of galaxy images as we observe them on the sky (see Bartelmann & Schneider, 2001; Hoekstra & Jain, 2008; Kilbinger et al., 2013, for reviews on the topic).

The main goal of this thesis is to investigate the possibility of studying dark energy/modified gravity models with current and future cosmic shear surveys. This kind of work entails the production of constraints on parameters that describe these alternatives to the cosmological constant, either by using currently available datasets or by calculating predictions for next generation Stage IV surveys. A significant part of the work implies the study of constraints achievable from weak gravitational lensing on a large class of dark energy/modified gravity theories: the Horndeski class.

The Horndeski class represents indeed an example of a remarkably large set of extensions to General Relativity. First discussed in 1974 by Horndeski (Horndeski, 1974) and subsequently rediscovered in Nicolis et al. (2009) and Deffayet et al. (2011), the Horndeski Lagrange density is the most general gravitational theory with one scalar degree of freedom, in addition to the metric tensor, with derivatives in the equations of motion not higher than second order; this guarantees safety from ghost-like degrees of freedom. This set of theories collects under its name many different models of dark energy/modified gravity (see Sec. 2.5.2 for a list of some of them).

Recently (Gleyzes et al., 2013; Bellini & Sawicki, 2014), it has been demonstrated that given a fixed background, the evolution of linear perturbations in Horndeski gravity can be described by a set of four functions of time only. This is a very appealing result from an observational perspective, as it reduces the range of possibilities to a few functions to constrain. In our work we will present different analyses aimed at producing constraints on these functions.

The first results of this thesis concern the two main techniques for analysing a cosmic shear survey; a tomographic method, where correlations between the lensing signal in different redshift bins are used to recover redshift information, and a 3D approach, where the full redshift information is carried through the entire analysis. In this thesis we compare the two methods, by forecasting cosmological constraints for future surveys like Euclid. We extend the 3D formalism to theories beyond the standard model belonging to the Horndeski class. We model the time evolution of the functions mentioned earlier, fully describing the evolution of linear perturbations in Horndeski gravity, assuming proportionality to the dark energy density fraction, and estimate expected constraints from future surveys on the proportionality coefficients. We find that a 3D analysis can constrain Horndeski theories better than a tomographic one, in particular with a decrease in the errors of the order of 20%. Our work shows for the first time a quantitative comparison on an equal footing between forecasts for both a fully 3D and a tomographic analysis of cosmic shear surveys. The increased sensitivity of the 3D formalism comes from its ability to retain information on the source redshifts along the entire analysis.

We then proceed to consider cosmic shear in combination with other observables, and present constraints on the aforementioned Horndeski functions, obtained from our joint analysis of the tomographic cosmic shear signal in $\sim 450\text{deg}^2$ of data from the Kilo Degree Survey (KiDS), the galaxy-

matter cross-correlation signal of galaxies from the Galaxies And Mass Assembly (GAMA) survey determined with KiDS weak lensing, and the angular correlation function of the same GAMA galaxies. As expected, the available data do not allow for tight constraints on the Horndeski functions, however the methodology and implementation that we have developed will allow for tighter constraints with future data releases of the KiDS survey and/or future Stage IV surveys such as Euclid.

Looking ahead towards the future, we proceed in this thesis considering correlations of the cosmic shear field with other observables, including galaxy clustering and Cosmic Microwave Background (CMB) primary anisotropies and lensing, and forecast expected constraints on Horndeski gravity attainable from a combination of future Stage IV surveys such as Euclid and CMB experiments such as CMB-S4.

The aforemenetioned 3D analysis of the cosmic shear field, based on a spherical-Bessel decomposition of the field ("3D cosmic shear") is one way to maximise the amount of redshift information in a lensing analysis and therefore provides a powerful tool to investigate in particular the growth of cosmic structure that is crucial for dark energy studies. However, the computation of simulated 3D cosmic shear covariance matrices presents numerical difficulties, due to the required integrations over highly oscillatory functions. In this thesis we present and compare two numerical methods and relative implementations to perform these integrations. We then show how to generate 3D Gaussian random fields on the sky in spherical coordinates, starting from the 3D cosmic shear covariances. To validate our field-generation procedure, we calculate the Minkowski functionals associated to our random fields, compare them with the known expectation values for the Gaussian case and demonstrate parameter inference from Minkowski functionals from a cosmic shear survey. This is a first step towards producing fully 3D Minkowski functionals for a lognormal field in 3D to extract Gaussian and non-Gaussian information from the cosmic shear field, as well as towards the use of Minkowski functionals as a probe of cosmology beyond the commonly used two-point statistics.

The detailed structure of this thesis is as follows: we start in Chapter 2 by reviewing the concordance cosmological model, as well as introducing theories and parameterizations of dark energy/modified gravity; we then proceed in Chapter 3 to the description of the statistical tools employed throughout our analysis. In Chapter 4 we review the theory of cosmic shear, with emphasis on the spherical Fourier-Bessel formalism and its relation to a tomographic approach. In Chapter 5 we present the expected constraints from future cosmic shear surveys on Horndeski gravity obtained in our comparison between 3D cosmic shear and tomography. In Chapter 6 we present our cross-correlation analysis of the KiDS and GAMA surveys and the forecasts produced considering cross-correlations of future Stage IV surveys. Chapter 7 is devoted to the description of the numerical techniques used for the evaluation of the 3D cosmic shear integrals and the generation of 3D lensing random fields, as well as to the calculation of the Minkowski Functionals associated to these fields. Finally, we draw our conclusions in Chapter 8.

2 Chapter 2

An introduction to dark energy/modified gravity cosmologies

The discovery of the accelerating universe from supernovae observations in the late 1990s (Riess et al., 1998; Perlmutter et al., 1999) unambiguously indicated the presence of a new component in the Universe dominating the energy density today, or of a modification of the laws of gravity. The new component quickly became a fundamental costituent of the new standard cosmological model, which also features baryonic matter, dark matter, and radiation (photons and relativistic neutrinos). The simplest and best-known candidate for dark energy is the energy of the vacuum, represented in Einstein's equations by the cosmological-constant term. Vacuum energy density, unchanging in time and spatially constant, is currently in good agreement with existing data and defines the standard concordance model for cosmology, denoted ΛCDM to highlight the presence of a cosmological costant Λ in addition to the (mainly Cold Dark) Matter component. Yet, there exists a rich set of other dark energy models, including evolving scalar fields and modifications to General Relativity as the theory of gravity on cosmological scales, which we will briefly introduce in this Chapter.

We will start by reviewing the basics of cosmology: Sec. 2.1 is dedicated to the background expansion, while Sec. 2.3 discusses the growth of density fluctuations generating cosmological structures as we observe them today, after reviewing in Sec. 2.2 the statistical tools to study the evolution of cosmic fields. Sec. 2.4 is dedicated to briefly reviewing the main problems with the Λ paradigm, while Sec. 2.5 introduces different possibilities to extend the concordance model to a dark energy component or a different law of gravity on cosmological scales. Sec. 2.6 reviews general approaches to describe gravity on cosmological scales, first through the possible background parameterizations for the dark energy component and then considering effective theory of dark energy and phenomenological parameterizations of the gravitational potentials acting at the perturbations level. We conclude in Sec. 2.7 with a discussion of the screening mechanism.

2.1. Homogeneous and isotropic background

The cosmological principle is the fundamental assumption underpinning modern cosmology. It states that the Universe on scales of a few hundred Mpc/h, i.e. greater than the largest observed cosmological structures, appears homogeneous and isotropic in the spatial dimensions to a freely falling observer. For this observer the line element $\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$, built from the metric tensor $g_{\mu\nu}$ and specifying the geometry of space-time, takes the Friedmann-Robertson-Walker (FRW) form,

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\left(\frac{1}{1-K\chi^2}\mathrm{d}\chi^2 + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\varphi^2\right), \tag{2.1}$$

where we used spherical spatial coordinates (χ, θ, ϕ). χ is the comoving distance, obtained factoring out from the radial coordinate the scale factor $a(t)$; the latter describes the scaling of all physical dis-

tances with cosmic time t measured by a freely falling observer on a clock. Usually, $a(t)$ is normalised to have value of unity at present time. The quantity K is the intrinsic curvature of three-dimensional space; $K = 0$ corresponds to a spatially flat universe with Euclidean geometry, while $K > 0$ corresponds to positive curvature (spherical geometry with finite volume) and $K < 0$ to negative curvature (hyperbolic geometry with infinite volume). The cosmological principle is an assumption, but also a testable hypothesis, as indeed there is excellent observational evidence that the Universe satisfies homogeneity and isotropy on its largest spatial scales (e.g. Laurent et al., 2016).

Knowledge of the metric allows us to solve the gravitational field equations that link the local and divergence-free curvature, given by the Einstein tensor $G_{\mu\nu}$, to the energy-momentum tensor $T_{\mu\nu}$

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \quad \text{with} \quad G_{\mu\nu} = R_{\mu\nu} - \frac{R}{2} g_{\mu\nu}. \tag{2.2}$$

The Riemann tensor $R_{\mu\nu}$ and Ricci scalar R describe the geometry of the Universe, as they are related to the metric tensor $g_{\mu\nu}$ and its first and second derivatives. The stress-energy tensor quantifies the energetic content of the Universe and contains contributions from all species; it obeys local energy-momentum conservation $\nabla^\mu T_{\mu\nu} = 0$, where ∇ denotes a covariant divergence. The source of the gravitational field is $T_{\mu\nu}$. The most general form of $T_{\mu\nu}$ for a fluid in the FRW metric is given by an ideal, relativistic fluid

$$T_{\mu\nu} = \left(\rho c^2 + p\right) \upsilon_\mu \upsilon_\nu - g_{\mu\nu} p. \tag{2.3}$$

The cosmic fluid can be completely parameterised specifying its density $\rho = \rho(t)$ and pressure $p = p(t)$, which can be only functions of time; isotropy does not permit any spatial dependence. In Eq.2.3 υ_μ denotes the 4-velocity.

The scale factor $a(t)$ is a function of the energy densities and pressures of the components that fill the Universe. Its evolution is governed by the Friedmann equations, which can be derived from the field equations Eq. 2.2 inserting the FRW metric Eq.2.1 and the expression for a general ideal relativistic fluid Eq.2.3:

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G \rho}{3} - \frac{k}{a^2} + \frac{\Lambda}{3}\,, \tag{2.4}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3}\,, \tag{2.5}$$

where H is the Hubble parameter, defined as the time derivative of the logarithmic scale factor, i.e., $H(t) = \dot{a}/a$, Λ is a cosmological constant term that in general can be added to the field equations Eq.2.2, ρ is the total energy density, and p is the pressure.

Th only free choice allowed by the FRW symmetries for a relation between pressure and density is the equation of state

$$p = w \rho c^2, \tag{2.6}$$

where c denotes the speed of light. Pressureless dark matter is characterised by $w = 0$, relativistic matter by $w = 1/3$ and the value $w \equiv -1$ corresponds to the cosmological constant Λ.

We can define the critical density $\rho_{\rm crit} \equiv 3H^2/(8\pi G)$ as the density that leads to a flat universe with $K = 0$. We can then express all densities as dimensionless density parameters:

$$\Omega_i(a) := \frac{\rho_i(a)}{\rho_{\rm crit}(a)} \quad \Omega_{i0} := \frac{\rho_i(a=1)}{\rho_{\rm crit}(a=1).} \tag{2.7}$$

The expansion rate of the universe $H \equiv \dot{a}/a$ from (2.4) can then be written as

$$H^2 = H_0^2[\Omega_{m0} a^3 + \Omega_{r0} a^4 + \Omega_\Lambda + \Omega_K a^2] = H_0^2 E(a) \tag{2.8}$$

where $H_0 = H(a = 1) = 100h\frac{\mathrm{km}}{\mathrm{sMpc}}$ is the present value of the Hubble parameter (the Hubble constant), Ω_m and Ω_r are the matter and radiation energy densities relative to critical, and the dimensionless curvature "energy density" Ω_k is defined such that $\sum_i \Omega_i = 1$. Since $\Omega_r \simeq 8 \times 10^{-5}$, we can typically ignore the radiation contribution for low-redshift ($z \lesssim 10$) measurements; however, near the epoch of recombination ($z \sim 1000$), radiation contributes significantly, and at earlier times ($z \gtrsim 3300$), it dominates. The closure relation $\sum_i \Omega_i = 1$ and the scaling with a of the different terms in Eq. 2.8 means that by specifying the density parameters and the Hubble constant, the background evolution of the Universe is completely determined.

2.2. Statistical properties of random fields in structure formation

The topic of this section is the formation of cosmic structure as we observe them today, deviating from the homogeneous background described in the previous Section. Such a cosmological background cannot originate the highly inhomogeneous structures that we observe around us: however, if there exists a mechanism able to generate small inhomogeneities in the density field, these can grow by gravitational instability and eventually form highly inhomogenenous structures. Currently the widest accepted mechanism for the generation of these primordial fuctuations is inflation (Guth, 1981): the fluctuations are seeded by quantum fluctuations of the inflaton field, which drives an era of rapid accelerated expansion in the early Universe.

Given initial perturbations in the density field we will now briefly review how we can describe the statistical properties of random fields and how the power spectrum of the initial perturbations looks like. The ultimate goal is to make statistical predictions about cosmic fields which depend on the statistical nature of the primordial fluctuations. The connection between these statistical descriptions will be clearer later when we discuss the linear growth: ultimately it is thanks to the homogeneous growth in the linear regime of structure formation that we can investigate initial conditions given inflationary processes in the cosmological structures, by observing the large-scale structure today. Let us start by summarising the basics of structure formation.

To describe the fluctuations of the cosmic density field $\rho(\mathbf{x})$ one can choose the density contrast, i.e. the relative deviation of the density field $\rho(\mathbf{x})$ from the mean background density $\langle\rho\rangle = \Omega_m\rho_{\mathrm{crit}}$,

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \langle\rho\rangle}{\langle\rho\rangle}. \tag{2.9}$$

These fluctuations are random variables i.e. follow a certain probability distribution: considering a hypothetical ensemble of statistically equivalent Universes, the probability of finding the amplitudes $\delta(\mathbf{x}_1)$ and $\delta(\mathbf{x}_2)$ and positions $\mathbf{x}_1$ and $\mathbf{x}_2$ is assumed to be a bivariate Gaussian, with a covariance matrix $C = \left\langle \delta^T \delta \right\rangle$ that contains the correlation function $\xi(\mathbf{x}_1, \mathbf{x}_2) \equiv \langle\delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\rangle$ of the random field. The brackets $\langle\ldots\rangle$ denote averages over many (hypothetical, in fact) realisations of the density field.

The correlation function $\xi(\mathbf{x}_1, \mathbf{x}_2)$ is the off-diagonal element of the covariance matrix C; if the correlation function vanishes, this means that the amplitudes are mutually uncorrelated and follow independently from univariate Gaussian distributions. In contrast, a finite correlation function defines a correlation length in the density random field. Due to statistical isotropy and homogeneity of the density field, the correlation function can depend only on the absolute value r of the separation $\mathbf{r} = \mathbf{x}_2 - \mathbf{x}_1$.

The most common descriptor of the density field is the power spectrum, a particularly convenient description for a statistically homogeneous field. The power spectrum is defined as the variance of the density contrast δ in Fourier space:

$$\langle\delta(\mathbf{k}_1)\delta^*(\mathbf{k}_2)\rangle = (2\pi)^3\delta_D(\mathbf{k}_1 - \mathbf{k}_2)P(\mathbf{k}_1). \tag{2.10}$$

Here, the Fourier transform is defined as follows:

$$\delta(\mathbf{k}) = \int \mathrm{d}^3x\, \delta(\mathbf{x}) \exp(-\mathrm{i}\mathbf{k}\mathbf{x}) \quad \leftrightarrow \quad \delta(\mathbf{x}) = \int \frac{\mathrm{d}^3k}{(2\pi)^3}\, \delta(\mathbf{k}) \exp(+\mathrm{i}\mathbf{k}\mathbf{x}). \tag{2.11}$$

$P_\delta(k)$ cannot depend on the orientation of the wavevector k due to isotropy; as long as the density field is statistically homogeneous, there is no coupling between different modes, which is expressed by the Dirac delta distribution in Eq.2.10. The power spectrum is the Fourier transform of the two-point correlation function:

$$P_\delta(k) = \int \mathrm{d}^3x \left\langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2 + \mathbf{r}) \exp(-\mathrm{i}\mathbf{k}\mathbf{r}) \right\rangle = 2\pi \int r^2 \mathrm{d}r\, \xi(r)\, j_0(kr), \tag{2.12}$$

where the second equality is valid only for statistically homogeneous and isotropic fields. In principle, $P_\delta(k)$ is a function of coordinate time, reflecting the evolution of the density field, and is sometimes expressed as a function of redshift, i.e. $P_\delta(k, z)$. As we will see in the following, statistical homogeneity is preserved in the linear regime, so that the redshift evolution changes the amplitude of the spectrum, but not its shape. Non-linear growth, in contrast, leads to mode coupling and moves power to smaller scales as structures collapse; therefore the redshift dependence in the non-linear regime is non-trivial.

All the above definitions can be extended to higher order correlators for statistically homogeneous fields:

$$(2\pi)^3 P_n(\mathbf{k}_1, \mathbf{k}_2, \cdots \mathbf{k}_n)\delta_D^{(3)}(\mathbf{k}_1 + \cdots + \mathbf{k}_n) := \langle \delta(\mathbf{k}_1, t) \cdots \delta(\mathbf{k}_n, t) \rangle_C \tag{2.13}$$

Here again the Dirac distribution ensures statistical homogeneity. The index C denotes the connected part, corresponding to the cumulants of the distribution. The $n = 2$ correlator corresponds to the power spectrum, while the $n = 3$ correlator is called the bispectrum. Gaussian random fields have the useful property that all cumulants of higher order than two vanish. In Eq. 2.13 all spectra with $n > 2$ will vanish. This also has some important consequences for higher order moments of the distribution: while odd n-point correlation functions of Gaussian random fields are equal to zero, even n-point functions can be decomposed into products of two-point functions by virtue of the Wick theorem,

$$\langle \delta(\mathbf{k}_1) \dots \delta(\mathbf{k}_n) \rangle = \sum_{\text{pairs}} \prod_{i,j \in \text{pairs}} \left\langle \delta(\mathbf{k}_i)\delta(\mathbf{k}_j) \right\rangle, \tag{2.14}$$

As explained e.g. in Schäfer (2018), for a zero-mean Gaussian distributed random variable all moments are proportional to the variance, therefore the latter is everything needed to reconstruct this probability distribution from its moments.

2.3. The growth of density fluctuations

On scales much smaller than those on which spacetime curvature becomes important (i.e. on sub-horizon scales $k \geq 2\pi/\chi_H$, where $\chi_H = c/H_0$ is the Hubble radius today), gravitational instability can be treated in the Newtonian framework. As the matter budget of the Universe is dominated by cold dark matter, the focus is on a collisionless, self-gravitating medium in a homogeneous and isotropic background expanding at the Hubble rate.

2.3.1. Gravitational instability of collisionless dark matter

Since the collision term in the Boltzmann equation vanishes in the case of collisionless dark matter, the equations governing the evolution of the matter density ρ, the gravitational potential Φ and the

velocity field υ have the same shape as the hydrodynamical equations describing the conservation of mass and momentum and the self-gravity of a perfect fluid with negligible pressure. As long as the overdensity is small, $\delta \ll 1$, the treatment can be restricted to the first order in the perturbations. We have the continuity equation

$$\frac{\partial}{\partial \eta}\delta + \mathrm{div}\upsilon = 0, \tag{2.15}$$

and the linearised Euler equation,

$$\frac{\partial}{\partial \eta}\upsilon + aH\upsilon = -\nabla\Phi, \tag{2.16}$$

which are valid only if the overdensity δ is small, $\delta \ll 1$. η is the conformal time, related to coordinate time t by $\mathrm{d}\eta = \mathrm{d}t/a$. The Newtonian Poisson-equation is instead always linear,

$$\Delta\Phi = \frac{3H_0^2\Omega_m}{2a}\delta. \tag{2.17}$$

These equations can be combined into a single differential equation for the density contrast (e.g. Linder & Jenkins, 2003):

$$\frac{\mathrm{d}^2}{\mathrm{d}a^2}D_+(a) + \frac{1}{a}\left(3 + \frac{\mathrm{d}\ln H}{\mathrm{d}\ln a}\right)\frac{\mathrm{d}}{\mathrm{d}a}D_+(a) = \frac{3}{2a^2}\Omega_m(a)D_+(a). \tag{2.18}$$

The second term, proportional to $\mathrm{d}\delta/\mathrm{d}a$, acts as a friction term and is sometimes called the 'Hubble drag', since it reflects the Hubble expansion counteracting the gravitative attraction. As temporal and spatial dependence decouple in the linear regime, the density factorises:

$$\delta(\mathbf{x}, a) = D_+(a)\delta(\mathbf{x}). \tag{2.19}$$

The function $D_+(a)$ describing the evolution is called the growth function and represents the growing solution of Eq.2.18. The second solution decays with increasing a and consequently cannot describe the amplification of density perturbations. The growth function is usually normalised to $D_+(a = 1) = 1$. In a matter-dominated universe ($\Omega_m = 1$), the Hubble rate is $H(a) \sim a^{-3/2}$, so that $D_+(a) = a$.

Knowledge of the Hubble function $H(a)$ or the growth function $D_+(a)$ allows us to calculate the other quantity: the two are not independent, in fact Peebles (1980) shows that $D_+(a)$ is given by

$$D_+(a) \propto H(a)\int_a^1 \frac{\mathrm{d}a}{(aH(a))^3}. \tag{2.20}$$

This happens (Schäfer, 2018) because in General Relativity the dependence $\Omega_m(a)$ on time is given entirely by the Hubble function $H(a)$, and that in Newtonian gravity the same term determines the term on the right hand side driving structure formation. This is not true in general in modified gravity theories, or in cosmologies with interactions between the cosmological fluids; different predictions from the two terms would then represent a possible signature of such an alternative cosmology.

Homogeneous structure formation corresponds to independently growing Fourier modes,

$$\delta(\mathbf{x}, a) = D_+(a)\delta(\mathbf{x}, a = 1) \longrightarrow \delta(\mathbf{k}, a) = D_+(a)\delta(\mathbf{k}, a = 1), \tag{2.21}$$

which conserves every statistical property of the initial conditions, in particular Gaussianity. The Gaussianity of the initial density perturbations is a consequence of inflation, where a large number of uncorrelated quantum fluctuations are superimposed, yielding a Gaussian amplitude distribution due to the central limit theorem.

2.3.2. Primordial power spectrum

The shape of the primordial spectrum is determined by the mechanism of inflation. Most models predict Gaussian perturbations with a scale-invariant dimensionless power spectrum

$$k^3 P_\Phi(k) \propto k^{n_s - 1} \tag{2.22}$$

where $n_s \sim 1$ is the spectral index. According to the Poisson equation, the scaling between Fourier modes of the density contrast and the gravitational potential is $\delta(\mathbf{k}) \sim k^2\Phi(\mathbf{k})$. The initial matter power spectrum then has the shape $P_{\delta,\mathrm{ini}}(k) \sim k^4 P_\Phi(k) \sim k^{n_s}$, which is the Harrison-Zeldovich-Peebles spectrum (Zeldovich, 1971). The suppression of growth during radiation-domination is accounted for by the introduction of the transfer function $T(k)$:

$$T(q) = \frac{\ln(1 + 2.34q)}{2.34q}\left(1 + 3.89q + (16.1q)^2 + (5.46q)^3 + (6.71q)^4\right)^{-\frac{1}{4}},$$

or the more accurate fit described by Eisenstein & Hu (1998) incorporating baryonic wiggles for flat cosmological models with low matter density Ω_m. The asymptotic behaviour of the transfer function is such that $T(k) \propto$ const for $k \ll 1$ and $T(k) \propto k^{-2}$ at $k \gg 1$, such that $P(k) \propto k^{n_s}$ on large scales and $P(k) \propto k^{n_s - 4}$ on small scales. The usual form of the linearly evolved power spectrum is then

$$P_{\mathrm{lin}}(k, a) = D_+^2(a) T^2(k) P_{\delta,\mathrm{ini}}(k). \tag{2.23}$$

In addition to the spectral index n_s and the transfer function, the amplitude of the power spectrum must be given. This is most commonly characterised by the parameter σ_8 , which is the variance of the smoothed density field, i.e. $\delta(x)$ convolved with a top hat filter of the radius $R = 8h^{-1}$ Mpc:

$$\sigma_R^2 = \frac{1}{2\pi^2}\int_0^\infty \mathrm{d}k\, k^2 P(k) W^2(kR), \tag{2.24}$$

$W(x) = j_1(x)/x$ is the three-dimensional Fourier transform of the top hat filter; $j_1(x)$ is the first spherical Bessel function.

2.3.3. Nonlinear structure formation

In the linear regime of structure formation the growth is homogeneous and conserves the Gaussianity of the initial conditions. Nonlinear structure formation implies instead inhomogeneous growth and the emergence of non-Gaussian features. Following Schäfer (2018), we can understand this tight bound in structure formation between non-gaussianity, non-linearity and inhomogeneity of the density field. From non-linearity follows inhomogeneity, because if e.g. the density contrast approaches $\delta \simeq -1$ in a deeply underdense region of the Universe, the linearisation fails and at the same time locally the growth must decrease. Inhomogeneity implies non-Gaussianity because if the amplitudes of the density contrast δ increase, their initially Gaussian distribution, $p(\delta)\mathrm{d}\delta$ must also become wider; however, the density contrast is bounded from below at the value -1, requiring the amplitude distribution $p(\delta)\mathrm{d}\delta$ to become asymmetric, therefore losing Gaussianity. Finally in inhomogeneous growth the Fourier-modes $\delta(\mathbf{k}, a)$ couple to each other since the growth $D_+(\mathbf{x}, a)$ now also depends on the position; coupling of Fourier modes implies that the central limit theorem does not hold any more because the superposition of the Fourier yields modes a non-Gaussian distribution.

Different approaches to modelling the nonlinear growth of structure

At late times during cosmological evolution the linear description breaks down and one can consider an expansion to higher order

$$\delta(\mathbf{x}, a) = \sum_{n=1}^{\infty} \delta^{(n)}(\mathbf{x}, a) \simeq \sum_{n=1}^{\infty} D_+^n(a)\delta^{(n)}(\mathbf{x}) \tag{2.25}$$

and similarly for the velocity divergence, such that $\delta^{(1)}$ corresponds to the linear solution. Each field $\delta^{(n)}$ grows homogeneously at the rate $D_+^n(a)$, but the sum does not. The second equality is exact if $\Omega_m = 1$ and also approximately valid in dark energy cosmologies. Plugging this into Eqs.2.15,2.16,2.17, we can see that the non-linearities in the continuity and Euler equation imply convolutions of the density and velocity fields in Fourier space. This couples the different Fourier modes. This perturbative expansion is the basis of Eulerian perturbation theory. An alternative approach is Lagrangian perturbation theory, where the focus is on the particle trajectories linking the initial positions $\mathbf{q}$ to their positions $\mathbf{x}$ at time η, rather than on the density and velocity fields. A recent approach (Kinetic Field Theory, Bartelmann et al., 2017) is based instead on a non-equilibrium statistical field theory for classical particles, which describes the dark matter particles as an ensemble occupying the phase space subject to Hamiltonian dynamics.

Another promising possibility (albeit numerically challenging) is to rely on numerical N-body simulations (Vogelsberger et al., 2014). These simulations use effective particles with different properties, for example dark matter particles interact only gravitationally and have typical masses of $10^9\, M_\odot$. The power spectrum of the Cosmic Microwave Background is used initially to correlate the position and momenta of the particles, subject to the continuity equation. They are then evolved with analytic methods up to a redshift of roughly 100. For smaller redshifts instead the full Newtonian dynamical equations is solved numerically on an expanding background. Modern simulations also contain baryonic physics and sophisticated high-resolution methods to describe different types of baryon feedback mechanism (Vogelsberger et al., 2014). The problem of these simulations is that they are computationally very expensive. For applications to future surveys, a complication comes from the fact that they will span an enormous amount of different scales, requiring a large volume needed for the simulations and, at the same time, a sufficiently high resolution. Furthermore, the observable properties of the LSS are of statistical nature, therefore one needs an ensemble of simulations for each individual set of cosmological parameters. The development of techniques to model correctly and efficiently nonlinear structure formation is of paramount importance for future surveys and in this sense a joint analysis with analytic and numerical methods needs to be investigated for next generation surveys, to leverage the strengths and reducing the weaknesses of both approaches.

2.4. Problems with the Cosmological Constant

The observed acceleration of the Universe (Riess et al., 1998; Perlmutter et al., 1999) can be ascribed to a dark energy component accounting for approximately 70% of the energy budget of the Universe. From a theoretical point of view, identifying the dark energy with a cosmological constant term Λ fits well the observations, but has been questioned in terms of naturalness and interpretation in terms of energy density of the vacuum (see Martin, 2012, for a recent review). Theoretical estimates for the vacuum density are many orders of magnitude larger than its observed value. In addition, Ω_Λ and Ω_m are of the same order of magnitude only at present, which marks our epoch as a special time in the evolution of the Universe (the "coincidence problem"). This lack of a clear theoretical understanding has motivated the development of a wide variety of alternative models, as described in Sec. 2.5.

The fine-tuning problem

A full solution of the cosmological constant problem would require a clear connection between Quantum Field Theory and General Relativity. From the first Friedmann equation Eq.2.4 follows that the cosmological constant is of the order of the square of the Hubble parameter today:

$$\Lambda \approx H_0^2 = (2.1h \times 10^{-42}\mathrm{GeV})^2 \tag{2.26}$$

corresponding to an energy density $\rho_\Lambda = \Lambda/8\pi G$ of:

$$\rho_\Lambda \approx 10^{-47}\mathrm{GeV}^4 \tag{2.27}$$

where $1/G = m_{\rm Pl}$ and the Planck mass is equal to $m_{\rm Pl} = 10^{19}$ GeV. If the vacuum energy density comes from the zero point energy of a single field with mass m and momentum k, with energy $E = \sqrt{k^2 + m^2}/2$, summing contributions from all momenta up to a cut-off scale $k_{\rm max}$ we obtain:

$$\langle \rho_{\rm vac} \rangle = \int_0^{k_{\rm max}} \frac{4\pi k^2 {\rm d}k}{(2\pi)^3} \frac{\sqrt{k^2 + m^2}}{2} \approx \frac{k_{\rm max}^4}{16\pi^2}, \tag{2.28}$$

since the integral will be dominated by large modes ($k \gg m$). The Planck mass $m_{\rm Pl}$ represents the limit of validity of General Relativity, therefore it appears sensible to use this value as a limit of integration for the integral. Doing so we find $\langle \rho_{\rm vac} \rangle \simeq 10^{74}{\rm GeV}^4$, which is 10^{121} times larger than the value in Eq. 2.27. This is a "fine-tuning" problem: we need a cancellation spanning 120 orders of magnitude to match the measured and the predicted value of Λ. Even though this was just a rough estimate, neglecting e.g. the equation of state of vacuum energy, performing the calculations more carefully does not alter the main conclusion: in the absence of a powerful symmetry able to cancel vacuum fluctuations, such as supersymmetry (e.g. Shadmi, 2017), an enormous fine tuning is required order by order in perturbation theory.

The coincidence problem

Given the progression of dominating contributors to the energy budget, it seems fortuitous that the influence of the cosmological constant is in fact comparable to that of matter at the present time despite the rapid transition. The redshift at which the energy densities of matter and cosmological costant coincide ($z_{\rm co}$) is:

$$z_{\rm co} = \left(\frac{\Omega_\Lambda}{1 - \Omega_\Lambda} \right)^{\frac{1}{3}} - 1 \tag{2.29}$$

Using a value for today's density $\Omega_\Lambda = 0.7$, the coincidence redshift is $z_{\rm co} \approx 0.3$. This number represent a very recent epoch, moreover it depends strongly on the ratio of Ω_Λ/Ω_m: if this ratio was just 10 times smaller or larger, accelerated expansion would not even be observed today.

Despite the tremendous observational progress in measuring dark energy properties, no fundamentally new insights into the physics behind this mysterious component have resulted. Remarkably, while the error bars have shrunk dramatically, current constraints are still roughly consistent with the specific model that was originally quoted as the best fit in the late 1990s — a component contributing about 70% to the current energy budget with an equation-of-state ratio $w \simeq -1$. This has led some in the particle physics and cosmology community to suspect that dark energy really is just the cosmological constant Λ and that its unnaturally small value is the product of a multiverse, such as would arise from the framework of eternal inflation or from the landscape picture of string theory, which generically features an enormous number of vacua, each with a different value for Λ. In this picture, we live in a vacuum which is able to support stars, galaxies, and life, making our tiny Λ a necessity rather than an accident or a signature of new physics. As such reasoning may be untestable and therefore arguably unscientific, many remain hopeful that cosmic acceleration can be explained by testable physical theory that does not invoke the anthropic principle. For now, improved measurements provide by far the best opportunity to better understand the physics behind the accelerating universe.

2.4.1. A distinction between dark energy and modified gravity

Alternatives to the cosmological constant can be generally grouped into two main categories. Either dark energy is a modification of gravity on the largest scales ("modified gravity" theories), or it is given by a scalar field that effectively behaves as a fluid with negative pressure (usually referred to as proper "dark energy" models). The distinction between the two scenarios is not rigidly set and here

we report only one possible way of operating such distinction, recently put forward by Joyce et al. (2016).

This distinction relies on the motion of bodies in the theory considered. To begin with, we recall the *weak equivalence principle* (WEP), which states that there exists some metric to which all matter species couple universally. Then, test bodies—regardless of their composition—fall along geodesics of this metric. This is usually stated as the equivalence of inertial and gravitational mass.

To distinguish dark energy from modified gravity, Joyce et al. (2016) further invoke the *strong equivalence principle* (SEP). The SEP extends the universality of free fall to massive bodies, i.e. to be completely independent of a body's composition, including gravitational binding energy, so compact objects like black holes also follow geodesics. Following Joyce et al. (2016) we can call anything which obeys the SEP dark energy, and anything which does not, modified gravity. The motivation for this definition is to classify models which influence ordinary matter only gravitationally as dark energy. In these models, the force felt between two bodies is only that of General Relativity (and possibly other Standard Model forces). However, in models of modified gravity, bodies may carry additional charges (e.g. scalar charge) which leads to them experiencing an additional force beyond that of gravity. The appeal to the SEP is an attempt to make this intuition precise. A theoretical motivation for this distinction based upon the SEP is that it is believed—though not proven—that General Relativity is the only metric theory which obeys the SEP (Will, 2014).

2.5. Theories of gravity and dark energy

The quest to test gravity and find alternatives to the cosmological constant has produced many theories beyond Einstein's General Relativity (GR) and other descriptions of gravity on cosmological scales. Those models which are close to ΛCDM are in broad agreement with current constraints on the background cosmology, but the perturbations may still evolve differently. Here we will classify the different means to modify Einstein's theory.

The starting point for a fully covariant extended theory of gravity is a generalization of the Einstein-Hilbert action

$$S_{GR} = \int d^4x \sqrt{-g} \frac{R[g_{\mu\nu}]}{16\pi G} + S_m[g_{\mu\nu}, \cdots], \tag{2.30}$$

where G is Newton's constant and S_m denotes the action of matter, universally and minimally coupled to the metric $g_{\mu\nu}$. Variation of the action (2.30) with respect to the metric leads to Einstein's field equations Eq.2.2.

At the classical level, the results of Lovelock imply that the Einstein-Hilbert action is unique in 4D (Lovelock, 1971). It follows that alternative theories of gravity can be classified into those that

- Break the fundamental assumptions.
- Include additional fields.
- Make the graviton massive.

Theories may even fall within several of these categories. For instance: bimetric gravity has an additional field (tensor) and contains a massive graviton, Einstein-Aether is both Lorentz-violating and includes a vector field, TeVeS has a scalar in addition to a vector, and many extra-dimensional models can be described in terms of additional fields in certain limits.

The most important case for our work is the addition of a scalar field: in Chapters 5 and 6 we will present constraints on a very large class of these so-called scalar-tensor theories, the Horndeski class.

2.5.1. Breaking fundamental assumptions

Lovelock's theorem assumes a four dimensional pseudo-Riemannian manifold and local interactions satisfying Lorentz invariance. Any departure from these principles can be the basis for the construction of a new modified theory of gravity.[5]

Extra dimensions: Additional spatial dimensions allow the inclusion of new operators constructed only from the metric tensor, such as the Gauss-Bonnet term (a topological term in 4 dimensions which does not contribute to the equations of motion). The lack of observation of extra dimensions requires however some mechanism to hide them. One example is compactification, when extra dimensions are sufficiently small that they are not accessible to experimental tests Overduin & Cooperstock (1998).

Lorentz Invariance Violation In many of these alternatives to General Relativity Lorentz invariance is broken by letting a preferred time direction emerge spontaneously (see Blas & Lim (2014) for a review). Horava gravity (Horava, 2009), for example, has a preferred foliation of space-time which implies violation of Lorentz invariance, but interestingly Lorentz symmetry can be recovered at low energies.

Non-local theories Non-local theories include inverse powers of the Laplacian operator in the action. These models can involve general functions (e.g. $R \cdot f(\Box^{-1}R)$, Deser & Woodard, 2007) or be linear (e.g. $R\frac{m^2}{\Box^2}R$, Jaccard et al., 2013). Non-local models are constructed using the Ricci scalar, since non-local terms involving contractions of the Ricci tensor give rise to cosmological instabilities (Nersisyan et al., 2017).

2.5.2. Additional fields

Gravity can be extended by the inclusion of additional fields that interact directly with the metric. These theories will vary according to which type of field is added (scalar, vector, tensor) and its interaction with gravity. Since theories with additional tensors (bigravity and multigravity) are extensions of massive gravity, we will group all of these in Sec. 2.5.2. We will assume a minimal universal coupling of matter to the metric. For a very complete review of gravity theories containing additional fields, see Heisenberg (2018).

Additional scalar field

Since a scalar field does not have a preferred orientation, a classical state can exist in the universe without affecting the isotropy of the space-time if it depends only on time: for this reason scalar fields are the simplest and most studied extensions to General Relativity. A potential term for the scalar can mimic a cosmological constant very closely in the limit in which the field is varying very slowly (e.g. if the potential is very flat), which is the foundation of the simplest single-field inflation and dark energy models (quintessence). We notice that scalar fields may also arise as limits of more complex theories, in other words as effective descriptions of more fundamental theories belonging to other categories, such as braneworld (de Rham et al., 2011).

Ostrogradski's theorem states that theories with second and higher (time) derivatives in the action generically introduce unstable degrees of freedom (Woodard, 2015). This leads to a classification (Ezquiaga & Zumalacárregui, 2018) based on the highest-order derivatives of the additional field present in the action and the equations of motion, with three generations of theories. This way of

[5] GR extensions including additional geometric elements like torsion or non-metricity can be viewed as either breaking the fundamental assumptions or including additional fields.

categorising the theories should be interpreted in the sense of subsets: a class of theories with higher derivatives includes the class of theories with lower derivatives.

(1) Simplest scalar tensor theories: 1st order derivatives in the action, 2nd order in equations.

These contain at most first derivatives of the scalar in the action. They can be seen as a generalization of the Jordan-Brans-Dicke theory of gravity (Brans & Dicke, 1961)

$$S = \int d^4x \sqrt{-g} \frac{M_{\rm Pl}^2}{2} \left[\omega(\phi)R - K(X,\phi)\right] + S_m\,, \tag{2.31}$$

where $X \equiv -\nabla^\nu\phi\nabla_\nu\phi/2$ is the canonical kinetic term of the scalar field. This theory includes GR ($\omega = 1, K = \Lambda$), quintessence ($\omega = 1, K = X - V$) (Wetterich, 1988; Ratra & Peebles, 1988), Brans-Dicke models (Brans & Dicke, 1961) ($\omega = \phi$, $K = \frac{\omega_{\rm BD}}{\phi}X - V(\phi)$), k-essence (Armendáriz-Picón et al., 1999) ($\omega = 1$, $K = K(\phi, X)$). Archetypical modified-gravity models such as $f(R)$ (Carroll et al., 2004) are equivalent to instances of these theories. Chameleons and symmetrons (Hinterbichler & Khoury, 2010) also belong to this class of theories.

(2) Horndeski theories (Horndeski, 1974): 2nd order derivatives in the action and 2nd order in equations.

The Horndeski class of modified gravity theories represents an example of a remarkably large set of extensions to General Relativity, and the most important for this thesis. First discussed in 1974 by Horndeski (Horndeski, 1974) and subsequently rediscovered in Nicolis et al. (2009) and Deffayet et al. (2011), the Horndeski Lagrange density is the most general way of writing the Lagrangian of a scalar-tensor theory of gravity that is four-dimensional, Lorentz-invariant, local, and has equations of motion with derivatives not higher than second order. This ensures the safety of the theory against Ostrogradski instabilities and subsequent ghost degrees of freedom (Woodard, 2007). The Horndeski action can be written as follows:

$$\begin{aligned}
S[g_{\mu\nu},\phi] &= \int \mathrm{d}^4x\sqrt{-g}\left[\sum_{i=2}^{5}\frac{1}{8\pi G_N}\mathcal{L}_i[g_{\mu\nu},\phi] + \mathcal{L}_m[g_{\mu\nu},\psi_M]\right] \\
\mathcal{L}_2 &= G_2(\phi,X), \\
\mathcal{L}_3 &= -G_3(\phi,X)\Box\phi, \\
\mathcal{L}_4 &= G_4(\phi,X)R + G_{4X}(\phi,X)\left[(\Box\phi)^2 - \phi_{;\mu\nu}\phi^{;\mu\nu}\right], \\
\mathcal{L}_5 &= G_5(\phi,X)G_{\mu\nu}\phi^{;\mu\nu} \\
&\quad - \frac{1}{6}G_{5X}(\phi,X)\left[(\Box\phi)^3 + 2\phi_{;\mu}{}^{\nu}\phi_{;\nu}{}^{\alpha}\phi_{;\alpha}{}^{\mu} - 3\phi_{;\mu\nu}\phi^{;\mu\nu}\Box\phi\right].
\end{aligned} \tag{2.32}$$

The four contributions $\mathcal{L}_i$ of the gravitational sector depend on arbitrary functions of the metric $g_{\mu\nu}$ and the kinetic term $K = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi$ of the additional scalar degree of freedom ϕ. The subscripts ϕ, X denote partial derivatives, e.g. $G_{iX} = \frac{\partial G_i}{\partial X}$. We write the normalization of the G_i functions following the convention implemented in the `hi_class` code (Zumalacárregui et al., 2017). We will consider only universal coupling between the metric and the matter fields (collectively described by ψ_m and contained in the matter Lagrangian $\mathcal{L}_m$), which are therefore uncoupled to the scalar field. Most of the universally coupled models with one scalar degree of freedom belong to the Horndeski class. These include for example quintessence (Wetterich, 1988; Ratra & Peebles, 1988), Brans-Dicke models (Brans & Dicke, 1961), k-essence (Armendáriz-Picón et al., 1999; Armendariz-Picon et al., 2001), kinetic gravity braiding (Deffayet et al., 2010; Kobayashi et al., 2010; Pujolàs et al., 2011), covariant galileons (Nicolis et al., 2009; Deffayet et al., 2009), disformal and Dirac-Born-Infeld gravity (de Rham & Gabadadze,

2010; Zumalacárregui et al., 2013; Bettoni & Liberati, 2013), Chameleons (Khoury & Weltman, 2004; Khoury & Weltman, 2004), symmetrons (Hinterbichler & Khoury, 2010; Hinterbichler et al., 2011), Gauss-Bonnet couplings (Ezquiaga et al., 2016) and models screening the cosmological constant (Charmousis et al., 2012; Martín-Moruno et al., 2015). Archetypical modified gravity-models such as all variants of $f(R)$ (Carroll et al., 2004) and $f(G)$ (Carroll et al., 2005) theories are also included. Models that are not within this broad class are those that contain higher derivatives in the equations of motion (Zumalacárregui & García-Bellido, 2014; Gleyzes et al., 2015), and modifications of gravity with non-scalar degrees of freedom, e.g. Einstein-Aether models (Jacobson & Mattingly, 2001) or ghost-free massive gravity (de Rham & Tolley, 2010; de Rham et al., 2011; Hassan & Rosen, 2012). The choice of the $G_i(g_{\mu\nu}, K)$ functions completely specifies the single modified gravity model that one considers.

(3) Beyond Horndeski: 2^{nd} order derivatives in the action and higher order in equations.

Theories beyond Horndeski have higher order equations of motion without including additional degrees of freedom. The first examples of these theories (Zumalacárregui & García-Bellido, 2014) were related to General Relativity by a metric redefinition involving derivatives of the scalar field (Bekenstein, 1993). The best known beyond Horndeski theory is given by the Gleyzes-Langlois-Piazza-Vernizzi (GLPV) action (Gleyzes et al., 2015).

The study of degeneracy conditions for scalar-tensor theories ultimately lead to the *degenerate higher-order scalar-tensor* (DHOST) (Langlois & Noui, 2016) paradigm classification of theories with the right number of degrees of freedom (also known as Extended Scalar-Tensor or EST) (Crisostomi et al., 2016).

Additional vector field

A background vector field does not satisfy the isotropy requirements of the cosmological background, unless it points in the time direction and only depends on time $A_\mu = (A_0(t), 0, 0, 0)$. Isotropy can also happen on average, if a vector with a space-like projection oscillates much faster than the Hubble time (Cembranos et al., 2012). In that case the background is isotropic on average but the perturbations (including gravitational waves) inherit a residual anisotropy (Cembranos et al., 2017). A large number of vectors can also lead to statistical isotropy (e.g. if the orientations are random) (Golovnev et al., 2008).

Additional tensor fields and massive gravity

Giving a mass to the graviton is another means to extend General Relativity, with gravity mediated by a particle with mass m_g, spin $s = 2$ and $2s + 1 = 5$ polarization states (see de Rham et al. (2017) for bounds on the graviton mass). The linear theory of massive gravity was formulated in 1939 by Fierz & Pauli (Fierz & Pauli, 1939) It was later found that Fierz-Pauli theory was discontinuous and gave different results from General Relativity in the limit where massive gravity should have recovered Einstein's theory. Considering non-linear interactions solved the apparent discontinuity, but made the theory affected by ghost instability. The apparent difficulties were overcome in de Rham-Gabadadze-Tolley theory (dRGT) (de Rham & Gabadadze, 2010), also known as ghost-free massive gravity (for current reviews on the theory see Hinterbichler (2012); de Rham (2014)).

In order to write a mass term for the metric, dRGT incorporates an additional, non-dynamical tensor. Massive gravity can be extended by including a kinetic term to the auxiliary metric, which becomes fully dynamical. This leads to the theory of bigravity (or bimetric gravity) (Hassan & Rosen, 2012), which contains two spin-2 particles: one massive and one massless. The same procedure can be extended to more than two interacting metrics, leading to multigravity theories (Hinterbichler & Rosen, 2012).

2.6. Parameterizations of dark energy

Given the lack of a consensus model for cosmic acceleration, it is a challenge to provide a simple yet unbiased and sufficiently general description of dark energy. The equation-of-state parameter w has traditionally been identified as one useful phenomenological description; being the ratio of pressure to energy density, it is also closely connected to the underlying physics. Many more general parameterizations exist, some of them with appealing statistical properties. We now review a variety of formalisms that have been used to describe and constrain dark energy (e.g. in Planck Collaboration et al., 2016b).

(1) **Background parameterizations**. In this approach the only quantities to be parameterized are the background-level ones. Perturbations are always included, but their evolution depends only on the background.

(2) **Perturbation parameterizations**. Here the perturbations too are parameterized. There are two main approaches within this category: effective field theory for DE (EFT, e.g. Gubitosi et al., 2013; Bloomfield et al., 2013; Gleyzes et al., 2013), which contains all symmetry operators in the Lagrangian, written in unitary gauge, i.e. in terms of metric perturbations only and therefore is theoretically well motivated. A second approach includes a more phenomenological class of models obtained by directly parameterizing two independent functions of the gravitational potentials. This approach can in principle probe *all* degrees of freedom at the background and perturbation level (e.g. Kunz, 2012) and is easier to handle in numerical codes. While the connection to physical models is less obvious here than in EFT, this approach allows us to gain a more intuitive understanding of the general constraining power of the data.

2.6.1. Background parameterizations

Constant equation of state

Assuming that dark energy is spatially smooth, its simplest parametrization is in terms of its equation-of-state

$$w \equiv \frac{p_{\rm de}}{\rho_{\rm de}c^2} = \text{constant}. \tag{2.33}$$

Vacuum energy is represented by the choice $w = -1$. Together with $\Omega_{\rm de}$, w provides a two-parameter description of the dark-energy sector.

Varying equation of state

Promoting either the dark energy density or the equation of state to a general function of redshift — $\Omega_{\rm de}(z)$ or $w(z)$ — would be the most general way to describe dark energy, still assuming its spatial homogeneity. In practice, however, either of these functions formally corresponds to infinitely many parameters to measure, and measuring even a few such parameters is a challenge. Perhaps not surprisingly, therefore, the most popular parametrizations of w have involved two free parameters. Other low-dimensional parametrizations have been proposed (Gerke & Efstathiou, 2002); for low redshift they are all essentially equivalent, but for large z they lead to different and often unphysical behavior. The parametrization (Linder & Jenkins, 2003; Chevallier & Polarski, 2001)

$$w(a) = w_0 + w_a(1-a) = w_0 + w_a\,\frac{z}{1+z}\,, \tag{2.34}$$

where $a = 1/(1+z)$ is the scale factor, avoids this problem, and it fits many scalar field and some modified gravity expansion histories. This therefore leads to the most commonly used description of dark energy, namely the three-parameter set $\{\Omega_{\rm de}, w_0, w_a\}$.

Principal components

The cosmological function that we would like to determine — $w(z)$, $\rho_{\rm de}(z)$, or $H(z)$ — can be expanded in terms of principal components, a set of functions that are uncorrelated and orthogonal by construction (Huterer & Starkman, 2003). In this approach, the data determine which parameters are measured best.

Suppose we parametrize $w(z)$ in terms of piecewise constant values w_i ($i = 1, \ldots, N$), each defined over a narrow redshift range $z_i < z < z_i + \Delta z$). In the limit of small Δz this recovers the shape of an arbitrary dark energy history), but the estimates of the w_i from a given dark energy probe will be very noisy. Principal component analysis (PCA) extracts from those noisy estimates the best-measured features of $w(z)$. One finds the eigenvectors $e_i(z)$ of the inverse covariance matrix for the parameters w_i and the corresponding eigenvalues λ_i. The equation-of-state parameter is then expressed as

$$1 + w(z) = \sum_{i=1}^{N} \alpha_i \, e_i(z) \, , \tag{2.35}$$

Generalized dark energy phenomenology

The simplest and by far the most studied class of models is dark energy that is spatially smooth and its only degree of freedom is its energy density — that is, it is fully described by either $\rho_{\rm de}(a)$ or $w(a) \sim -1$.

Another possibility is that dark energy has the speed of sound that allows clustering at sub-horizon scales, that is, $c_s^2 \equiv \delta p_{\rm de}/\delta\rho_{\rm de} < 1$ (where c_s is quoted in units of the speed of light). Unfortunately, the effects of the speed of sound are small, and become essentially negligible in the limit when the equation of state of dark energy w becomes close to -1, and are difficult to discern with late-universe measurements even if w deviates from the cosmological constant value at some epoch. It will therefore be essentially impossible to measure the speed of sound even with future surveys.

Another possibility is the presence of "early dark energy" (Wetterich, 2004; Pettorino et al., 2013), component that is non-negligible at early times, typically around recombination or even earlier. The early component is motivated by various theoretical models, and could imprint signatures via the early-time Integrated Sachs-Wolfe effect.

Finally, there is a possibility that dark energy is coupled to dark matter (breaking adiabaticity), or other components or particles (Amendola, 2000). This is a much richer — though typically very model-dependent — set of possibilities, with many opportunities to test them using data.

As yet, there is no observational evidence for generalized dark energy beyond the simplest model but, as with modified gravity, studying these extensions is important to understand how dark energy phenomenology can be searched for by cosmological probes.

2.6.2. Parameterizations of cosmological gravity at the perturbations level

A great disadvantage of the full covariant approach described in Sec.2.5 is that the predictions for every model/theory have to be obtained from scratch, which makes the exploration of the theory space a daunting task.

An alternative route is to constrain deviations from General Relativity, without reference to any fundamental theory. The tradeoff is to keep the theory of gravity as general as possible at the expense of dealing with a very simple space-time. The simplest situation is where the background space-time is flat and maximally symmetric (Minkowski), a setup useful to model gravity in the Solar System. In this simple case one can define a series of quantities, known as Parameterized Post-Newtonian (PPN) coefficients, that describe general modifications of gravity over Minkowski space (see Will (2014) for

details, including constraints and additional assumptions). These PPN parameters can be constrained by experiments (such as the deflection of light by massive bodies) and computed for any theory, and thus provide a very efficient phenomenological dictionary.

When dealing with linear perturbations acting on a Friedmann-Robertson-Walker metric in modified gravity one can assume spatial flatness and, considering only scalar perturbations (see Durrer & Tansella, 2016; Adamek et al., 2016, for vector and tensor perturbations), write the line element in Newtonian gauge as

$$\mathrm{d}s^2 = \left(1 + 2\frac{\Phi}{c^2}\right)c^2\mathrm{d}t^2 - a^2(t)\left(1 - 2\frac{\Psi}{c^2}\right)\left(\mathrm{d}\chi^2 + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\varphi^2\right) \tag{2.36}$$

with the Bardeen potentials Φ and Ψ. In General Relativity $\Phi = \Psi$ in absence of anisotropic stress, while this is in general not true in modified gravity. The time-evolution of the cosmological background makes an extension of PPN approach to cosmology a difficult task, as instead of constant coefficients one needs to deal with functions of time due to the evolution of the universe.

Describing the perturbations requires more functional freedom than for the background. Here we will review two common procedures, namely the effective theory of dark energy and the modified gravitational "constants".

Effective theory of dark energy

The effective (field) theory of dark energy (Gubitosi et al., 2013; Bloomfield et al., 2013; Gleyzes et al., 2013) can be used to systematically describe general theories of gravity over a cosmological background. The original formulation applies to theories with a scalar field ϕ and uses the unitary "gauge": a redefinition of the time coordinate as the constant ϕ hypersurfaces (this is always possible if $\phi_{,\mu}$ is time-like and non-degenerate, as in perturbed cosmological backgrounds, but not in general). One then constructs all the operators compatible with the symmetries of the background (recalling that the time translation invariance is broken by the cosmological evolution).

The action obtained with this procedure reads:

$$\begin{aligned} S \quad = \quad & \int d^4x\sqrt{-g}\left\{\frac{m_0^2}{2}\left[1+\Omega(\tau)\right]R + \Lambda(\tau) - a^2c(\tau)\delta g^{00}\right. \\ & + \frac{M_2^4(\tau)}{2}\left(a^2\delta g^{00}\right)^2 - \bar{M}_1^3(\tau)2a^2\delta g^{00}\delta K^\mu_\mu \\ & - \frac{\bar{M}_2^2(\tau)}{2}\left(\delta K^\mu_\mu\right)^2 - \frac{\bar{M}_3^2(\tau)}{2}\delta K^\mu_\nu\delta K^\nu_\mu + \frac{a^2\hat{M}^2(\tau)}{2}\delta g^{00}\delta R^{(3)} \\ & \left. + m_2^2(\tau)\left(g^{\mu\nu}+n^\mu n^\nu\right)\partial_\mu\left(a^2g^{00}\right)\partial_\nu\left(a^2g^{00}\right)\right\} \\ & + S_{\mathrm{m}}[\chi_i, g_{\mu\nu}]. \end{aligned} \tag{2.37}$$

Here R is the Ricci scalar, $\delta R^{(3)}$ is its spatial perturbation, K^μ_ν is the extrinsic curvature, and m_0 is the bare (reduced) Planck mass. The matter part of the action, S_{m}, includes all fluid components except dark energy, i.e., baryons, cold dark matter, radiation, and neutrinos. The action in Eq. 2.37 depends on nine time-dependent functions, here $\{\Omega, c, \Lambda, \bar{M}_1^3, \bar{M}_2^4, \bar{M}_3^2, M_2^4, \hat{M}^2, m_2^2\}$, whose choice specifies the theory. In this way, EFT provides a direct link to any scalar field theory.

In Gleyzes et al. (2013) and Bellini & Sawicki (2014) it has been shown that one can parametrize the evolution of linear cosmological perturbations in Horndeski theories by means of four functions of (conformal) time only, which we will collectively refer to here as α functions. Each of them carries a physical meaning, which we describe briefly here, referring to Bellini & Sawicki (2014) and references therein for a more complete description:

- α_K is the *kineticity* term, i.e. the kinetic energy of the scalar perturbations arising directly from the action. Increasing this term suppresses the sound speed of scalar perturbations. This makes the sound horizon smaller than the cosmological horizon, allowing the scalar field to enter a quasi-static configuration on smaller scales, below the sound horizon (Sawicki & Bellini, 2015). In the quasi-static approximation where time derivatives are considered to be sub-dominant with respect to space derivatives, α_K does not enter the equations of motion and is therefore largely unconstrained by cosmic shear (Bellini et al., 2016; Alonso et al., 2016, although see Kreisch & Komatsu, 2017).

- α_B is the *braiding* term, which describes mixing of the scalar field with the metric kinetic term, leading to what is typically interpreted as a fifth force between massive particles.

- α_M is the *Planck-mass run rate*, defined by

$$\alpha_M \equiv \frac{\mathrm{d}\ln M_*^2}{\mathrm{d}\ln a}, \tag{2.38}$$

  where M_*^2 is the dimensionless product of the normalization of the kinetic term for gravitons and $8\pi G_N$ measured on Earth. This function describes the rate of evolution of the effective Planck mass.

- α_T is the *tensor speed excess,* indicating deviations from the speed of light in the propagation speed of gravitational waves. This can lead to anisotropic stress even in the absence of scalar field perturbations, as a result of a change in the response of the Newtonian potential to matter sources. Recently, very strong constraints have been placed on α_T by the measurement of the gravitational waves speed derived by the detection of the binary neutron star merger GW170817 and the gamma ray burst GRB170817A (Abbott et al., 2017a,b; Baker et al., 2017; Creminelli & Vernizzi, 2017; Ezquiaga & Zumalacárregui, 2017; Sakstein & Jain, 2017; Lombriser & Lima, 2017; Bettoni et al., 2017). Since the speed has been found to be very close to the speed of light, α_T has been consequently constrained to be very close to zero at the present time. We remark that the other three functions (as well as α_T's past value), are instead still free to vary. Ezquiaga & Zumalacárregui (2017) identify the models within the Horndeski classes that are still viable after GW170817; Peirone et al. (2018) show that, even with the strict bound on the present-day gravitational wave speed, there is still room within Horndeski theories for nontrivial signatures of modified gravity that can be measured at the level of linear perturbations.

The specific model considered within the Horndeski class is defined by the choice of the α_i functions. The ΛCDM model corresponds to the choice $\alpha_K = \alpha_B = \alpha_M = \alpha_T = 0$. Once the α functions are set, Bellini & Sawicki (2014) show that it is sufficient to solve the equations of motion for the background and perturbations to fully determine the evolution of linear perturbations at the linear level.

In our work we will need to choose a parametrization for the time evolution of the α functions. Following a common procedure, already implemented in `hi_class`, we will often (though not always: see Sec. 6.2.4) choose to parameterize these functions such that they trace the evolution of the dark energy component, to which they are proportional

$$\boxed{\alpha_i = \hat{\alpha}_i\, \Omega_{\rm DE}(\tau)} \tag{2.39}$$

This choice is the simplest and the most common in the literature (as used e.g. in Planck Collaboration et al., 2016b) and, despite not being the only one, can already provide a lot of information on Horndeski gravity, as remarked by (Gleyzes, 2017) who showed that simple parametrizations are sufficient to describe the theory space in Effective Field Theory of dark energy (Gubitosi et al., 2013; Bloomfield et al., 2013; Gleyzes et al., 2013), which the α parameterization belongs to.

Modified Gravitational "constants"

A very commonly used approach employs general modifications of the equations relating the gravitational potentials to the matter density contrast

$$\nabla^2\Psi = 4\pi G a^2 \mu(t,k)\rho\delta\,, \tag{2.40}$$
$$\nabla^2(\Phi+\Psi) = 8\pi G a^2 \Sigma(t,k)\rho\delta \tag{2.41}$$

(note that different conventions exist in the literature). Here δ is the density contrast in Newtonian gauge and the functions μ, Σ parameterize the evolution of the gravitational potentials as a function of time a and scale k. The functions μ, Σ are often referred to as $G_{\rm matter}$, $G_{\rm light}$ because gradients of Ψ determines the force felt by non-relativistic particles and those of $\Psi + \Phi$ the geodesics of massless particles (and thus the lensing potential). The ratio of the gravitational potentials,

$$\eta \equiv \frac{\Phi}{\Psi} = \frac{2\Sigma}{\mu} - 1\,, \tag{2.42}$$

is of particular interest, since GR predicts that it is exactly one in the absence of radiation and any sizable deviation could be an indication of modified gravity.

This approach has numerous advantages as a test of gravity against data. It is completely theory agnostic, not requiring any information on the ingredients or laws of the theories being tested. Most importantly, it is completely general for universally coupled theories: given any solution $\Delta, \Psi, \Phi(a,k)$ it is possible to obtain μ, Σ through Eqs. 2.40, 2.41. In this sense, any finding of $\mu, \Sigma \neq 1$ might point towards deviations from GR and warrant further investigation.

The main shortcoming of this approach is its great generality: any practical attempt to implement Eqs. 2.40, 2.41 requires a discretization of the functional space, introducing $2 \cdot N_k \cdot N_z$ free parameters for a homogeneous binning. In contrast, the EFT approach for Horndeski theories requires only $4 \cdot N_z$ parameters, making it a more economic parameterization for all but the simplest scale-dependencies ($N_k = 1, 2$). Capturing the full scale dependence of μ, Σ requires either a large parameter space or assumptions about the k-dependence.

A common practice to overcome this limitation is to choose a functional form for μ, Σ as a function of scale. For Horndeski theories the functional form is a ratio of quadratic polynomials in k (Amendola et al., 2013)

$$\mu = h_1 \frac{1 + h_5 k^2}{1 + h_3 k^2}\,, \quad \eta = h_2 \frac{1 + h_4 k^2}{1 + h_5 k^2}\,, \tag{2.43}$$

for functions h_i that depend on redshift through the theory and the scalar field evolution. The mapping is exact on small scales in which the field dynamics can be neglected, below scalar sound horizon.

Another main shortcoming of the completely general approach is that there is no information from other regimes. The major setback with respect to EFT is the lack of information from gravitational wave observables, while in EFT the tensor and scalar sectors are modified accordingly i.e. GW data restrict the modifications available to scalar perturbations, for instance, theories with $\eta \neq 1$ require either α_M or α_T to be non-zero. Attempts to explore the connections between μ, Σ and the EFT approach in Horndeski-like theories have used very general parameterizations: connecting theoretical viability conditions of the theory with the behavior of μ, η (Perenon et al., 2015), including the case with $\alpha_T = 0$ to address the impact of the GW speed measurement (Peirone et al., 2018). General properties of Horndeski theories could be inferred from detailed measurements of μ, Σ (Pogosian & Silvestri, 2016).

2.7. Screening mechanism

An important ingredient for all modified gravity theories which have a significant impact on cosmological scales is a screening mechanism which effectively screens modifications at small scales to provide agreement with small scale tests of General Relativity such as in the Solar System. Screening mechanisms act as non-linear effects by virtue of which the departures from General Relativity fade in short scales or high-density environments. They can be generally divided into potential and kinetic screening mechanisms. If one uses an additional scalar field ϕ (e.g. in the Brans-Dicke theory), the interaction potential is given by a Yukawa potential with characteristic scale m_ϕ, which gives the range of the interaction. If the mass of the scalar, m_ϕ depends on the environment in such a way that it is large if the density is high, General Relativity will be recovered in these regions. This can be achieved by a suitable choice of the interaction potential of the scalar field, since the mass corresponds to its second derivative. This is called Chameleon mechanism (Khoury & Weltman, 2004). A kinetic screening is for example the Vainshtein mechanism (Babichev & Deffayet, 2013), for which the kinetic term X screens modifications on small scales. The screening feature of modified gravity has been studied in a model-by-model basis using non-linear techniques, but has been often ignored in e.g. forecasts for future experiments. In this thesis, we will use linear cosmological perturbation theory and hence the screening effects have to be included in a phenomenological fashion. In this work we will always model the small-scale recovery of General Relativity through a scale dependence of the α EFT functions for Horndeski gravity or by acting on the effective Newtonian coupling, the gravitational slip and the linear growth factor. Our prescription is certainly not the unique that can be followed: however, we note that this or other prescriptions have to be introduced whenever screened scales are included in the analysis. If they can be properly modeled, non-linear scales contribute greatly to constraint other cosmological parameters, but not accounting for screening can in principle largely overestimate the surveys capacity to test gravity.

Chapter 3

Statistical methods in cosmology

The archetypical question of empirical science is: given a set of observations $\mathbf{d}$ and any information or prejudices we had prior to making these observations (denoted $\mathcal{I}$), what is our degree of belief in some hypothesis $\mathcal{H}$? That is, under the logical interpretation of probabilities, what is the probability $P(\mathcal{H}|\mathbf{d}, \mathcal{I})$? Any testable theory will make predictions about the observations we should expect given a certain hypothesis, i.e., $P(\mathbf{d}|\mathcal{H})$. Hence making statements about our state of belief in a certain hypothesis, given our observations, requires us to take our forward model from theory to data $P(\mathbf{d}|\mathcal{H})$ and solve the inverse problem to obtain $P(\mathcal{H}|\mathbf{d}, \mathcal{I})$. Bayes' theorem is precisely the tool for solving this inverse problem, relating theoretical forward model, prior information and prejudices and inferences:

$$P(\mathcal{H}|\mathbf{d}, \mathcal{I}) = \frac{P(\mathbf{d}|\mathcal{H})P(\mathcal{H}|\mathcal{I})}{P(\mathbf{d})}. \tag{3.1}$$

Bayesian inference encompasses any inference process making use of Bayes' theorem and taking the logical interpretation of probabilities as degrees of plausibility. By allowing us to access the probability of a proposition given our observations, Bayes' theorem is of central importance in drawing clear scientific statements from data.

Inference problems tend to fall into two categories: parameter inference and model comparison. Parameter inference assumes a model $\mathcal{M}$ with a set of parameters θ and asks the question: given my observations, model assumptions and prior beliefs, what are my beliefs about the parameters θ of my model $\mathcal{M}$? Model comparison poses the more abstract question: given two models $\mathcal{M}_A$ and $\mathcal{M}_B$, what is my relative belief in these models in light of my observations (and given my prior relative belief in the two models)? Empirical science can be reduced to making observations and then asking one of these two questions. In this Chapter we will focus on the former, since it is central to the understanding of some aspects of this thesis, but we will also briefly mention the latter in Sec. 3.4.3, as it helps understand one of the sampling methods for the posterior distribution that will be described in the following.

We begin by reviewing the parameter inference process within a Bayesian framework in Sec. 3.1; we then expand in Sec. 3.2 on the posterior distribution, the main target of any Bayesian inference process, describing products that can be obtained from the posterior, such as errors on model parameters. In Sec. 3.3 we comment extensively on an increasingly popular tool in statistical cosmology, the Fisher matrix; finally, in Sec. 3.4 we review some of the sampling methods for the posterior distribution widely used in Bayesian inference.

3.1. Bayesian parameter inference

Parameter inference is concerned with making probabilistic statements about the values of the parameters of a chosen model, given some observations. The ultimate goal of parameter inference is the

distribution of parameter values given the observations, model assumptions and prior beliefs, i.e. the posterior distribution $P(\boldsymbol{\theta}|\mathbf{d}, \mathcal{M})$.

In order to make statements about model parameters $\boldsymbol{\theta}$ from observations $\mathbf{d}$, we must have a forward model that relates the parameters to the data, specifying the probability of obtaining data $\mathbf{d}$ given our model $\mathcal{M}$ and parameters $\boldsymbol{\theta}$, i.e. $P(\mathbf{d}|\boldsymbol{\theta}, \mathcal{M})$. The forward model describes how the data were generated from the theory and measurement process, covering two key elements: (1) what are the physical predictions of the model $\mathcal{M}$ (i.e., the physical theory) relevant for the observations, and (2) how does the measurement process introduce uncertainties into the data (i.e., a physical model for the measurement process/instrument). With a forward model specifying $P(\mathbf{d}|\boldsymbol{\theta}, \mathcal{M})$ in hand, Bayes' theorem allows us to turn this around and find the probability of the parameters given the observations (and model assumptions),

$$\boxed{P(\boldsymbol{\theta}|\mathbf{d}, \mathcal{M}, \mathcal{I}) = \frac{P(\mathbf{d}|\boldsymbol{\theta}, \mathcal{M}) \times P(\boldsymbol{\theta}|\mathcal{M}, \mathcal{I})}{P(\mathbf{d}|\mathcal{M}, \mathcal{I})}} \tag{3.2}$$

The distribution of the data given the parameters (and model) $P(\mathbf{d}|\boldsymbol{\theta}, \mathcal{M})$ is called the likelihood, the prior $P(\boldsymbol{\theta}|\mathcal{M}, \mathcal{I})$ encodes our prior beliefs about the parameters (assuming the model is true), and the normalization constant $P(\mathbf{d}|\mathcal{M}, \mathcal{I})$ is called the evidence and has no dependence on the model parameters (but is important for model comparison, see Sec. 3.4.3). The posterior distribution of the parameters $P(\boldsymbol{\theta}|\mathbf{d}, \mathcal{M}, \mathcal{I})$ is a complete description of our beliefs about the parameter values given the observations and is the ultimate goal of parameter inference. The parameter inference process can hence be summarized as follows:

(1) Write down a forward model $\mathcal{M}$ from theory, and from it derive the likelihood $P(\mathbf{d}|\boldsymbol{\theta}, \mathcal{M})$,

(2) Specify our prior beliefs about the parameters (under the model) $P(\boldsymbol{\theta}|\mathcal{M}, \mathcal{I})$ and use Bayes' theorem to write down the posterior distribution $P(\boldsymbol{\theta}|\mathbf{d}, \mathcal{M})$,

(3) Compute the posterior $P(\boldsymbol{\theta}|\mathbf{d}, \mathcal{M})$, i.e. our degree-of-belief in the parameter values.

Step (1) is the combined effort of theorists developing predictive theories for observable quantities, and experimentalists developing a model for their designed measurement process. Step (2), employing Bayes' theorem, requires the choice of a prior encoding our prior beliefs about the parameters before conducting the new experiment. In the absence of previous data, we can attempt to define an 'uninformative' prior encoding our ignorance of the parameters – typically, a flat prior. If the new observations are successor to a series of previous experiments, the prior may simply be the accumulated inference from all previous data, as we shall see in the folowing.

3.1.1. Updating the probability distribution for a parameter

If we obtain some more information from a new experiment, then we can use Bayes' theorem to update our estimate of the probabilities associated with each parameter. For Bayesian statistics to be logically coherent, it should not matter how we gain our information, meaning that the effect on the probability of the parameters should be the same if we either add the results of a new experiment to the probability of the parameters, or if we carry out the two experiments first, and then see how they both affect the probability of the parameters.

Let us demonstrate that these two interpretations are equivalent. We start with Bayes' expression for the posterior probability of a parameter (or more generally of some hypothesis), where we put explicitly that all probabilities are conditional on some prior information $\mathcal{I}$,

$$p(\boldsymbol{\theta}|\mathbf{d}\mathcal{I}) = \frac{p(\boldsymbol{\theta}|\mathcal{I})p(\mathbf{d}|\boldsymbol{\theta}\mathcal{I})}{p(\mathbf{d}|\mathcal{I})}. \tag{3.3}$$

Let us imagine we carry out a new experiment with new data, $\mathbf{d}'$. We have two ways to analyse the new data:

- Interpretation 1: we regard $\mathbf{d}'$ as the dataset, and $\mathbf{d}\mathcal{I}$ (means $\mathbf{d}$ *and* $\mathcal{I}$) as the new prior information;
- Interpretation 2: we put all the data together, and call it $\mathbf{d}'\mathbf{d}$, and interpret it with the old prior information $\mathcal{I}$.

Let us start with Interpretation 1. We rewrite Bayes' theorem, equation (3.3) by changing datasets $\mathbf{x} \to \mathbf{x}'$, and letting the old data become part of the prior information $I \to \mathcal{I}' = \mathbf{d}\mathcal{I}$. Bayes' theorem is now

$$p(\boldsymbol{\theta}|\mathbf{d}'\mathcal{I}') = \frac{p(\boldsymbol{\theta}|\mathbf{d}\mathcal{I})p(\mathbf{d}'|\boldsymbol{\theta}\mathbf{d}\mathcal{I})}{p(\mathbf{d}'|\mathbf{d}\mathcal{I})}. \tag{3.4}$$

We notice that the new prior in this expression is just the old posterior probability from equation (3.3), and that the new likelihood is just

$$p(\mathbf{d}'|\mathbf{d}\boldsymbol{\theta}\mathcal{I}) = \frac{p(\mathbf{d}'\mathbf{d}|\boldsymbol{\theta}\mathcal{I})}{p(\mathbf{d}|\boldsymbol{\theta}\mathcal{I})}. \tag{3.5}$$

Substituting this expression for the new likelihood:

$$p(\boldsymbol{\theta}|\mathbf{d}\mathcal{I}') = \frac{p(\boldsymbol{\theta}|\mathbf{d}\mathcal{I})p(\mathbf{d}'\mathbf{d}|\boldsymbol{\theta}\mathcal{I})}{p(\mathbf{d}'|\mathbf{d}\mathcal{I})p(\mathbf{d}|\boldsymbol{\theta}\mathcal{I})}. \tag{3.6}$$

Using Bayes' theorem again on the first term on the top and the second on the bottom, we find

$$p(\boldsymbol{\theta}|\mathbf{d}\mathcal{I}') = \frac{p(\boldsymbol{\theta}|\mathcal{I})p(\mathbf{d}'\mathbf{d}|\boldsymbol{\theta}\mathcal{I})}{p(\mathbf{d}'|\mathbf{d}\mathcal{I})p(\mathbf{d}|\mathcal{I})}, \tag{3.7}$$

and simplifying the denominator gives finally

$$p(\boldsymbol{\theta}|\mathbf{d}\mathcal{I}') = \frac{p(\boldsymbol{\theta}|\mathcal{I})p(\mathbf{d}'\mathbf{d}|\boldsymbol{\theta}\mathcal{I})}{p(\mathbf{d}'\mathbf{d}|\mathcal{I})} = p(\boldsymbol{\theta}|\,[\mathbf{d}\mathbf{d}']\mathcal{I}), \tag{3.8}$$

which is Bayes' theorem in Interpretation 2, i.e. it has the same form as equation (3.3), the outcome from the initial experiment, but now with the data $\mathbf{d}$ replaced by $\mathbf{d}'\mathbf{d}$. In other words, we have shown that $\mathbf{d} \to \mathbf{d}'$ and $\mathcal{I} \to \mathbf{x}\mathcal{I}$ is equivalent to $\mathbf{d} \to \mathbf{d}'\mathbf{d}$. This shows us that how we add in new information makes no difference: Bayes' theorem gives us a natural way of improving our statistical inferences as our state of knowledge increases.

3.1.2. Frequentist versus Bayesian approach

Here we want to illustrate some differences and similarities between the frequentist maximum likelihood and the Bayesian approach to inference.

In general, observations in an experiment follow a certain distribution for a fixed parameter value. This distribution is the probability distribution over all possible observation values for the given value of the parameter. We define 'parameter space' the set of all possible parameter values, which has a number of dimensions equal to the number of parameters. We will call 'sample space' the set of all possible values of the observations, whose dimension is the number of observations. We define the 'inference universe' of the problem to be the Cartesian product of parameter space and sample space. It is the dimensional space where the first dimensions are the parameter space, and the remaining dimensions are the sample space. We do not ever observe the parameter, so the position in those coordinates are always unknown. However, we do observe the sample, so we know the last coordinates.

We will let the dimensions be $p = 1$ and $n = 1$ for simplicity, i.e. we consider a single parameter and a single observation. Figs. 3.1, 3.2, and 3.3 are exact in this case, and serve as instructional examples for the higher dimensional case.

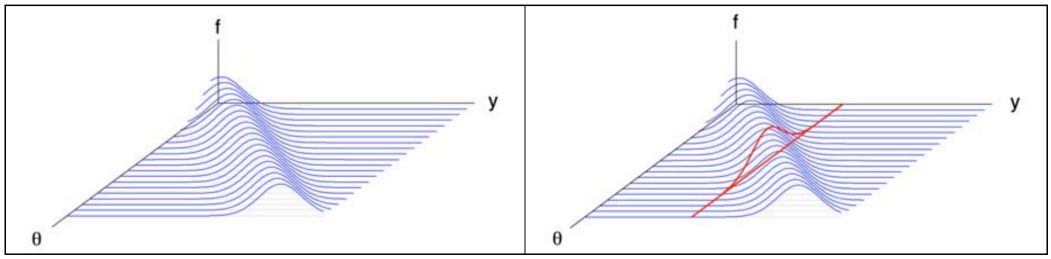

Figure 3.1.: The observation distribution (*left*) and the likelihood function (*right*)

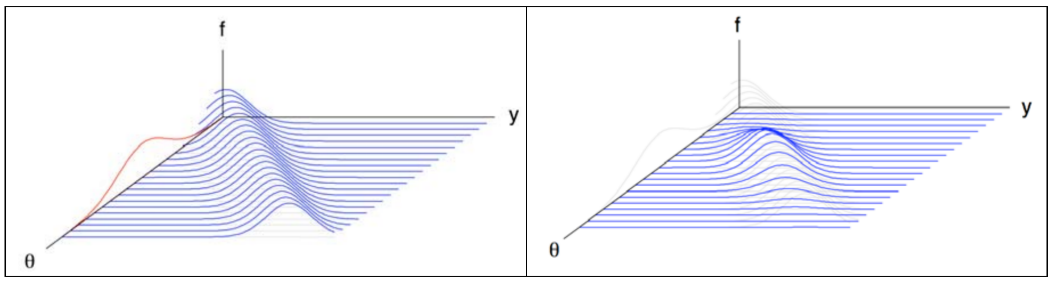

Figure 3.2.: The prior and likelihood (*left*) and the joint density of θ and y (*right*)

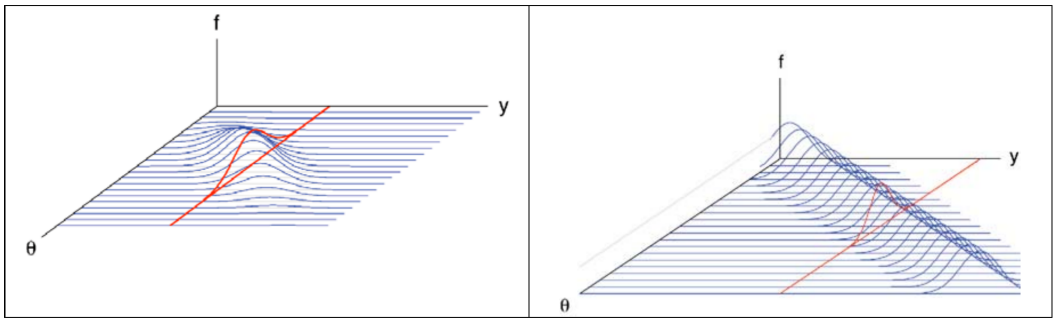

Figure 3.3.: Posterior (*left*) and posterior when flat prior is used (*right*)

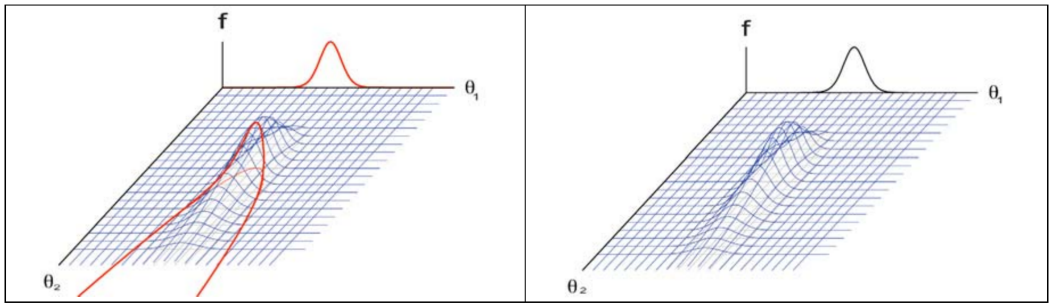

Figure 3.4.: Profile likelihood (*left*) and marginal posterior (*right*) for θ_1

Frequentist Maximum Likelihood estimation

The sampling distribution $f(y|\theta)$ is a function of both the value of the observation and the parameter value. Given the value θ, it gives the probability distribution of the observation y. It is defined for all points in the inference universe. Thus, it forms a surface defined on the inference universe. It is a probability distribution in the observation dimension for each particular value of the parameter. However, it is not a probability distribution in the parameter dimension. The first panel of Fig.3.1 shows the sampling distribution surface in 3D perspective. The likelihood function has the same functional form as the sampling distribution, but y is held at the observed value and θ is allowed to vary over all possible values. It is found by cutting the sampling distribution surface with a vertical plane parallel to the axis through the observed value, as shown in the second panel of Fig.3.1. Likelihood inference is based on the likelihood function. Since it is not a probability density, in the frequentist approach the best estimator of the parameter is taken to be the value that has the highest value of the likelihood function.

Bayesian estimation

Bayesian estimation requires that we have a probability distribution defined on the parameter space before we look at the data: the prior distribution. It gives our belief weights for each of the possible parameter values before we see the data. This requires that we allow a different interpretation of probability on parameter space than on sample space. Probability on sample space has the usual interpretation of relative frequency, while on parameter space it has a degree-of-belief interpretation. The prior distribution of the parameter is shown with the sampling distribution surface in the first panel of Fig.3.2. The joint distribution of the parameter and the observation is found by multiplying each value of the sampling distribution surface by the corresponding height of the prior distribution. This is shown in the second panel of Fig.3.2. To find the posterior distribution of the parameter given the observed value we cut the joint distribution of the parameter and the observation with a vertical plane parallel to the parameter axis through the observed value of y. This is shown in the first panel of Fig.3.3. The posterior distribution summarizes the belief we can have about all possible parameter values, given the observed data. It will always be a probability distribution, conditional on the observed data.

Using a flat prior the posterior has same shape as the likelihood

If we decide to use a flat prior distribution that gives equal weight to all values of the parameters, the joint distribution on the inference universe will be the same as the sampling distribution surface. This is shown in the second panel of Fig.3.3.

3.1.3. Nuisance parameters, marginalization

In many cases we are forced to infer parameters from the data that we are both uncertain about and/or not interested in, which nonetheless impact our analysis; these are the so–called nuisance parameters. We want to make sure that the nuisance parameters do not interfere with our inference on the parameter of interest. Since, in the Bayesian approach, the joint posterior distribution is a probability distribution, while using the likelihood approach the joint likelihood function is not a probability distribution, the two approaches have different ways of dealing with nuisance parameters, even when we use independent joint flat priors and the posterior distribution and likelihood function have the same shape.

Likelihood inference in the presence of nuisance parameters

For instance, suppose that θ_1 is the parameter of interest, and θ_2 is a nuisance parameter. Frequentist likelihood inference on θ_1 is often based on the profile likelihood function given by:

$$L_p(\theta_1|\mathbf{d}) = \max_{\theta_2} L(\theta_1, \theta_2|\mathbf{d}) \tag{3.9}$$

where $L(\theta_1, \theta_2|\mathbf{d})$ is the joint likelihood function. Essentially, the nuisance parameter has been eliminated by plugging $\hat{\theta}_2|\theta_1$, to be interpreted as the conditional maximum likelihood value of θ_2 given θ_1, into the joint likelihood. Hence

$$L_p(\theta_1|\mathbf{d}) = L(\theta_1, \hat{\theta}_2|\theta_1 \mid \mathbf{d}). \tag{3.10}$$

This is shown in the left panel of Fig.3.4. The profile likelihood function may lose some information about θ_1, compared to the joint likelihood function. Note that the maximum profile likelihood value of θ_1 will be the same as its maximum likelihood value. However confidence intervals based on profile likelihood may not be the same as those based on the joint likelihood.

Bayesian Inference in the presence of nuisance parameters

Bayesian statistics has a single way of dealing with nuisance parameters. Uncertainties in nuisance parameters can be formally propagated through the inference process by jointly inferring the parameters of interest θ_1 and nuisance parameters θ_2 together, and then marginalising (integrating) over the nuisance parameters, i.e.,

$$P(\theta_1|\mathbf{d}) = \int P(\theta_1, \theta_2|\mathbf{d}) \mathrm{d}\theta_2 = \int \frac{P(\mathbf{d}|\theta_1, \theta_2) P(\theta_1, \theta_2|\mathcal{M}) \mathrm{d}\theta_2}{P(\mathbf{d}|M)} \tag{3.11}$$

where we used Bayes' theorem in the second equality. In this way our prior beliefs and uncertainties on θ_2 have been updated in light of the new observations and explicitly propagated to the inference on θ_1, where the "prior information" $\mathcal{I}$ implicit on the right hand side of the posterior $P(\theta_1, \theta_2|\mathbf{d}, \mathcal{M}, \mathcal{I})$ explicitly includes our prior beliefs (uncertainties) about the nuisance parameters. The marginal posterior is found by integrating the nuisance parameter out of the joint posterior, a process referred to as marginalization.

3.2. Posterior interpretation

The interpretation of a posterior density is straightforward: $P(\theta|\mathbf{d}, \mathcal{M}, \mathcal{I})\mathrm{d}\theta$ is our degree-of-belief that θ is in the interval $[\theta, \theta + \mathrm{d}\theta]$, given my observations $\mathbf{d}$, assuming model $\mathcal{M}$ and some prior beliefs/information $\mathcal{I}$. Whilst the full posterior density represents all of our beliefs about parameters θ, it is often useful to define quantities derivable from the posterior that make simplified statements about our inference. For example, Bayesian credible intervals allow us to declare a region $\mathcal{R}_X$ that contains X% of the posterior density, where the boundary of $\mathcal{R}_X$ is an isoprobability contour of the posterior (or an isoprobability level in 1D). It is commonplace to quote 68%, 95% and 99% credible intervals, since these (roughly) correspond to 1-, 2- and 3-σ credible regions for a one-dimensional Gaussian.

3.2.1. Errors

Let us assume we have a posterior probability distribution, which is single-peaked. Two common estimators (indicated by a hat: $\hat{\boldsymbol{\theta}}$) of the parameters are the peak (most probable) values, or the mean,

$$\hat{\boldsymbol{\theta}} = \int d\boldsymbol{\theta}\, \boldsymbol{\theta}\, p(\boldsymbol{\theta}|\vec{x}). \tag{3.12}$$

An estimator is *unbiased* if its expectation value is the true value $\boldsymbol{\theta}_0$:

$$\langle \hat{\boldsymbol{\theta}} \rangle = \boldsymbol{\theta}_0. \tag{3.13}$$

Let us assume for now that the prior is flat, so the posterior is proportional to the likelihood. This can be relaxed. Close to the peak, a Taylor expansion of the log likelihood implies that locally it is a multivariate Gaussian *in parameter space*:

$$\ln L(\vec{x};\boldsymbol{\theta}) = \ln L(\vec{x};\boldsymbol{\theta}_0) + \frac{1}{2}(\boldsymbol{\theta}_\alpha - \boldsymbol{\theta}_{0\alpha})\frac{\partial^2 \ln L}{\partial \boldsymbol{\theta}_\alpha \partial \boldsymbol{\theta}_\beta}(\boldsymbol{\theta}_\beta - \boldsymbol{\theta}_{0\beta}) + \dots \tag{3.14}$$

or

$$L(\vec{x};\boldsymbol{\theta}) = L(\vec{x};\boldsymbol{\theta}_0)\exp\left[-\frac{1}{2}(\boldsymbol{\theta}_\alpha - \boldsymbol{\theta}_{0\alpha})H_{\alpha\beta}(\boldsymbol{\theta}_\beta - \boldsymbol{\theta}_{0\beta})\right]. \tag{3.15}$$

The Hessian matrix $H_{\alpha\beta} \equiv -\frac{\partial^2 \ln L}{\partial\theta_\alpha \partial\theta_\beta}$ controls whether the estimates of θ_α and θ_β are correlated or not. If it is diagonal, the estimates are uncorrelated. Note that this is a statement about estimates of the quantities, not the quantities themselves, which may be entirely independent, but if they have a similar effect on the data, their estimates may be correlated. Note that in cases of practical interest, the likelihood may not be well described by a multivariate Gaussian at levels which set the interesting credibility levels (e.g. 68%). We turn later to how to proceed in such cases.

3.2.2. Conditional and marginal errors

If we fix all the parameters except one, then the error is given by the curvature along a line through the likelihood (or the posterior, if the prior is not flat):

$$\sigma_{\text{conditional},\alpha} = \frac{1}{\sqrt{H_{\alpha\alpha}}}. \tag{3.16}$$

This is called the *conditional error*, and is the minimum error bar attainable on θ_α if all the other parameters are known.

3.2.3. Marginalising over a Gaussian likelihood

The marginal distribution of $\boldsymbol{\theta}_1$ is obtained by integrating over the other parameters:

$$p(\boldsymbol{\theta}_1) = \int d\boldsymbol{\theta}_2 \dots d\boldsymbol{\theta}_N p(\boldsymbol{\theta}). \tag{3.17}$$

The triangle contour plots typically presented in forecast papers (some examples will be presented in Sec. 5.3) are plots of marginal distributions of all parameters in pairs, as a way to present some complex results. In that case two variables are left out of the integration.

When plotting these error ellipses it is important to specify which contours are plotted, specifcally if they are for the joint distribution (i.e. 68% of the probability lies within the inner contour), or whether 68% of the probability of a single parameter lies within the bounds projected onto a parameter axis. The latter is a 1σ, single-parameter error contour (and corresponds to $\Delta\chi^2 = 1$), whereas the former is a 1σ contour for the joint distribution, and corresponds to $\Delta\chi^2 = 2.3$.

Note that $\Delta\chi^2 = \chi^2 - \chi^2(\textit{minimum})$, where

$$\chi^2 = \sum_i \frac{(x_i - \mu_i)^2}{\sigma_i^2} \tag{3.18}$$

for data x_i with $\mu_i = \langle x_i \rangle$ and variance σ_i^2. If the data are correlated, this generalises to

$$\chi^2 = \sum_{ij}(x_i - \mu_i)C_{ij}^{-1}(x_j - \mu_j) \tag{3.19}$$

where $C_{ij} = \langle (x_i - \mu_i)(x_j - \mu_j) \rangle$.

A multivariate Gaussian likelihood is a common assumption, so it is useful to compute marginal errors for this rather general situation. The simple result is that the marginal error on parameter θ_α is

$$\sigma_\alpha = \sqrt{(H^{-1})_{\alpha\alpha}}. \tag{3.20}$$

Note that we invert the Hessian matrix, and then take the square root of the diagonal components. Let us prove this important result. In practice it is often used to estimate errors for a future experiment, where we deal with the expectation value of the Hessian, called the *Fisher Matrix*:

$$F_{\alpha\beta} \equiv \langle H_{\alpha\beta} \rangle = \left\langle -\frac{\partial^2 \ln L}{\partial \boldsymbol{\theta}_\alpha \partial \boldsymbol{\theta}_\beta} \right\rangle. \tag{3.21}$$

The expected error on θ_α is thus

$$\sigma_\alpha = \sqrt{(F^{-1})_{\alpha\alpha}}. \tag{3.22}$$

It is always at least as large as the expected conditional error. This result applies for Gaussian-shaped likelihoods, and is useful for experimental design; for real data, one would perform the marginalisation a different way, as explained later.

To prove that the marginal error for a Gaussian likelihood is given by 3.22, we will use characteristic functions.

Characteristic functions

In probability theory the Fourier Transform of a probability distribution function is known as the *characteristic function*. For a multivariate distribution with N parameters, it is defined by

$$\phi(\mathbf{k}) = \int d^N \boldsymbol{\theta}\, p(\boldsymbol{\theta}) e^{-i\mathbf{k}\cdot\boldsymbol{\theta}} \tag{3.23}$$

with reciprocal relation

$$p(\boldsymbol{\theta}) = \int \frac{d^N \mathbf{k}}{(2\pi)^N}\, \phi(\mathbf{k}) e^{i\mathbf{k}\cdot\boldsymbol{\theta}} \tag{3.24}$$

(note the choice of where to put the factors of 2π is not universal). Hence the characteristic function is also the expectation value of $e^{-i\mathbf{k}\cdot\boldsymbol{\theta}}$:

$$\phi(\mathbf{k}) = \langle e^{-i\mathbf{k}.\boldsymbol{\theta}} \rangle. \tag{3.25}$$

Part of the power of characteristic functions is the ease with which one can generate all of the moments of the distribution by differentiation:

$$\langle \boldsymbol{\theta}_\alpha^{n_\alpha} \dots \boldsymbol{\theta}_\beta^{n_\beta} \rangle = \left[\frac{\partial^{n_\alpha + \dots + n_\beta} \phi(\mathbf{k})}{\partial(-i\mathbf{k}_\alpha)^{n_\alpha} \dots \partial(-i\mathbf{k}_\beta)^{n_\beta}} \right]_{\mathbf{k}=\mathbf{0}}. \tag{3.26}$$

This can be seen if one expands $\phi(\mathbf{k})$ in a power series, using

$$\exp(\alpha) = \sum_{i=0}^{\infty} \frac{\alpha^n}{n!}, \tag{3.27}$$

giving

$$\phi(\mathbf{k}) = 1 - i\mathbf{k} \cdot \langle \boldsymbol{\theta} \rangle - \frac{1}{2} \sum_{\alpha\beta} \mathbf{k}_\alpha \mathbf{k}_\beta \langle \boldsymbol{\theta}_\alpha \boldsymbol{\theta}_\beta \rangle + \dots. \tag{3.28}$$

Hence for example we can compute the mean

$$\langle \boldsymbol{\theta}_\alpha \rangle = \left[\frac{\partial \phi(\mathbf{k})}{\partial(-i\mathbf{k}_\alpha)} \right]_{\mathbf{k}=\mathbf{0}} \tag{3.29}$$

and the covariances, from

$$\langle \boldsymbol{\theta}_\alpha \boldsymbol{\theta}_\beta \rangle = \left[\frac{\partial^2 \phi(\mathbf{k})}{\partial(-i\mathbf{k}_\alpha)\partial(-i\mathbf{k}_\beta)} \right]_{\mathbf{k}=\mathbf{0}}. \tag{3.30}$$

(Putting $\alpha = \beta$ yields the variance of $\boldsymbol{\theta}_\alpha$ after subtracting the square of the mean).

3.2.4. The expected marginal error on θ_α is $\sqrt{(\mathbf{F}^{-1})_{\alpha\alpha}}$

The likelihood is here assumed to be a multivariate Gaussian, with expected Hessian given by the Fisher matrix. Thus (suppressing ensemble averages)

$$L(\boldsymbol{\theta}) = \frac{1}{(2\pi)^{M/2}\sqrt{\det \mathbf{F}}} \exp\left(-\frac{1}{2}\boldsymbol{\theta}^T \mathbf{F} \boldsymbol{\theta}\right), \tag{3.31}$$

where T indicates transpose, and for simplicity we have assumed the parameters have zero mean (if not, just redefine $\boldsymbol{\theta}$ as the difference between $\boldsymbol{\theta}$ and the mean). We proceed by diagonalising the quadratic, then computing the characteristic function, and compute the covariances using Eq. 3.30. This is achieved in the standard way by rotating the parameter axes:

$$\boldsymbol{\psi} = R\boldsymbol{\theta} \tag{3.32}$$

for a matrix $\mathbf{R}$. Since $\mathbf{F}$ is real and symmetric, $\mathbf{R}$ is orthogonal, $\mathbf{R}^{-1} = \mathbf{R}^T$. Diagonalising gives

$$\boldsymbol{\theta}^T \mathbf{F} \boldsymbol{\theta} = \boldsymbol{\psi}^T \mathbf{R}\mathbf{F}\mathbf{R}^T \boldsymbol{\psi}, \tag{3.33}$$

and the diagonal matrix composed of the eigenvalues of $\mathbf{F}$

$$\boldsymbol{\Lambda} = \mathbf{R}\mathbf{F}\mathbf{R}^T, \tag{3.34}$$

Note that the eigenvalues of $\mathbf{F}$ are positive, as $\mathbf{F}$ must be positive-definite.

The characteristic function is

$$\phi(\mathbf{k}) = \frac{1}{(2\pi)^{M/2}\sqrt{\det \mathbf{F}}} \int d^M\boldsymbol{\psi} \exp\left(-\frac{1}{2}\boldsymbol{\psi}^T \boldsymbol{\Lambda}\mathbf{F}\boldsymbol{\psi}\right) \exp(-i\mathbf{k}^T \mathbf{R}^T \boldsymbol{\psi}) \tag{3.35}$$

where we exploit the fact that the rotation has unit Jacobian to change $d^M\boldsymbol{\theta}$ to $d^M\boldsymbol{\psi}$. If we define $\mathbf{K} \equiv \mathbf{R}\mathbf{k}$,

$$\phi(\mathbf{k}) = \frac{1}{(2\pi)^{M/2}\sqrt{\det \mathbf{F}}} \int d^M\boldsymbol{\psi} \exp\left(-\frac{1}{2}\boldsymbol{\psi}^T \boldsymbol{\Lambda}\boldsymbol{\psi}\right) \exp(-i\mathbf{K}^T \boldsymbol{\psi}) \tag{3.36}$$

and since $\boldsymbol{\Lambda}$ is diagonal, the first exponential is a sum of squares, which we can integrate separately, using

$$\int_{-\infty}^{\infty} d\psi \exp(-\Lambda\psi^2/2) \exp(-iK\psi) = \sqrt{2\pi/\Lambda} \exp[-K^2/(2\Lambda)]. \tag{3.37}$$

All multiplicative factors cancel (since the rotation preserves the eigenvalues, so $\det(\mathbf{F}) = \prod \boldsymbol{\Lambda}_\alpha$), and we obtain

$$\phi(\mathbf{k}) = \exp\left(-\sum_i K_i^2/(2\Lambda_i)\right) = \exp\left(-\frac{1}{2}\mathbf{K}^T \boldsymbol{\Lambda}^{-1} \mathbf{K}\right) = \exp\left(-\frac{1}{2}\mathbf{k}^T \mathbf{F}^{-1} \mathbf{k}\right) \tag{3.38}$$

where the last result follows from $\mathbf{K}^T \boldsymbol{\Lambda}^{-1} \mathbf{K} = \mathbf{k}^T (\mathbf{R}^T \boldsymbol{\Lambda}^{-1} \mathbf{R})\mathbf{k} = \mathbf{k}^T \mathbf{F}^{-1} \mathbf{k}$. Having obtained the characteristic function, Eq. 3.22 follows immediately from Eq. 3.30.

3.3. Fisher Matrix Analysis

We have already mentioned that Fisher matrices can be useful for experimental design. Now we can ask ourselves: how accurately can we estimate model parameters from a given data set?

Suppose for definiteness that our data set consists of N real numbers x_1, x_2, ..., x_N, which we arrange in an N-dimensional vector $\vec{x}$. These numbers could for instance denote the N coefficients of a Fourier expansion of an observed galaxy density field. Before collecting the data, we think of $\vec{x}$ as a random variable with some probability distribution $L(\vec{x};\boldsymbol{\theta})$, which depends in some known way on a vector of M model parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_M)$.

Such model parameters might for instance be the cosmological parameters. We will let $\boldsymbol{\theta}_0$ denote the true parameter values and let $\boldsymbol{\theta}$ refer to our estimate of $\boldsymbol{\theta}$. Since $\boldsymbol{\theta}$ is some function of the data vector $\vec{x}$, it too is a random variable. For it to be a good estimate, we would of course like it to be unbiased, *i.e.*,

$$\langle \boldsymbol{\theta} \rangle = \boldsymbol{\theta}_0, \tag{3.39}$$

and give as small error bars as possible, *i.e.*, minimize the standard deviations

$$\Delta\theta_\alpha \equiv \left(\left\langle \theta_\alpha^2 \right\rangle - \langle \theta_\alpha \rangle^2 \right)^{1/2}. \tag{3.40}$$

In statistics jargon, we want the BUE θ_α, which stands for the "Best Unbiased Estimator" (Tegmark et al., 1997).

A key quantity in this context is the so-called *Fisher information matrix*, defined as

$$\boxed{\mathbf{F}_{\alpha\beta} \equiv \left\langle \frac{\partial^2 \mathcal{L}}{\partial\theta_\alpha \partial\theta_\beta} \right\rangle} \tag{3.41}$$

where

$$\mathcal{L} \equiv -\ln L. \tag{3.42}$$

Another key quantity is the *maximum likelihood estimator*, or *ML-estimator* for brevity, defined as the parameter vector $\boldsymbol{\theta}_{\rm ML}$ that maximizes the likelihood function $L(\vec{x};\boldsymbol{\theta})$.

Using this notation, a number of powerful theorems have been proven (see e.g. Kenney & Keeping, 1951):

(1) For any unbiased estimator, $\Delta\theta_\alpha \geq 1/\sqrt{\mathbf{F}_{\alpha\alpha}}$ (the *Cramér-Rao* inequality).

(2) If an unbiased estimator attaining ("saturating") the Cramér-Rao bound exists, it is the ML estimator (or a function thereof).

(3) The ML-estimator is asymptotically BUE.

The first of these theorems thus places a firm lower limit on the error bars that one can attain, regardless of which method one is using to estimate the parameters from the data.

The normal case is that the other parameters are estimated from the data as well, in which case, as we have seen, the minimum standard deviation rises to

$$\Delta\theta_\alpha \geq (\mathbf{F}^{-1})_{\alpha\alpha}^{1/2}. \tag{3.43}$$

This is called the *marginal error*.

The second theorem shows that maximum-likelihood (ML) estimates have quite a special status: if there is a best method, then the ML-method is the one. Finally, the third result basically tells us that in the limit of a very large data set, the ML-estimate for all practical purposes is the best estimate, the one for which the Cramér-Rao inequality becomes an equality[6]. It is these nice properties that have made ML-estimators so popular.

Note that conditional and marginal errors coincide if $\mathbf{F}$ is diagonal. If it is not, then the *estimates* of the parameters are correlated (even if the parameters themselves are uncorrelated).

[6]This is sometimes called 'saturating the Cramér-Rao bound'

3.4. Posterior sampling

Step (3) described in Sec.3.1 involves computing the posterior distribution. For a small number of parameters, the posterior can simply be evaluated exactly on a grid of points in the model-parameter space. However, this approach quickly becomes computationally challenging as the number of parameters increases beyond $\gtrapprox 3$. In this case, we can instead draw samples from the posterior, using Monte Carlo methods. Given a (sufficiently large) set of samples of the posterior density, we can reconstruct the full density function, or compute summary statistics (mean, mode, covariance etc) for making clear scientific statements. In the following section we review the use of some Monte Carlo sampling schemes for practical sampling from posterior densities.

3.4.1. Markov Chain Monte Carlo methods

Sampling from a (high-dimensional) probability density is non-trivial. We expect properly drawn samples to come from regions of high probability, and these cannot be known a priori. Even if they were (roughly) known, there is no obvious way to directly draw samples without evaluating the density everywhere.

Markov Chain Monte Carlo (MCMC) methods provide an efficient way of sampling (high-dimensional) probability densities and are now well established as a cornerstone of Bayesian inference. MCMC methods work by generating a Markov chain of points $\{\theta_i\}$, where each step θ_i depends probabilistically on the previous θ_{i-1}, and the Markov process is carefully constructed so that the stationary distribution of $\{\theta\}$ converges to the target density in our case, the posterior distribution.

MCMC Markov chains need to satisfy two conditions to ensure convergence to the target distribution:

- Firstly the transition probability of the Markov chain $T(\theta'|\theta)$ and the target (posterior) density must satisfy detailed balance,

$$P(\theta|\mathbf{d})T(\theta'|\theta) = P(\theta'|\mathbf{d})T(\theta|\theta') \tag{3.44}$$

  to ensure that $P(\theta|\mathbf{d})$ is a stationary distribution of the Markov chain, i.e., for some starting point(s) θ_0, the equilibrium distribution of points in the chain θ converges to $P(\theta|\mathbf{d})$.

- Secondly, the Markov process must be ergodic – that is, any point can be reached from any other point in the chain (although not necessarily in a single step). This ensures the stationary distribution of the chain is unique; if any point can be reached from any other, the stationary distribution can no longer depend on the starting point θ_0, so combined with detailed balance the stationary distribution is necessarily the target density, as required.

If we can construct a Markov chain that satisfies detailed balance with the posterior and is ergodic, we have an efficient and generally applicable way of generating a set of samples from a posterior density. We provide a brief discussion of practical MCMC implementations below, focusing on those that have been employed in this thesis.

Convergence of MCMC chains and burn-in

MCMC methods guarantee convergence of the samples to the target density in the limit $t \to \infty$. Given that we cannot run infinitely long chains, we require a convergence test to determine when we can consider our samples to be representative of the target and we can stop the CPUs. A wide range of diagnostic tests are available but perhaps the most commonly used convergence test is the Gelman-Rubin test. They define a statistic R computed from M chains of length N (with dispersed independent

starting points),

$$R = \sqrt{(1 - 1/N) + \frac{M+1}{MN}\frac{B_i}{W}} \tag{3.45}$$

where B is the sample variance of the individual chain means, W is the sample mean of the individual chain sample variances. R is computed on a parameter-by-parameter basis. Heuristically speaking, in the limit of converged chains $R \to 1$ and it is commonplace to take $R < 1.03$ (or even smaller) as a convergence criterion. It should be noted that convergence tests are inevitability heuristic, and there does not exist a formal convergence test that absolutely guarantees MCMC convergence. However, in practice for well-behaved posterior densities without pathologies – such as two widely separated narrow peaks – tests such as the Gelman-Rubin test are practical and effective.

The final MCMC implementation issue is burn-in. In the case where the starting point of the MCMC chain is in a region of low target density, the chain will have to random-walk over (aka "burn in") to where the bulk of the probability density is. These initial samples are an artefact of a way-out starting point and will not be representative of the target density whilst in the limit of long chains, this burn-in period will pale into insignificance, it is commonplace to identify and remove burn-in samples, for example the first 30% of the chain.

3.4.2. Metropolis-Hastings

The Metropolis-Hastings (MH) algorithm is the simplest and most widespread MCMC sampling scheme, and the basis for a multitude of more sophisticated methods. MH follows a simple iterative procedure (illustrated in Algorithm 1 below) for sampling a target density $P(\theta_i|\mathbf{d})$:(1) Given the current position θ_i, propose a new position θ_{i+1} from a proposal distribution $Q(\theta_{i+1}|\theta_i)$. (2) Accept the proposed step with probability $p = min[1, P(\theta_{i+1}|\mathbf{d})Q(\theta_i|\theta_{i+1})/P(\theta_i|\mathbf{d})Q(\theta_{i+1}|\theta_i)]$, else remain at the current position, i.e., $\theta_{+1} = \theta_i$. The form of the acceptance probability ensures that detailed balance is satisfied, and for a sensible choice of proposal distribution the resulting Markov chain will be ergodic. The MH algorithm is guaranteed to converge to a stationary set of samples from the posterior in the limit $t \to \infty$. However, how fast it converges depends critically on the choice of proposal density Q and this inevitably requires some tuning; and for inference problems with > 10 parameters the acceptance rates tend to become too low for MH to be practical. There are a cornucopia of extensions and alternatives to the vanilla MH sampling scheme that achieve faster convergence, are practical for high-dimensional sampling problems, and often require less tuning. A notable example is Hamiltonian Monte Carlo (Neal, 2012) that is now widely used in astronomy and has proven effective for ultra-high dimensional $> 10^6$ inference problems, with fast and flexible public codes available.

Algorithm 1 The procedure for a single Metropolis-Hastings MCMC step.

1: Draw a proposal $Y \sim Q(Y; X(t))$
2: $q \leftarrow \frac{[p(Y)\,Q(X(t);Y)]}{[p(X(t))\,Q(Y;X(t))]}$ *// This line is generally expensive*
3: $r \leftarrow R \sim [0, 1]$
4: **if** $r \leq q$ **then**
5: $\quad X(t+1) \leftarrow Y$
6: **else**
7: $\quad X(t+1) \leftarrow X(t)$
8: **end if**

3.4.3. Model selection and Nested Sampling

We already mentioned that in parameter estimation the normalising evidence factor in Bayes' theorem (Eq.3.1) is usually ignored, since it is independent of the parameters θ, and inferences are obtained by taking samples from the (unnormalised) posterior using standard MCMC sampling methods, where at equilibrium the chain contains a set of samples from the parameter space distributed according to the posterior. This posterior constitutes the complete Bayesian inference of the parameter values, and can be marginalised over each parameter to obtain individual parameter constraints.

In contrast to parameter estimation problems, in model selection the evidence takes the central role and is simply the factor required to normalize the posterior over θ:

$$\mathcal{Z} = \int \mathcal{L}(\theta)\pi(\theta)d^D\theta, \tag{3.46}$$

where D is the dimensionality of the parameter space. As the average of the likelihood over the prior, the evidence automatically implements Occam's razor: a simpler theory with compact parameter space will have a larger evidence than a more complicated one, unless the latter is significantly better at explaining the data. The question of model selection between two models H_0 and H_1 can then be decided by comparing their respective posterior probabilities given the observed data set $\mathbf{D}$, as follows

$$\frac{\Pr(H_1|\mathbf{D})}{\Pr(H_0|\mathbf{D})} = \frac{\Pr(\mathbf{D}|H_1)\Pr(H_1)}{\Pr(\mathbf{D}|H_0)\Pr(H_0)} = \frac{\mathcal{Z}_1}{\mathcal{Z}_0}\frac{\Pr(H_1)}{\Pr(H_0)}, \tag{3.47}$$

where $\Pr(H_1)/\Pr(H_0)$ is the a priori probability ratio for the two models, which can often be set to unity but occasionally requires further consideration.

Evaluation of the multidimensional integral (3.46) is a challenging numerical task. The standard technique of thermodynamic integration draws MCMC samples not from the posterior directly but from $\mathcal{L}^\lambda\pi$ where λ is an inverse temperature that is slowly raised from ≈ 0 to 1 according to some annealing schedule. It is possible to obtain accuracies of within 0.5 units in log-evidence via this method, but in cosmological model selection applications it typically requires of order 10^6 samples per chain (with around 10 chains required to determine a sampling error). This makes evidence evaluation at least an order of magnitude more costly than parameter estimation.

Nested Sampling

Nested sampling (Skilling, 2006) is a Monte Carlo method used for the computation of the evidence that can also provide posterior inferences. It transforms the multi-dimensional integral of Eq. 3.46 into a one-dimensional integral over the prior volume. This is done by defining the prior volume X as $dX = \pi(\theta)d^N\theta$. Therefore,

$$X(\lambda) = \int_{\mathcal{L}(\theta)>\lambda} \pi(\theta)d^N\theta. \tag{3.48}$$

This integral extends over the region of parameter space contained within the likelihood contour $\mathcal{L}(\theta) = \lambda$. The evidence integral, Eq. 3.46, can then be written as

$$\mathcal{Z} = \int_0^1 \mathcal{L}(X)dX, \tag{3.49}$$

where $\mathcal{L}(X)$ is the inverse of Eq. 3.48 and is a monotonically decreasing function of X. Thus, if we evaluate the likelihoods $\mathcal{L}_i = \mathcal{L}(X_i)$, where X_i is a sequence of decreasing values,

$$0 < X_M < \cdots < X_2 < X_1 < X_0 = 1. \tag{3.50}$$

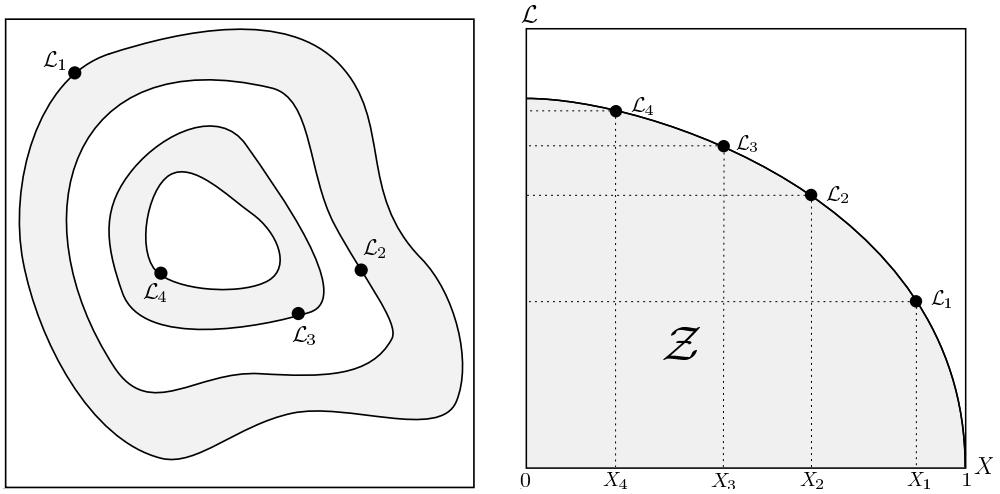

Figure 3.5.: Cartoon illustrating (a) the posterior of a two dimensional problem; and (b) the transformed $\mathcal{L}(X)$ function where the prior volumes X_i are associated with each likelihood $\mathcal{L}_i$. Originally published in Feroz et al. (2009).

The evidence can then be approximated numerically as a weighted sum

$$\mathcal{Z} = \sum_{i=1}^{M} \mathcal{L}_i w_i, \tag{3.51}$$

where the weights w_i for the simple trapezium rule are given by $w_i = \frac{1}{2}(X_{i-1} - X_{i+1})$. An example of a posterior in two dimensions and its associated function $\mathcal{L}(X)$ is shown in Fig. 3.5.

The fundamental operation of nested sampling begins with the initial, 'live', points being chosen at random from the entire prior volume. The lowest likelihood live point is removed and replaced by a new sample with higher likelihood. This removal and replacement of live points continues until a stopping condition is reached (MultiNest uses a tolerance on the evidence calculation). The difficult task lies in finding a new sample with higher likelihood than the discarded point. As the algorithm goes up in likelihood, the prior volume that will satisfy this condition decreases until it contains only a very small portion of the total parameter space, making this sampling potentially very inefficient. MultiNest tackles this problem by enclosing all of the active points in clusters of ellipsoids. New points can then be chosen from within these ellipsoids using a fast analytic function. Since the ellipsoids will decrease in size along with the distribution of live points, their surfaces in effect represent likelihood contours of increasing value; the algorithm climbs up these contours seeking new points. As the clusters of ellipsoids are not constrained to fit any particular distribution, they can easily enclose curving degeneracies and are able to separate out to allow for multimodal distributions. This separation also allows for the calculation of the 'local' evidence associated with each mode. MultiNest has been shown to be of substantial use in astrophysics and particle physics (see Feroz et al., 2009), typically showing great improvement in efficiency over traditional MCMC techniques.

3.4.4. Affine invariant sampling

Most uses of MCMC in the astrophysics literature are based on slight modifications to the Metropolis-Hastings (M–H) method introduced above. Each step in a M–H chain is proposed using a compact proposal distribution centered on the current position of the chain (normally a multivariate Gaussian or something similar). Since each term in the covariance matrix of this proposal distribution is an unspecified parameter, this method has $N\,[N+1]/2$ tuning parameters (where N is the dimension of

the parameter space). To make matters worse, the performance of this sampler is very sensitive to these tuning parameters and there is no fool-proof method for choosing the values correctly. As a result, many heuristic methods have been developed to attempt to determine the optimal parameters in a data-driven way. Unfortunately, these methods all require a lengthy "burn-in" phase where shorter Markov chains are sampled and the results are used to tune the hyperparameters. This extra cost is unacceptable when the likelihood calls are computationally expensive.

The problem with traditional sampling methods can be visualized by looking at the simple but highly anisotropic density

$$p(\mathbf{x}) \propto f\left(-\frac{(x_1 - x_2)^2}{2\,\epsilon} - \frac{(x_1 + x_2)^2}{2}\right) \tag{3.52}$$

which would be considered difficult (in the small-ϵ regime) for standard MCMC algorithms. In principle, it is possible to tune the hyperparameters of a M–H sampler to make this sampling converge quickly, but if the dimension is large and calculating the density is computationally expensive the tuning procedure becomes intractable. Also, since the number of parameters scales as $\sim N^2$, this problem gets much worse in higher dimensions. Eq. 3.52 can, however, be transformed into the much easier problem of sampling an isotropic density by an *affine transformation* of the form

$$y_1 = \frac{x_1 - x_2}{\sqrt{\epsilon}}, \qquad y_2 = x_1 + x_2 \quad . \tag{3.53}$$

This motivates affine invariance: an algorithm that is *affine invariant* performs equally well under all linear transformations; it will therefore be insensitive to covariances among parameters.

The stretch move (Goodman & Weare, 2010) proposed an affine-invariant ensemble sampling algorithm informally called the "stretch move." This algorithm significantly outperforms standard M–H methods producing independent samples with a much shorter autocorrelation time. For completeness and for clarity of notation, we summarize the algorithm here and refer the interested reader to the original paper for more details. This method involves simultaneously evolving an ensemble of K *walkers* $S = \{X_k\}$ where the proposal distribution for one walker k is based on the current positions of the $K-1$ walkers in the *complementary ensemble* $S_{[k]} = \{X_j,\ \forall\, j \neq k\}$. Here, "position" refers to a vector in the N-dimensional, real-valued parameter space.

To update the position of a walker at position X_k, a walker X_j is drawn randomly from the remaining walkers $S_{[k]}$ and a new position is proposed:

$$X_k(t) \rightarrow Y = X_j + Z\,[X_k(t) - X_j] \tag{3.54}$$

where Z is a random variable drawn from a distribution $g(Z = z)$. It is clear that if g satisfies

$$g(z^{-1}) = z\,g(z), \tag{3.55}$$

the proposal of Eq. 3.54 is symmetric. In this case, the chain will satisfy detailed balance if the proposal is accepted with probability

$$q = \min\left(1,\, Z^{N-1}\,\frac{p(Y)}{p(X_k(t))}\right) \quad , \tag{3.56}$$

where N is the dimension of the parameter space. This procedure is then repeated for each walker in the ensemble *in series* following the procedure shown in Algorithm 2.

Goodman & Weare (2010) advocate a particular form of $g(z)$, namely

$$g(z) \propto \begin{cases} \dfrac{1}{\sqrt{z}} & \text{if } z \in \left[\dfrac{1}{a}, a\right], \\ 0 & \text{otherwise} \end{cases} \tag{3.57}$$

where a is an adjustable scale parameter that Goodman & Weare (2010) set to 2.

Algorithm 2 A single stretch move update step from Goodman & Weare (2010)

1: **for** $k = 1, \ldots, K$ **do**
2: Draw a walker X_j at random from the complementary ensemble $S_{[k]}(t)$
3: $z \leftarrow Z \sim g(z)$, Eq. 3.57
4: $Y \leftarrow X_j + z\,[X_k(t) - X_j]$
5: $q \leftarrow z^{N-1}\, p(Y)/p(X_k(t))$ *// This line is generally expensive*
6: $r \leftarrow R \sim [0, 1]$
7: **if** $r \leq q$, Eq. 3.56 **then**
8: $X_k(t+1) \leftarrow Y$
9: **else**
10: $X_k(t+1) \leftarrow X_k(t)$
11: **end if**
12: **end for**

The parallel stretch move It is tempting to parallelize the stretch move algorithm by simultaneously advancing each walker based on the state of the ensemble instead of evolving the walkers in series. Unfortunately, this subtly violates detailed balance. Instead, we must split the full ensemble into two subsets ($S^{(0)} = \{X_k,\ \forall k = 1, \ldots, K/2\}$ and $S^{(1)} = \{X_k,\ \forall k = K/2+1, \ldots, K\}$) and simultaneously update all the walkers in $S^{(0)}$ — using the stretch move procedure from Algorithm 2 — based *only* on the positions of the walkers in the other set ($S^{(1)}$). Then, using the new positions $S^{(0)}$, we can update $S^{(1)}$. In this case, the outcome is a valid step for all of the walkers. The pseudocode for this procedure is shown in Algorithm 3. This code is similar to Algorithm 2 but now the computationally expensive inner loop (starting at line 2 in Algorithm 3) can be run in parallel.

The performance of this method — quantified by the autocorrelation time — is comparable to the serial stretch move algorithm but the fact that one can now take advantage of generic parallelization makes it extremely powerful.

Algorithm 3 The parallel stretch move update step

1: **for** $i \in \{0, 1\}$ **do**
2: **for** $k = 1, \ldots, K/2$ **do**
3: *// This loop can now be done in parallel for all k*
4: Draw a walker X_j at random from the complementary ensemble $S^{(\sim i)}(t)$
5: $X_k \leftarrow S^{(i)}_k$
6: $z \leftarrow Z \sim g(z)$, Eq. 3.57
7: $Y \leftarrow X_j + z\,[X_k(t) - X_j]$
8: $q \leftarrow z^{n-1}\, p(Y)/p(X_k(t))$
9: $r \leftarrow R \sim [0, 1]$
10: **if** $r \leq q$, Eq. 3.56 **then**
11: $X_k(t + \frac{1}{2}) \leftarrow Y$
12: **else**
13: $X_k(t + \frac{1}{2}) \leftarrow X_k(t)$
14: **end if**
15: **end for**
16: $t \leftarrow t + \frac{1}{2}$
17: **end for**

All the sampling methods presented in Secs. 3.4.3 and 3.4.4 have been used in the analyses shown in Chapter 6. Specifically, Nested Sampling has been employed in Sec. 6.2 through its MultiNest implementation (Feroz et al., 2009), while affine invariant sampling has been used in Secs. 6.2 and 6.3, through both the CosmoHammer suite (Akeret et al., 2012) and an independent implementation.

Chapter 4

4 The theory of weak gravitational lensing

In this Chapter we review the formalism of weak gravitational lensing, starting in Sec. 4.1 with an introduction to the phenomen and general theory of lensing; we then specialise our treatment in Sec. 4.2 to cosmic shear, the gravitational lensing caused by the large-scale structure, which we describe with two different formalisms, tomography and a spherical Fourier-Bessel expansion of the shear field.

4.1. An introduction to the weak gravitational lensing effect

In this section we introduce the phenomenon of gravitational lensing and show how the general-relativistic treatment of light deflection leads to the lens equation. We introduce the fields characterising gravitational lenses and their effect on image shapes: the lensing potential; the deflection angle; the convergence, shear and reduced shear; and the magnification. In our review of the basics of gravitational lensing we follow closely the presentation given in Bartelmann & Schneider (2001) and we refer to them for details and rigorous derivations of the equations presented here.

4.1.1. Light propagation in perturbed spacetime

The propagation of a light bundle in arbitrary spacetimes is governed by the equation of geodesic deviation, or Jacobi equation, which links the deformation of the cross-section of the bundle to the optical tidal matrix, quantifying spacetime curvature. In the following we consider a homogeneous and isotropic spacetime described by a Friedmann-Lemaitre-Robertson-Walker metric (already introduced in Sec. 2.1) with 'small' perturbations such that the gravitational potential of the inhomogeneities is small ($\phi \ll c^2$) peculiar velocities are small, ($\upsilon \ll c$) and inhomogeneities are localised, i.e. ϕ varies on a scale that is small compared to the Hubble scale. Then the local neighbourhood of inhomogeneities can be approximated as flat space and the perturbed metric has the form

$$\mathrm{d}s^2 = a(\eta)^2\Big\{-\Big(1+\frac{2\Phi}{c^2}\Big)c^2\mathrm{d}\eta^2 + \Big(1-\frac{2\Phi}{c^2}\Big)[\mathrm{d}\chi^2 + \chi^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\Phi^2)]\Big\} \tag{4.1}$$

where η is the conformal time, related to coordinate time t by $\mathrm{d}\eta = \mathrm{d}t/a$.

The optical tidal matrix can then be split up into two contributions, from the homogeneous and isotropic background on the one hand and inhomogeneities on the other. The Jacobi equation can be brought into the following form:

$$\Big(\frac{\mathrm{d}^2}{\mathrm{d}\chi^2} + K\Big)x^i = -\frac{2}{c^2}\partial^i\Phi. \tag{4.2}$$

$(\chi_1, \chi_2)^T$ is the comoving separation between a fiducial ray in the light bundle and a closely neighbouring ray, measured perpendicular to the tangent vector to the fiducial ray; χ is the comoving distance

from the observer measured along the fiducial ray; K is the spatial curvature of the background. The derivative $\partial^i\Phi$ is the transverse gradient of the gravitational potential, perpendicular to the fiducial ray. If the two rays enclose an angle $(\theta_1, \theta_2)^T$ at the observer, the boundary conditions are

$$x^i|_{\chi=0} = 0, \quad \left.\frac{\mathrm{d}x^i}{\mathrm{d}\chi}\right|_{\chi=0} = \theta^i. \tag{4.3}$$

Eq. 4.2 is then solved by

$$x^i(\chi) = \chi\theta^i - \frac{2}{c^2}\int_0^{\chi_s} \mathrm{d}\chi'(\chi - \chi')\partial^i\Phi[x^j(\chi'),\chi']. \tag{4.4}$$

In the Born approximation the integration is carried out along the unperturbed path with $x^j(\chi') \approx \chi'\theta^j$. Denoting by χ_s the comoving distance to the light source and transforming from the comoving separation x^i to the angle $\beta_i = x^i(\chi_s)/\chi_s$, one finds

$$\beta^i = \theta^i - \frac{2}{c^2}\int_0^{\chi_s} \mathrm{d}\chi' \frac{\chi_s - \chi'}{\chi_s}\partial^i\Phi[\chi'\theta^j,\chi']. \tag{4.5}$$

4.1.2. The lens mapping

Eq. 4.5 is the lens equation

$$\beta^i = \theta^i - \alpha^i(\theta^j), \tag{4.6}$$

where

$$\alpha^i(\theta^j) = \frac{2}{c^2}\int_0^{\chi_s} \mathrm{d}\chi' \frac{\chi_s - \chi'}{\chi_s}\partial^i\Phi[\chi'\theta^j,\chi'] \tag{4.7}$$

is the reduced deflection angle, which can be written as the transverse gradient of the lensing potential

$$\phi(\theta^j) = \frac{2}{c^2}\int_0^{\chi_s} \mathrm{d}\chi' \frac{\chi_s - \chi'}{\chi_s\chi'}\Phi[\chi'\theta^j,\chi']. \tag{4.8}$$

Note that 'gradient' here refers to the derivative with respect to angular coordinates:

$$\alpha^i(\theta^j) = \frac{\partial}{\partial\theta_i}\phi(\theta^j), \tag{4.9}$$

hence the additional factor of χ'^{-1} in the integrand. The deflection angle α^i cannot be observed because the true position of a source is not normally known. If, however, α^i varies across the extent of a source, it will appear deformed. The linear variation is described by the Jacobian of the lens mapping, Eq.4.6:

$$A = \left(\frac{\partial\beta^i}{\partial\theta^j}\right) = \begin{pmatrix} 1-\kappa & 0 \\ 0 & 1-\kappa \end{pmatrix} - \begin{pmatrix} \gamma_+ & \gamma_\times \\ \gamma_\times & \gamma_+ \end{pmatrix} \tag{4.10}$$

Here A has been split up into a diagonal part proportional to the trace and a trace-free part. In this linear approximation a circle is mapped onto an ellipse. The convergence κ quantifies the isotropic dilation of the image, while the components of the shear tensor γ describe its elongation. The inverse of the determinant of the Jacobian is the magnification

$$\mu = (\det A)^{-1} = \left[(1-\kappa)^2 - \gamma_+^2 - \gamma_\times^2\right]^{-1}. \tag{4.11}$$

Convergence and shear can be expressed in terms of the lensing potential:

$$\kappa = \frac{1}{2}\left(\frac{\partial^2\phi}{\partial(\theta^1)^2} + \frac{\partial^2\phi}{\partial(\theta^2)^2}\right), \tag{4.12}$$

$$\gamma_+ = \frac{1}{2}\left(\frac{\partial^2\phi}{\partial(\theta^1)^2} - \frac{\partial^2\phi}{\partial(\theta^2)^2}\right), \tag{4.13}$$

$$\gamma_\times = \frac{\partial^2\phi}{\partial\theta^1\partial\theta^2}. \tag{4.14}$$

The derivation above has made use of the flat-sky approximation, measuring separations in Cartesian coordinates θ^1, θ^2. On scales on which the curvature of the celestial sphere becomes important, a suitable basis must be defined and derivatives are expressed in terms of the ð operator and its complex conjugate, as we shall see in detail in next Section. The convergence κ is closely related to the mass density ρ: taking the appropriate derivative of the lensing potential given in Eq. 4.8,

$$\kappa(\theta^j) = \frac{1}{c^2}\int_0^{\chi_s} \mathrm{d}\chi' \frac{(\chi_s - \chi')\chi'}{\chi_s\chi'}\frac{\partial^2\Phi}{\partial\theta_i\partial\theta^i}[\chi'\theta^j, \chi']. \tag{4.15}$$

Poisson's equation states that

$$\partial^i\partial_i\Phi = \nabla_\perp\Phi + \nabla_\parallel\Phi = 4\pi G\rho, \tag{4.16}$$

where the Laplacian has been split up into a transverse part and the derivative along the light path. The latter averages out in the line-of-sight integration, so that the derivative perpendicular to the light path

$$\nabla_\perp\Phi = \chi'^{-2}\frac{\partial^2\Phi}{\partial\theta_i\partial\theta^i} \tag{4.17}$$

can safely be replaced by the three-dimensional Laplacian. Then

$$\kappa(\theta^j) = \frac{4\pi G}{c^2}\int_0^{\chi_s} \mathrm{d}\chi' \frac{(\chi_s - \chi')\chi'}{\chi_s}\rho[\chi'\theta^j, \chi']. \tag{4.18}$$

i.e. the convergence is a weighted projection of the density, integrated along the unperturbed light path. The shear tensor parametrises the elliptical distortion of an image. Writing the complex shear as $\gamma = |\gamma|e^{2i\phi}$, the angle ϕ is the position angle of the elliptical image of a circle. The shear tensor is invariant under rotations of π and therefore has spin 2, which is intuitive, since an ellipse is mapped onto itself under such rotations.

Information about shape can be extracted from a brightness distribution $I(\theta)$ in terms of its moments. Defining the centre of brightness as

$$\bar{\theta} = \frac{\int \mathrm{d}^2\theta\,\theta\, I(\theta)}{\int \mathrm{d}^2\theta\, I(\theta)}. \tag{4.19}$$

The second moments are given by

$$Q_{ij} = \frac{\int \mathrm{d}^2\theta(\theta_i - \bar{\theta}_i)(\theta_j - \bar{\theta}_j)I(\theta)}{\int \mathrm{d}^2\theta\, I(\theta)}, \quad i, j \in [1, 2] \tag{4.20}$$

An ellipticity parameter is derived from these moments as

$$\epsilon = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22} + 2\sqrt{Q_{11}Q_{22} - Q_{12}^2}}. \tag{4.21}$$

Defining the reduced shear as

$$g = \frac{\gamma}{1-\kappa}, \tag{4.22}$$

the transformation between the (complex) source ellipticity ϵ' and the image ellipticity ϵ by lensing is given by

$$\epsilon = \frac{\epsilon' + g}{1 + g^*\epsilon'} \quad \text{for } |g| \leq 1 \quad \epsilon = \frac{1 + \epsilon'^* g}{\epsilon'^* - g^*} \quad \text{for } |g| \geq 1. \tag{4.23}$$

4.2. 3D and tomographic cosmic shear

We begin by reviewing here the two main methods used to analyse a cosmic shear survey, so called 'tomography' and '3D cosmic shear' as they are commonly referred to, whose cosmological constraining power we will later compare in Sec. 5.3. The main idea behind both techniques is to retrieve redshift information in the analysis: this is a crucial feature to investigate the evolution of cosmic structure, particularly important for dark energy/modified gravity studies. After a brief introduction to these two methods, we derive explicitly the equations for the power spectrum of the lensing potential in both formalisms, since from that we can derive the spectra of all lensing observables. We show in particular how the tomographic case can be derived from the 3D approach by means of a series of approximations. We assume General Relativity throughout this section. In Sec. 5.1 we will extend our discussion to alternative theories of gravity, as well as compare the performance of the two methods in constraining Horndeski parameters in Sec. 5.3.

4.2.1. Including redshift information in a cosmic shear analysis

Cosmic shear is the weak gravitational lensing effect caused by the large-scale structure of the Universe: the differential deflections in light bundles from distant galaxies caused by fluctuations of the gravitational fields of the large-scale structure result in a coherent distortion of galaxy images as we observe them on the sky (see Bartelmann & Schneider, 2001; Hoekstra & Jain, 2008; Kilbinger et al., 2013, for reviews on the topic).

In contrast to studies of weak lensing by individual haloes of galaxies or galaxy clusters, which typically aim at constraining the mass or the density profile of the object, in cosmic shear analyses the focus lies on the statistical properties of the distribution of matter in the Universe. The lensing effect is not associated with a particular intervening lens, but rather corresponds to small distortions (of the order of 1%) by all potential fluctuations along the line of sight; detecting the extremely faint cosmic shear signal requires averaging over many background galaxies. In this sense, cosmic shear measurements are of a statistical nature and the statistical properties of the shear field reflect those of the underlying density field by virtue of the gravitational field equations.

The cosmic shear field has zero mean; at the level of one-point statistics, cosmological information can be extracted from e.g. peak counts (Lin, 2016; Peel et al., 2017; Fluri et al., 2018), while for two-point statistics one looks in configuration space at the angular correlation function of the shear field, or its equivalent in Fourier space, the cosmic shear angular power spectrum. At higher order, cosmic shear can break degeneracies between the dark sector and neutrinos (Peel et al., 2018). As far as two-point statistics are concerned, there are two main advantages in using the power spectrum over the correlation function – sparsity of the covariance matrix, as a consequence of isotropy (as discussed in Sec.5.1.2), and the property of isolating more easily contributions from different k-scales. We will come back to this last point in the next Chapter, where we will briefly discuss the advantages of using 3D cosmic shear power spectra within a Bayesian Hierarchical Model for cosmological inference. Here we concentrate on the cosmic shear angular power spectrum and derive formal expressions to describe it in the tomographic and 3D approach, which we will motivate and introduce in the following.

Since the first detections in early 2000s (e.g. Bacon et al., 2000; Van Waerbeke et al., 2000; Brown et al., 2003), cosmic shear analyses have developed within a well-established theoretical and experimental framework. Cosmic shear is appealing in particular as one of the most promising probes of dark energy (Jain & Taylor, 2003; Bernstein & Jain, 2004; Hannestad et al., 2006; Amendola et al., 2008; Huterer, 2010). To understand why, let us start from considering the lensing potential $\phi(\chi, \hat{\mathbf{n}})$, as it is the quantity from which all lensing observables are derived (cf. Eqs. 4.12, 4.13, 4.14). The lensing potential is defined as a weighted projection of the gravitational potential (cf. Eq. 4.8):

$$\phi(\chi, \hat{\mathbf{n}}) = \frac{2}{c^2} \int_0^{\chi} \mathrm{d}\chi' \frac{\chi - \chi'}{\chi\chi'} \Phi(\chi', \hat{\mathbf{n}}). \tag{4.24}$$

where General Relativity and spatial flatness have been assumed, and the integration is carried out in Born's approximation, i.e. along the unperturbed light path. The weight function $(\chi - \chi')/(\chi\chi')$ quantifies the influence of geometry and represents the 'efficiency' of the lens: like that of an optical lens, the efficiency of a gravitational lens is determined by the distances between source, lens and observer, implying that observed variations with redshift will be sensitive to the cosmological model. In addition, as the shear is linked to the tidal field of the gravitational potential, it probes the matter distribution in the Universe.

Thus cosmic shear is sensitive to the growth rate of the perturbations of the gravitational potential (through the matter power spectrum) and to the geometry of the Universe (through the distance-redshift relation). These features are crucial for dark energy studies, as they allow us to study the dark energy component as a dynamical effect in redshift (or equivalently, time). It follows that the sensitivity of cosmic shear to dark energy can be fully exploited only if the analysis performed is able to recover information on the evolution in redshift of the large-scale structure. This is only to a little extent achieved in a 2-dimensional analysis: galaxy shapes are observed on the 2-dimensional celestial sphere and the shear components are line-of-sight projected quantities, with the projection causing loss of information on the redshift evolution (Jain & Seljak, 1997; Takada & Jain, 2003a,b; Munshi & Kilbinger, 2006; Jee et al., 2013; Kilbinger et al., 2013).

We can understand this by looking at Eq. 4.24: since the lensing observables, i.e. convergence and cosmic shear, are derived from the lensing potential with linear relations (cf. Eqs. 4.12, 4.13, 4.14), already at the level of the lensing potential we see how cosmic shear provides an integrated measurement of the evolution of the cosmic density field weighted by the lensing efficiency function. In turn, the weak lensing convergence and shear fields, being line of sight-averaged quantities, are statistically not as constraining as the full 3-dimensional density field, implying loss of information. This is due to the mixing of spatial scales that takes place in the projection and to the fact that integrating the signal along the line of sight averages out the sensitivity to those parameters that, entering the model in a nonlinear way, may produce different effects on the lensing signal at different redshifts (Schäfer & Heisenberg, 2012).

For this reason, as an improvement with respect to a pure 2D projection, a tomographic analysis based on a binning in redshift of the sources has been first proposed in Hu (1999) and has since become the standard technique for cosmological weak lensing studies (Takada & White, 2004; Simon et al., 2004; Takada & Jain, 2004; Hollenstein et al., 2009; Kilbinger et al., 2009; Schäfer & Heisenberg, 2012; Heymans et al., 2013). Galaxies are assigned to different bins according to their redshifts, so that intra- and inter-bin correlations of the binned shear field can be computed. This reduces the range of the projection to the width of the bins and allows for some gain in redshift information through the inter-bin correlations. Despite its success in providing some sensitivity to the growth of structure with its '2D$\frac{1}{2}$' nature, as it has sometimes been relabelled, tomography has still the disadvantage of representing a compression of data: while the 2D analysis performed within a single bin is such that the range of the projection is smaller than in the pure 2D case, being restricted to the width of the bin, it is also true that this does not represent yet a fully 3D treatment of the shear field, still entailing some loss of information that needs to be accurately evaluated and possibly avoided.

As an alternative to tomography, a method to retain information on the redshift of each source galaxy along the entire weak lensing analysis, based on a spherical Fourier-Bessel decomposition of the shear field, has been first proposed in Heavens (2003) and subsequently refined in Castro et al. (2005), Heavens et al. (2006) and Kitching et al. (2011). In addition to avoiding any binning and averaging in redshift, the spherical Fourier-Bessel formalism allows for a cleaner separation between angular (ℓ) and radial (k) modes (Kitching et al., 2014), in a sense that will be specified in Sec.5.1, where we describe this 3D formalism. This feature makes it easier than in tomography to reduce the impact of problematic small scales, where models for the non-linear growth of structure (Smith et al., 2003; Takahashi et al., 2012; Mead et al., 2015) or baryon feedback (Semboloni et al., 2011; van Daalen et al., 2011; Semboloni et al., 2013) do not yet provide a fully reliable description. These advantages compensate for the extra computational time required by the more complicated integrations in the covariance of the shear modes, which represent a challenge for standard numerical approaches; in Chapter 7 we will discuss in detail the numerical techniques that have been implemented to tackle these issues in order to produce the forecasts presented in Sec. 5.3.

4.2.2. Shear as a spin-2 field, convergence as a spin-0 field

A particularly elegant way of representing the shear and convergence fields is through spin raising $\eth$ and lowering $\bar{\eth}$ operators acting on the lensing potential (see Castro et al., 2005; Goldberg et al., 1967; Newman & Penrose, 1962, for detailed discussions). The $\eth$-derivative (and its conjugate $\bar{\eth}$) acts as a covariant differentiation operator on the celestial sphere and relates quantities of different spin, raising (lowering) the spin s of a function, a number which characterises its transformation properties under rotations. The shear components γ_1 and γ_2 for example, which we defined in Sec. 4.1.2, can be conveniently packaged into a single complex scalar field $\gamma = \gamma_1 + i\gamma_2$. Under $SO(2)$ rotations (on the sky), the complex shear field transforms as $\gamma \to e^{-is\alpha}\gamma$ with $s = 2$ and a rotation angle α - this can be straightforwardly verified by rotating the shear part of the distortion matrix A introduced in Eq. 4.10. This phase dependence expresses the fact that the complex shear field $\gamma(\vec{\chi})$ is invariant under a rotation over π radians. The complex shear is therefore a spin-2 field[7].

In general, any spin-s field can be generated from a spin-0 scalar through successive application of spin raising $\eth$ and lowering $\bar{\eth}$ operators. Acting twice on ϕ, for example, the $\eth$ operator relates the scalar (spin-0) lensing potential to the spin-2 shear field γ, so that we can define: [8]

$$\gamma(\chi, \hat{\mathbf{n}}) := {}_2\gamma(\chi, \hat{\mathbf{n}}) = \frac{1}{2}\eth\eth\phi(\chi, \hat{\mathbf{n}}), \quad \bar{\gamma}(\chi, \hat{\mathbf{n}}) := {}_{-2}\gamma(\chi, \hat{\mathbf{n}}) = \frac{1}{2}\bar{\eth}\bar{\eth}\phi(\chi, \hat{\mathbf{n}}). \tag{4.25}$$

The $\eth$ formalism becomes particularly important when considering big portions of the sky, where the curvature of the celestial sphere is not negligeable: in the flat-sky approximation, we recover the expressions already introduced in Sec. 4.1.2 (cf. Eq. 4.13):

$$\eth = \frac{\partial}{\partial\theta^1} + i\frac{\partial}{\partial\theta^2}. \tag{4.26}$$

Similarly, the convergence field κ is invariant under $SO(2)$ rotations and is hence a spin-0 field; it is also sourced from the lensing potential by application of the $\eth$ operators:

$$\kappa = \frac{1}{4}[\eth\bar{\eth} + \bar{\eth}\eth]\phi(\mathbf{r}). \tag{4.27}$$

[7] This fact turns out to be intuitive if we think that the shear transforms the image of an ideal perfectly circular source into an ellipse; we recognise that an ellipse gets mapped into itself by rotations of 180 degrees.

[8] This can be verified by comparing the combinations of covariant derivatives contributing to the shear components γ_1 and γ_2 in their definition to the explicit form taken by the $\eth$ differential operators on the 2D spherical full-sky. For a detailed derivation see Castro et al. (2005).

4.2.3. Lensing fields in the spherical Fourier-Bessel basis

Since we are interested in computing the power spectra of the lensing fields, we must choose a convenient basis for the spectral expansion. The natural choice of basis functions are eigenfunctions of the Laplace operator in spherical coordinates, i.e. products of spherical harmonics and spherical Bessel functions $\{Y_{\ell m}(\theta,\phi)\, j_\ell(k\chi)\}$[9] , since this choice leads to a very simple relationship between the coefficients of the gravitational potential and those of the density contrast field, related by Poisson's equation (cf. Eq. 4.74). Spherical coordinates are a sensible choice for a number of reasons: the shear field has spin-weight 2 on the angular sky, and by virtue of being an integrated effect, the lensing fields are isotropic on the sky but not homogeneous. Furthermore, when comparing to observations, the selection function for a survey can usually be separated into a radial and angular part, with photometric redshifts introducing purely radial errors in galaxy positions (cf. Eqs. 5.24 and 7.28).

The Bessel-harmonic transform in the basis $\{Y_{\ell m}(\theta,\phi)\, j_\ell(k\chi)\}$ for a scalar field ϕ is given by

$$\phi(\vec{\chi}) = \sqrt{\frac{2}{\pi}} \sum_\ell \sum_{m=-\ell}^{\ell} \int \mathrm{d}k\, k^2\, j_\ell(k\chi)\, \phi_{\ell m}(k)\, Y_{\ell m}(\theta,\phi) \tag{4.28}$$

$$\phi_{\ell m}(k) = \sqrt{\frac{2}{\pi}} \int \mathrm{d}^3\vec{\chi}\, \phi(\chi)\, j_\ell(k\chi)\, Y^*_{\ell m}(\theta,\phi). \tag{4.29}$$

This expression is valid e.g. for the lensing potential ϕ and the convergence field κ, since they are both spin-0 scalars. Using the Bessel-harmonic expansion of the potentials in Eq. 4.24, we find that the lensing and gravitational potential expansion coefficients are related by

$$\phi_{\ell m}(k) = \sqrt{\frac{2}{\pi}} \int \phi(\vec{\chi})\, Y^*_{\ell m}(\hat{\mathbf{n}})\, j_\ell(k\chi)\, \mathrm{d}^3\chi \tag{4.30}$$

$$= \frac{2}{c^2}\sqrt{\frac{2}{\pi}} \int \mathrm{d}^3\chi \int_\chi^\infty \frac{\chi-\chi'}{\chi\chi'}\, \Phi(\vec{\chi}')\, \mathrm{d}^3\chi'\, Y^*_{\ell m}(\hat{\mathbf{n}})\, j_\ell(k\chi) \tag{4.31}$$

$$= \frac{2}{c^2}\sqrt{\frac{2}{\pi}} \int \mathrm{d}^3\chi \int_\chi^\infty \frac{\chi-\chi'}{\chi\chi'}\, \sqrt{\frac{2}{\pi}} \sum_{\ell m} \int_0^\infty k'^2 \mathrm{d}k' \Phi_{\ell' m'}(k')\, Y_{\ell' m'}(\hat{\mathbf{n}}')\, j_{\ell'}(k'\chi')\, \mathrm{d}\chi'\, Y^*_{\ell' m'}(\hat{\mathbf{n}})\, j_\ell(k\chi) \tag{4.32}$$

$$= \frac{4}{\pi c^2} \int_0^\infty \int_0^\infty \int_\chi^\infty \frac{\chi-\chi'}{\chi\chi'} j_{\ell'}(k'\chi')\mathrm{d}\chi' \chi^2 j_\ell(k\chi)\mathrm{d}\chi k'^2 \mathrm{d}k'\, \Phi_{\ell m}(k') \tag{4.33}$$

$$= \frac{4}{\pi c^2} \int_0^\infty \mathrm{d}k' k'^2 \int_0^\infty \mathrm{d}\chi\, \chi\, j_\ell(k\chi) \int_0^\chi \mathrm{d}\chi' \frac{\chi-\chi'}{\chi'} j_\ell(k'\chi')\Phi_{\ell m}(k') \tag{4.34}$$

$$= \frac{4}{\pi c^2} \int \mathrm{d}k'\, k'^2 \eta_\ell(k,k')\Phi_{\ell m}(k'). \tag{4.35}$$

[9] The spherical Bessel functions $j_\ell(x)$ of order $\ell \in \mathbb{N}$ can be derived from the Bessel functions $J_\ell(x)$

$$j_\ell(x) := \frac{\pi}{2x} J_{\ell+\frac{1}{2}}(x),$$

or, using another representation,

$$j_\ell(x) = (-1)^\ell x^\ell \left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^\ell j_0(x) = (-1)^\ell x^\ell \left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^\ell \frac{\sin x}{x}.$$

The J_ℓ functions are defined as solutions of the Bessel differential equation

$$x^2\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + x\frac{\mathrm{d}y}{\mathrm{d}x} + (x^2-\ell^2)y = 0.$$

where we made us of the normalization of the spherical harmonics $\int \mathrm{d}\Omega Y_{\ell m}(\hat{\mathbf{n}}) Y^*_{\ell' m'}(\hat{\mathbf{n}}) = \delta^K_{\ell\ell'}\delta^K_{mm'}$ and we defined

$$\eta_\ell(k,k') = \int \chi^2 d\chi \int d\chi' \frac{\chi-\chi'}{\chi\chi'} j_\ell(k\chi) j_\ell(k'\chi'). \tag{4.36}$$

The transform 4.28 is suitable for scalars like the lensing potential ϕ, the gravitational potential Φ and the density contrast δ. It may also be applied to each component of the shear $\gamma_{1,2}$. It is, however, more elegant to introduce a set of basis functions that share the spin properties of the field that is to be expanded. By this means the expansion coefficients are scalars, while the behaviour of the tensor under transformations is encoded in the basis functions. For a generic spin-s field the spin-weighted spherical harmonics ${}_sY_{\ell m}(\hat{\mathbf{n}})$ (Newman & Penrose, 1962) are the natural choice. As the derivative ð acts as a spin-raising operator, they are defined as

$${}_sY_{\ell m} = \sqrt{\frac{(l-s)!}{(l+s)!}} \eth^s Y_{\ell m}. \tag{4.37}$$

This description encompasses the scalar spherical harmonics for $s = 0$ and generally defines fields of spin s. The derivative ð and its complex conjugate $\bar{\eth}$ raise and lower the spin by one, respectively:

$$\eth({}_sY_{\ell m}) = +\sqrt{(l-s)(l+s+1)}\,{}_{s+1}Y_{\ell m}, \tag{4.38}$$

$$\bar{\eth}({}_sY_{\ell m}) = -\sqrt{(l+s)(l-s+1)}\,{}_{s-1}Y_{\ell m}. \tag{4.39}$$

Like the standard spherical harmonics, the spin-weighted spherical harmonics are orthonormal,

$$\int \mathrm{d}\Omega\, {}_sY_{\ell m}(\hat{\mathbf{n}})\, {}_sY^*_{\ell' m'}(\hat{\mathbf{n}}) = \delta^K_{\ell\ell'}\delta^K_{mm'}. \tag{4.40}$$

The spin-2 shear field can therefore be expanded in the basis $\{{}_{\pm2}Y_{\ell m}(\theta,\phi) j_\ell(kr)\}$, defining

$${}_{\pm2}\gamma_{\ell m}(k) = \sqrt{\frac{2}{\pi}} \int \chi^2 \mathrm{d}\chi \int \mathrm{d}\Omega \gamma(\chi,\hat{\mathbf{n}}) j_\ell(k\chi)\, {}_{\pm2}Y^*_{\ell m}(\hat{\mathbf{n}}). \tag{4.41}$$

It is convenient to rotate the shear expansion coefficients ${}_{\pm2}\gamma(\mathbf{r})$ of Eq. 4.25 into parity-$(-1)^\ell$ E-mode and parity-$(-1)^{\ell+1}$ B-mode components (see Fig. 4.1),

$$\gamma^E_{\ell m}(k) = \frac{1}{2}[{}_2\gamma_{\ell m}(k) +_{-2} \gamma_{\ell m}(k)] = \frac{1}{2}\sqrt{\frac{2}{\pi}} \int \mathrm{d}^3\chi\, \phi(\vec{\chi})\, k\, j_\ell(k\chi)\, ({}_2Y^*_{\ell m} +_{-2} Y^*_{\ell m}) \tag{4.42}$$

$$\gamma^B_{\ell m}(k) = \frac{i}{2}[{}_2\gamma_{\ell m}(k) -_{-2} \gamma_{\ell m}(k)] = \frac{1}{2}\sqrt{\frac{2}{\pi}} \int \mathrm{d}^3\chi\, \phi(\vec{\chi})\, k\, j_\ell(k\chi)\, ({}_2Y^*_{\ell m} -_{-2} Y^*_{\ell m}) \tag{4.43}$$

Gravitational lensing sourced from a real scalar potential ϕ predicts E-mode- only shear, whilst B-modes are expected to enter only at second order in ϕ/c^2. Recovering B-modes hence provides a powerful test of residual systematics in a weak lensing analysis, which could introduce spurious B-modes as a leading order effect. Owing to the simple relationships between shear, convergence and the lensing potential (in terms of ð operators) the shear and convergence expansions are related to the lensing potential coefficients as [10]

$${}_{\pm2}\gamma_{\ell m}(k) = \left[\frac{(l+2)!}{(l-2)!}\right]^{1/2} \phi_{\ell m}(k) \simeq \frac{1}{2} l^2 \phi_{\ell m}(k), \tag{4.44}$$

$$\gamma^E_{\ell m}(k) = \left[\frac{(l+2)!}{(l-2)!}\right]^{1/2} \phi_{\ell m}(k) \simeq \frac{1}{2} l^2 \phi_{\ell m}(k), \quad \gamma^B_{\ell m}(k) = 0 \tag{4.45}$$

$$\kappa_{\ell m}(k) \frac{1}{2}\ell(\ell+1)\phi_{\ell m}(k) \simeq \frac{1}{2}\ell^2 \phi_{\ell m}(k). \tag{4.46}$$

[10] The ℓ-dependent factors come from the action of ð operators on spherical harmonics: $\eth\eth Y_{\ell m} = \left[\frac{(l+2)!}{(l-2)!}\right]^{1/2}_2 Y_{\ell m}$ and $(\eth\bar{\eth} + \bar{\eth}\eth) Y_{\ell m} = \ell(\ell+1) Y_{\ell m}$.

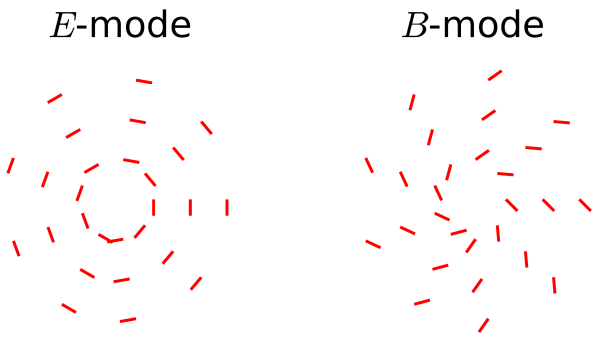

Figure 4.1.: Schematic demonstration of irrotational/pure-divergence E-mode pattern versus the divergenceless/pure-curl B-mode patterns. Weak lensing predicts pure-E-mode shear to leading order in Φ/c^2 , where B-modes enter only at second order. The lack of B-modes predicted from weak lensing makes B-modes a useful test of residual systematic effects, which can introduce spurious B-modes.

The two coefficients ${}_{+2}\gamma_{\ell m}$ and ${}_{-2}\gamma_{\ell m}$ are therefore identical, as expected since the potential ϕ is a real function.

4.2.4. Lensing spectra in the spherical case

Armed with the expansion in the Bessel-harmonic basis of the lensing potential, the shear and convergence field we set out to evaluate their power spectra. In this section we will focus on the potential power spectra, noting that the simple mappings from ϕ to κ and γ in Bessel-harmonic space given in Eqs. 4.44, 4.45, 4.46 can be used to convert between the various lensing power spectra.

Lensing potential power spectrum

We would like to calculate the covariance between the spherical Fourier coefficients $\phi_{\ell m}(k)$

$$\left\langle \phi_{\ell m}(k)\phi^*_{\ell' m'}(k')\right\rangle = \int d\Omega \int d\Omega' \left\langle \phi(\vec{k})\phi^*(\vec{k}')\right\rangle Y_{\ell m}(\hat{k})Y^*_{\ell m}(\hat{k}') \tag{4.47}$$

and we would like to link it to the power spectrum $P_{\phi\phi}$ of the Fourier modes, defined by

$$\left\langle \phi(\vec{k})\phi^*(\vec{k}')\right\rangle = (2\pi)^3\delta_D^{(3)}(\vec{k}-\vec{k}')P_{\phi\phi}(k). \tag{4.48}$$

In our conventions, the Fourier transform of the potential and its inverse are defined as

$$\hat{\Phi}(\vec{k}) = \int \mathrm{d}^3\chi\, \Phi(\vec{\chi})\mathrm{e}^{\mathrm{i}\vec{k}\cdot\vec{\chi}} \tag{4.49}$$

$$\Phi(\vec{\chi}) = \int \frac{\mathrm{d}^3k}{(2\pi)^3}\hat{\Phi}(\vec{k})\mathrm{e}^{-\mathrm{i}\vec{\chi}\cdot\vec{k}}, \tag{4.50}$$

where the integration range for both integrals is $\mathbb{R}^3$. We notice that for the normalised Dirac delta function δ_D the following holds:

$$\int \mathrm{d}^3k\delta_D(\vec{k}) = 1 = \int_0^\infty \mathrm{d}k \int_\Omega \mathrm{d}\Omega_k k^2\delta_D(\mathbf{k}) = \int_0^\infty \mathrm{d}k k^2\delta_D(k)\delta_D(\hat{\mathbf{k}}), \tag{4.51}$$

from which we see that

$$\delta_D^{(3)}(\vec{k}) = \frac{1}{k^2}\delta_D(k)\delta_D(\hat{k}). \tag{4.52}$$

Also, we notice that starting from the plane wave expansion, which decomposes a plane wave into a sum of spherical waves

$$e^{-i\vec{k}\cdot\vec{\chi}} = \sum_{\ell}(2\ell+1)i^{\ell}\, j_{\ell}(k\chi)P_{\ell}(\cos(\gamma)), \qquad \gamma = \hat{k}\cdot\hat{\chi}, \tag{4.53}$$

using the vector addition theorem for the spherical harmonics $Y_{\ell m}$

$$P_{\ell}(\cos(\gamma)) = \frac{4\pi}{2\ell+1}\sum_{m} Y_{\ell m}(\hat{k})Y^{*}_{\ell m}(\hat{\chi}) \tag{4.54}$$

we can explicitly derive the Rayleigh expansion for spherical waves

$$e^{-i\vec{k}\cdot\vec{\chi}} = 4\pi\sum_{l} i^{l}\, j_{l}(k\chi)\sum_{m} Y_{lm}(\hat{k})Y^{*}_{lm}(\hat{\chi}). \tag{4.55}$$

For a generic field (in our case the lensing potential) $\phi(\vec{\chi})$

$$\phi(\vec{\chi}) = \frac{1}{(2\pi)^3}\int \mathrm{d}^3k\,\phi(\vec{k})\, e^{\mathrm{i}\vec{k}\cdot\vec{\chi}} \tag{4.56}$$

$$= \frac{1}{(2\pi)^3}\int k^2\mathrm{d}\Omega_k\,\mathrm{d}k\,\phi(k)\,4\pi\sum_{\ell m} i^{\ell}\, j_{\ell}(k\chi)Y_{\ell m}(\hat{n}_k)Y_{\ell m}(\hat{n}_{\chi}) \tag{4.57}$$

$$= \frac{1}{2\pi^2}\int k^2\mathrm{d}\Omega_k\,\mathrm{d}k\,\phi(k)\sum_{\ell m} i^{\ell}\, j_{\ell}(k\chi)Y_{\ell m}(\hat{n}_k)Y_{\ell m}(\hat{n}_{\chi}). \tag{4.58}$$

It follows that

$$\phi_{\ell m}(k) = i^{\ell}k\int \mathrm{d}\Omega_k \frac{1}{\sqrt{8\pi^3}}\,\phi(k)\,Y_{\ell m}(\hat{n}_k). \tag{4.59}$$

Therefore, we can write for the covariance of the spherical Fourier-Bessel modes

$$\left\langle \phi_{\ell m}(k)\phi^{*}_{\ell' m'}(k')\right\rangle = i^{\ell}i^{*\ell'}\int \mathrm{d}\Omega_k \frac{1}{\sqrt{8\pi^3}}Y_{\ell m}(\hat{n}_k)\int \mathrm{d}\Omega_{k'}\frac{1}{\sqrt{8\pi^3}}Y^{*}_{\ell' m'}(\hat{n}'_k)(2\pi)^3 P_{\phi\phi}(k)\delta_D(k-k') \tag{4.60}$$

$$= \int \mathrm{d}\Omega_k\, Y_{\ell m}(\hat{n}_k)\, Y^{*}_{\ell' m'}(\hat{n}'_k)\, P_{\phi\phi}(k)\,\delta_D(k-k') \tag{4.61}$$

$$= P_{\phi\phi}(k)\,\delta_{\ell\ell'}\,\delta_{mm'}\,\delta_D(k-k'). \tag{4.62}$$

This expression is valid for the lensing potential power spectrum $P_{\phi\phi}$ and an analogous expression is valid for the gravitational potential power spectrum $P_{\Phi\Phi}$.

3D lensing potential power spectra

Taking the covariance of the expansion coefficients in Eq. 4.29, using Eq. 4.35 to connect the coefficients of the lensing and gravitational potential and appealing to isotropy, $\langle\phi_{\ell m}(k)\phi_{\ell' m'}(k')\rangle = C_{\ell}(k,k')\delta_{\ell\ell'}\delta_{mm'}$, we obtain the 3D weak lensing power spectrum coefficients $C_{\ell}(k_1,k_2)$,

$$\boxed{C_{\ell}(k_1,k_2) = \frac{16}{\pi^2c^4}\int_0^{\infty} k^2\,\mathrm{d}k\,\eta_{\ell}(k_1,k)\,\eta_{\ell}(k_2,k)} \tag{4.63}$$

where

$$\eta_{\ell}(k,k') = \int \chi^2 d\chi \int d\chi' \frac{\chi-\chi'}{\chi\chi'}\, j_{\ell}(k\chi)\, j_{\ell}(k'\chi')\,\sqrt{P_{\Phi\Phi}(k;\chi')}. \tag{4.64}$$

The η matrices represent the lensing 'efficiency' in the spherical Fourier-Bessel formalism.

At this point we could use the Poisson equation Eq. 4.74 to relate the gravitational potential power spectrum to the matter power spectrum, which describes the correlation between the Fourier modes of the density contrast. We will present this connection in Sec. 5.1.2, to highlight the modifications to the Poisson equation required in a modified gravity context. Here we will stick to the lensing potential power spectrum to highlight how all lensing spectra can be derived from this one, even in a modified gravity context as we show in Sec. 5.1.1.

Before proceeding, we note that in Eq. 4.63 we have actually assumed for convenience that the gravitational potential power spectrum can be accurately approximated by

$$P_{\Phi\Phi}(k;\chi,\chi') \simeq \sqrt{P_{\Phi\Phi}(k;\chi)\,P_{\Phi\Phi}(k;\chi')}. \tag{4.65}$$

This can be done because the correlations in the potential field are significantly non-zero for small separations (much smaller than the speed of light times the timescale over which it evolves), which can be seen by noting that Poisson's equation implies that

$$\Phi(\vec{k};\chi) = -\frac{3\Omega_m H_0^2}{2a(t)k^2}\,\delta(\vec{k};\chi) \tag{4.66}$$

and so

$$\left\langle \Phi(\vec{k};\chi)\Phi^*(\vec{k'};\chi')\right\rangle = \left(\frac{3\Omega_m H_0^2}{2}\right)^2 \int \frac{d^3\chi d^3\chi'}{a(t)a(t')} \frac{\left\langle \delta(\vec{k};\chi)\delta^*(\vec{k'};\chi')\right\rangle}{k^2 k'^2} e^{-i\vec{k}\cdot\vec{\chi}+i\vec{k'}\cdot\vec{\chi'}} \tag{4.67}$$

and the correlation of δ is restricted to small scales $|\vec{\chi}-\vec{\chi'}| \leq 100$Mpc. The lookback time over such a distance is small, so we can approximate $\vec{\chi} \simeq \vec{\chi'}$ (or $t \simeq t'$). Therefore we can replace the power spectra $P_{\Phi\Phi}(k;\chi,\chi')$ by either $P_{\Phi\Phi}(k;\chi)$ or $P_{\Phi\Phi}(k;\chi')$. For algebraic convenience, we choose the geometric mean of the power spectra, which allows us to separate two internal integrals. A further justification will be apparent from looking at the shape of the covariance matrices in Sec. 5.1.2 (cf. Fig. 5.4): the Bessel functions cut off long-wavelength contributions with $k \leq \ell/\chi_{\rm max}$ where $\chi_{\rm max}$ is the extent of the survey.

The type of cross-power spectrum in Eq. 4.65 between different cosmological epochs χ and χ' was introduced in Castro et al. (2005) and recently studied in Kitching & Heavens (2017). The oscillating Bessel functions in equation (4.63) ensure that only relatively close epochs contribute to the lensing potential correlation. This is sensible since observed light rays from two galaxies at different positions on the sky that necessarily converge at the observer today, pick up the density fluctuations at similar times while propagating through the large-scale structure. A similar argument has been made in Bartelmann & Schneider (2001): since the matter power spectrum scales with k for $k \to 0$, there is decreasing power towards larger and larger scales. In particular, the correlation of cosmic fields decreases strongly above a coherence scale $|\chi - \chi'| \gtrsim L_{\rm coh}$ which is significantly smaller than the horizon scale $\chi_{\rm H} = c/H_0$.

Tomographic lensing potential power spectra

By virtue of being an integrated effect along the line-of-sight (cf. Eq. 4.24), at moderate redshifts the lensing fields vary relatively slowly with z. An analysis that bins the field into a small number of broad redshift bins can bring in a substantial compression of the 3D fields with relatively small information loss. Tomographic weak lensing involves dividing the observed galaxies into a number of redshift slices i, where the sub-populations in each slice follow redshift distributions $n^{(i)}(z)\mathrm{d}z = p^{(i)}(\chi)\mathrm{d}\chi$. The lensing field probed by each tomographic slice is hence the full 3D field ϕ averaged over the redshift distributions $n^{(i)}(z)$, i.e.,

$$\phi^{(i)}(\theta,\phi) = \int_0^\infty \phi(\chi)p^{(i)}(\chi)\mathrm{d}\chi, \tag{4.68}$$

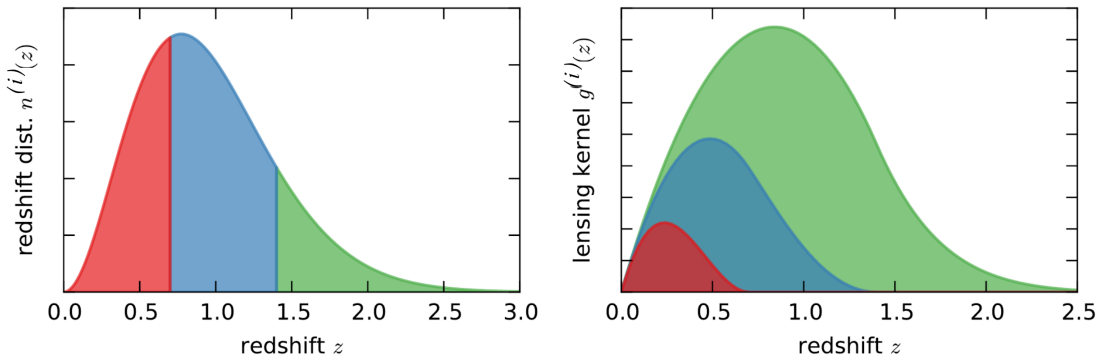

Figure 4.2.: Typical source redshift distribution separated into three tomographic slices $n^{(i)}$ (left), and their associated lensing kernels $g^{(i)}$ (right). The lensing efficiency peaks around half way between the observer ($z = 0$) and the source plane.

where the $p^{(i)}(\chi)$ are normalized to unity $\int p^{(i)}(\chi)\mathrm{d}\chi = 1$. Hence, the full 3D lensing field is compressed into a set of 2D tomographic fields $\{\phi^{(i)}(\theta,\phi)\}$. Replacing $\phi(\chi)$ with its Bessel-harmonic expansion of Eq. 4.28, and again appealing to isotropy $\left\langle \phi^{(i)}_{\ell m}\phi^{(j)}_{\ell' m'}\right\rangle = C_{\ell,ij}\delta_{\ell\ell'}\delta_{mm'}$, we find the tomographic power spectra between bins i and j:

$$\boxed{C_{\ell,ij} = \frac{8}{\pi c^2}\int k^2\,\mathrm{d}k\,\mathrm{d}\chi\,\mathrm{d}\chi' g^{(i)}(\chi)\,g^{(j)}(\chi')\,j_\ell(k\chi)\,j_\ell(k\chi')\,\sqrt{P_{\Phi\Phi}(k,\chi)}\,\sqrt{P_{\Phi\Phi}(k,\chi')}} \tag{4.69}$$

where the lensing kernels $\{g^{(i)}(\chi)\}$ are given by,

$$g^{(i)}(\chi) = \chi\int_\chi^\infty \mathrm{d}\chi' p^{(i)}(\chi')\frac{\chi'-\chi}{\chi'}. \tag{4.70}$$

The tomographic weak lensing power spectra are integrals of the gravitational potential power spectrum, over a geometric kernel. The lensing kernels determine which structures along the line-of-sight the lensing power spectra are most sensitive to, for given source redshift populations. Fig. 4.2 shows the lensing kernels for three redshift bins in a typical 3-bin tomography situation; for a given source plane redshift, most of the lensing signal comes from sources around half-way between the source plane and the observer.

Limber approximation

The exact expressions for the power spectra, Eqs. 4.63 and 4.69 are time consuming to evaluate accurately at high multipoles due to the rapid oscillations of the spherical Bessel functions. Many analyses adopt the Limber approximation instead, which is accurate at large ℓ and much easier to compute. To leading order in ℓ^{-1} we can expand

$$\int f(r)j_\ell(kr)\mathrm{d}r = \sqrt{\frac{\pi}{2\ell}}f\left(\frac{\ell}{k}\right) + O(\ell^{-2}). \tag{4.71}$$

In the Limber approximation, we effectively replace the spherical Bessel function with a delta-function,

$$j_\ell(k\chi) \to \sqrt{\frac{\pi}{2\ell}}\delta_D(\ell - k\chi). \tag{4.72}$$

The tomographic power spectra are then given by:

$$C_{\ell,ij} = \frac{4}{\ell^5 c^4}\int k^4\mathrm{d}k g^{(i)}\left(\frac{\ell}{\chi}\right)g^{(j)}\left(\frac{\ell}{\chi}\right)P_\Phi(k;\ell/k). \tag{4.73}$$

Shear power spectrum

Our ultimate goal is to provide expressions connecting explicitly the power spectrum of the shear and convergence field to the matter power spectrum. We will present only the equations in the case of tomography here, showing explicitly how to derive them from the power spectrum of the lensing potential. We could repeat the same exercise in the spherical Fourier-Bessel formalism, however in the 3D case there are additional observational effects that we would like to include in the formalism. We will comment extensively on them in Sec. 5.1.2, within our extended modified gravity scenario.

The 3D potential is related to the density contrast δ via the Poisson equation. Assuming General Relativity, this relation is written in Fourier space as

$$\hat{\Phi}(\vec{k};\chi) = -\frac{3}{2}\Omega_{\rm m}H_0^2 k^{-2}a^{-1}(\chi)\hat{\delta}(\vec{k};\chi), \tag{4.74}$$

where $\Omega_{\rm m}$ is the matter density parameter, H_0 the Hubble constant, $\vec{k}$ a 3D Fourier wave vector with modulus k being the comoving wave number, and a the scale factor with $a = 1$ today.

We recall from Sec. 4.2.3 that ${}_{\pm2}\gamma_{\ell m} = \frac{1}{2}\sqrt{\frac{(\ell+2)!}{(\ell-2)!}}\phi_{\ell m}$. The tomographic shear power spectrum is defined by

$$\left\langle {}_2\gamma_{\ell m,i}\ {}_2\gamma^*_{\ell' m',j}\right\rangle = \delta_{\ell\ell'}\delta_{mm'}C^{\gamma}_{ij}(\ell). \tag{4.75}$$

This is given by

$$\begin{aligned}\boxed{C^{\gamma}_{ij}(\ell)} =& \frac{1}{4}\frac{(\ell+2)!}{(\ell-2)!}\,C^{\phi}_{ij}(\ell)\\ =& \boxed{\frac{2}{\pi}\frac{(\ell+2)!}{(\ell-2)!}\left(\frac{3}{2}\Omega_{\rm m}\left(\frac{H_0}{c}\right)^2\right)^2\int_0^\infty\frac{{\rm d}\chi}{\chi}\frac{q_i(\chi)}{a(\chi)}\int_0^\infty\frac{{\rm d}\chi'}{\chi'}\frac{q_j(\chi')}{a(\chi')}\int_0^\infty\frac{{\rm d}k}{k^2}\,P_\delta(k,\chi,\chi')\,{\rm j}_\ell(k\chi)\,{\rm j}_\ell(k\chi')}\end{aligned} \tag{4.76}$$

where the matter power spectrum $P_\delta(k,\chi,\chi')$ is defined by

$$\left\langle\hat{\delta}(\vec{k};\chi)\hat{\delta}^*(\vec{k}';\chi')\right\rangle = (2\pi)^3\delta_{\rm D}(\vec{k}-\vec{k}')P_\delta(k;\chi,\chi'). \tag{4.77}$$

Eq. 4.77 reaches our goal to connect the shear power spectrum with the matter power spectrum.

Convergence power spectrum

The convergence is related to the lensing potential on the sphere via the product of spin-raising and spin-lowering ð operators, which are identical to the spherical Laplacian differential operator.

$$\kappa(\theta,\varphi) = \frac{1}{2}\eth\bar{\eth}\phi(\theta,\varphi) = \frac{1}{2}\nabla^2\phi(\theta,\varphi). \tag{4.78}$$

The spherical harmonics are eigenfunctions of the Laplacian,

$$\nabla^2{\rm Y}_{\ell m}(\theta,\varphi) = -\ell(\ell+1){\rm Y}_{\ell m}(\theta,\varphi). \tag{4.79}$$

The convergence power spectrum is then similar to the shear power spectrum (Eq. 4.75) with a different spherical pre-factor,

$$\boxed{C^{\kappa}_{ij}(\ell) = \frac{\ell(\ell+1)}{(\ell-1)(\ell+2)}C^{\gamma}_{ij}(\ell).} \tag{4.80}$$

The convergence power spectrum is thus larger than the shear power spectrum, but only by 10% for $\ell = 4$, 1% for $\ell = 14$, and less than 0.1% for $\ell > 45$.

4.2.5. A first comparison between tomographic and 3D cosmic shear

We conclude this section with an outlook on tomography and 3D cosmic shear. Loosely speaking, we can think of these two techniques in the following terms:

- in tomography the lensing effect first acts on the density field and subsequently a decomposition of the galaxies into redshift bins is performed;
- in 3D analyses we first decompose the density field in the spherical Fourier-Bessel formalism and then let the lensing effect take place.

Let us consider for example the convergence field κ in a pure 2D analysis, describing a change in size of the images of the sources. We know that it is related to the overdensity field δ by an integral along the line of sight: schematically,

$$\kappa(\chi_s) = \int_0^{\chi_s} W(\chi_s, \chi)\delta(\chi)\mathrm{d}\chi \tag{4.81}$$

where χ_s is the comoving radial distance of the source. As we discussed already, the lensing efficiency W is a function that peaks at about half the distance between the observer and the source (cf. Fig. 4.2). Thus, the main contribution to the integral is given by the "lenses" localised at that point, and little information can be gained by such a naive 2D approach. This is the motivation to introduce a binning in redshift in tomography. To the purpose of extracting more information on the localisation of the lenses, the convergence κ now takes values that are different for each bin i, therefore it acquires an index i:

$$\kappa_i(\chi_s) = \int_0^{\chi_s} W_i(\chi_s, \chi)\delta(\chi)\mathrm{d}\chi \tag{4.82}$$

where it is important to notice that also the W_i functions are now different from the original W: even if they are still defined on the entire range from the observer to the source, they are now modified so that they can peak at a different location for each redshift bin, enhancing the contribution given to the integral by the overdensity field in the corresponding bin. The binning in redshift allows now for a calculation of the auto- and cross-correlations between the convergence field in the bins, producing a covariance matrix $C_{ij} = \langle \kappa_i \kappa_j \rangle$, in general different from zero as a result of the overlap of the efficiency functions W_i.

Kitching et al. (2011) show that to convert from 3D cosmic shear to weak lensing tomography (in Limber approximation, as typically used) the Limber approximation must be applied to the full 3D shear estimator and a discretisation of physical modes through $k = \ell/r_i$ needs to be performed. Interestingly for a specific redshift bin at χ_i and a specific azimuthal ℓ-mode the tomographic approximation only probes a single physical k-mode $k_i = \ell/\chi_i$ from the full 3D shear field; in contrast in 3D cosmic shear we have control over the k and ℓ modes over the entire redshift range.

In tomography by fixing the distances of the tomographic binning we lose some flexibility over the physical wavenumbers probed, so there is a risk that either not all useful modes are included (increasing statistical errors), or that, for the nearby shells, the physical wavenumber range sampled extends to too high a value of k, where theoretical uncertainties become a potential source of systematic error. A way to reduce this contamination requires that the ℓ range chosen should be redshift-dependent – increasing $\ell_{\max} = \chi[z]k_{\max}$ for the distant shells, and reducing it for nearby shells.

This is the reason that lead Casas et al. (2016, 2017) to use a diagonal "cutoff" matrix to be added to the covariance matrix in their tomographic cosmic shear analysis, a matrix whose entries increase to very high values at the scales where the power spectrum $P(k)$ has to be cut to avoid the inclusion of uncertain non-linear scales. The matrix is chosen therefore to have further control on the inclusion

of non-linearities. Without that matrix, due to the redshift-dependent relation between k and ℓ coming from the Limber approximation, a very high $\ell_{\rm max}$ would correspond at low redshifts to a very high $k_{\rm max}$ where the accuracy of predictions for the matter power spectrum cannot be trusted anymore. Therefore the maximum multipole $\ell_{\rm max}$ considered in their analysis, to obtain the Fisher matrix as a sum over multipoles of many contributions (cf. Eq. 5.23) is limited by the minimum scale imposed either by $\ell_{\rm max}$ or $k_{\rm max}$, which is the maximum wavenumber considered in the matter power spectrum.

Isolation of scales

Let us now see why we expect the 3D weak lensing approach to be more efficient than tomography in isolating features of the matter power spectrum. This has been claimed many times in the literature (e.g. in Kitching et al., 2014) and is a crucial point in weak lensing analysis, deeply connected to the nature itself of the weak gravitational lensing effect: the lensing fields (both convergence and shear) are integrated quantities along the line of sight. Particularly, they are integrated measurements of the density contrast field and this integration procedure does involve a mixing of scales, because the integration kernels are famously 'broad'.

In lensing analysis it is not the convergence or shear fields itself to be probed, but their covariance. We showed e.g. in Sec. 4.2.4 that they are connected to the matter power spectrum, which makes cosmic shear an excellent tool to study in particular the growth of cosmic structure. One of the highest desirable properties of the covariance matrices would be to have a connection with the matter power spectrum that allows for isolation of features at different scales in P_δ. This way, one would for example be very sensitive to deviations in the matter power spectrum as predicted e.g. by different dark energy/modified gravity models. Schematically, one would like to have a relation lensing covariance - matter power spectrum of the following kind

$$C_\ell = \int \mathrm{d}k\, \delta(k)\, \delta(k)\, P_\delta(k) \tag{4.83}$$

where the delta functions help isolate features of the matter power spectrum. It turns out, as we show here, that the weights associated to the 3D cosmic approach indeed have this "delta"-like behaviour, at least in comparison with the weights associated to tomography. This explains why e.g. Grassi & Schäfer (2014) were able to isolate some of the Baryon Acoustic Oscillations features in the matter power spectrum using 3D weak lensing, while this is very hard to accomplish with tomography (Zhang et al., 2009).

Chapter 5

Testing dark energy and modified gravity with weak gravitational lensing

In this Chapter we investigate the constraining power of future cosmic shear surveys like Euclid (Laureijs et al., 2011) on the Horndeski class of modified gravity theories. We forecast in particular the expected sensitivity to the functions, introduced in Sec. 2.6.2, that fully describe the evolution of linear perturbations in Horndeski gravity. We produce our forecasts using the two main methods employed to analyse a cosmic shear survey: a traditional tomographic technique, where correlations between the lensing signal in different redshift bins are used to recover redshift information, and a less common 3D spherical Fourier-Bessel decomposition of the shear field, which allows for the inclusion of the full redshift information throughout the entire analysis.

We have already reviewed in Sec. 4.2 the tomographic and 3D approach assuming initially a standard General Relativity scenario, showing how to derive the expressions for the cosmic shear power spectrum in both formalisms. In Sec. 5.1 we proceed by extending the 3D approach to theories beyond the standard cosmological model, producing equations valid for very general modifications of gravity; we later concentrate on the Horndeski class. We present the methodology followed to obtain our forecasts in Sec. 5.2, before showing our results in Sec. 5.3 and drawing our conclusions in Sec. 5.4.

Our findings represent the first quantitative comparison on an equal footing between Fisher matrix forecasts on cosmological parameters (describing the concordance model as well as Horndeski gravity) for both a fully 3D and a tomographic analysis of cosmic shear surveys. The increased sensitivity of the 3D formalism, which we explicitly demonstrate by means of our Fisher matrix forecasts, comes from its ability to retain information on the source redshifts along the entire analysis.

Part of the content of this Chapter is based on the paper "Testing (modified) gravity with 3D and tomographic cosmic shear", by A. Spurio Mancini, R. Reischke, V. Pettorino, B.M. Schäfer, M. Zumalacárregui (Spurio Mancini et al., 2018a), accepted for publication in MNRAS and available online at arXiv:1801.04251.

5.1. 3d cosmic shear and tomography in modified gravity

To date, the 3D weak lensing approach has been applied to real data only in Kitching et al. (2007, 2014, 2016) for a ΛCDM model. Grassi & Schäfer (2014) investigated the possibility of detecting Baryon Acoustic Oscillation features in the cosmic matter distribution by 3D weak lensing; Zieser & Merkel (2016) studied the cross-correlation between the 3D weak lensing signal and the integrated Sachs-Wolfe effect; Camera et al. (2011) investigated the constraining power of 3D cosmic shear on a class of Unified Dark Matter models, where a single scalar field mimics both dark matter and dark energy, whereas Ayaita et al. (2012) employed 3D cosmic shear to explore the capability of future surveys to constrain dark energy clustering. While 3D weak lensing has been partially studied in

the context of modified gravity theories in Pratten et al. (2016) to constrain $f(R)$ chameleon models and environmentally dependent dilaton models, showing that for an all-sky spectroscopic survey the $f(R)$ parameter f_{R_0} can be constrained in the range $f_{R_0} < 5 \times 10^{-6}(9 \times 10^{-6})$ for $n = 1(2)$ with a 3σ confidence level, there has not been any application to a larger class of modified gravity theories. Alonso et al. (2016) forecast the sensitivity of future surveys to Horndeski theories using different probes, among which tomographic weak lensing.

In the following we will move to the description of the 3D cosmic shear formalism in a general modified gravity context. The aim is to produce an expression for the 3D cosmic shear power spectrum, for a general theory where the two Bardeen potentials Φ and Ψ may be different from each other. When dealing with linear perturbations acting on a Friedmann-Robertson-Walker metric in modified gravity one can assume spatial flatness and, considering only scalar perturbations (see Durrer & Tansella, 2016; Adamek et al., 2016, for vector and tensor perturbations), write the line element in Newtonian gauge as

$$\mathrm{d}s^2 = -\left(1 + 2\frac{\Phi}{c^2}\right)c^2\mathrm{d}t^2 + a^2(t)\left(1 - 2\frac{\Psi}{c^2}\right)\mathrm{d}\mathbf{x}^2 \tag{5.1}$$

with the Bardeen potentials Φ and Ψ. In General Relativity $\Phi = \Psi$ in absence of anisotropic stress, while this is in general not true in modified gravity.

5.1.1. 3D weak lensing signal from lensing potential

Information on the gravitational potential is encoded in a weighted projection along the line of sight, the lensing potential. In a modified gravity context, considering perturbations at the linear level, the lensing potential ϕ is related to the Bardeen potentials Ψ and Φ by

$$\phi(\chi, \hat{\mathbf{n}}) = \int_0^{\chi} \mathrm{d}\chi' \frac{\chi - \chi'}{\chi\chi'} \frac{\Phi(\chi, \hat{\mathbf{n}}) + \Psi(\chi, \hat{\mathbf{n}})}{c^2}, \tag{5.2}$$

where χ is a comoving distance, and the normalized vector $\hat{\mathbf{n}}$ selects a direction on the sky. Here and throughout the Chapter spatial flatness will be assumed, and the integration in Eq. 5.2 is carried out in Born approximation, i.e. along the unperturbed light path.

We investigate the relation between the coefficients $\phi_{\ell m}$ of the lensing potential and those of the Bardeen potentials $\Phi_{\ell m}$ and $\Psi_{\ell m}$. We insert Eq. 5.2 in the spherical Fourier transform of the lensing potential, Eq. 4.29, obtaining

$$\phi_{lm}(k) = \int k'^2 dk' \left(\eta_\ell^\Phi(k, k')\Phi_{\ell m}(k') + \eta_\ell^\Psi(k, k')\Psi_{\ell m}(k')\right), \tag{5.3}$$

where the lensing efficiency η has been defined in Eq. 4.36 and the superscript Φ, Ψ refers to the Bardeen potential on which the efficiency acts. Therefore the covariance of the lensing potential takes the form

$$C_l^{\phi\phi}(k, k') = \langle \phi_{\ell m}(k)\phi_{\ell m}(k')\rangle \tag{5.4}$$

$$= \langle \left[\eta_\ell^\Phi(k, k_1)\Phi(k_1) + \eta_\ell^\Psi(k, k_2)\Psi(k_2)\right]\left[\eta_l^\Phi(k', k_3)\Phi(k_3) + \eta_l^\Psi(k', k_4)\Psi\right]\rangle. \tag{5.5}$$

where the multiplication over a repeated k wavenumber stands short for integration over $k^2\mathrm{d}k$. We can rewrite this expression, exploiting the definition of the power spectra $P_{\Phi\Phi}$, $P_{\Phi\Psi}$ and $P_{\Psi\Psi}$ and making implicit the dependecies of the matrices on the wavectors for clarity of notation

$$C_\ell^{\phi\phi} = \eta_\ell^\Phi\eta_\ell^\Phi P_{\Phi\Phi} + \eta_\ell^\Psi\eta_\ell^\Phi P_{\Psi\Phi} + \eta_\ell^\Phi\eta_\ell^\Psi P_{\Phi\Psi} + \eta_\ell^\Psi\eta_\ell^\Psi P_{\Psi\Psi}. \tag{5.6}$$

Exploiting the symmetry of $P_{\Phi\Psi}$ and $P_{\Psi\Phi}$ and the equality of η^{Φ} and η^{Ψ}, the previous expression can be rewritten as

$$C_\ell^{\phi\phi} = \eta\,(P_{\Phi\Phi} + 2P_{\Psi\Phi} + P_{\Psi\Psi})\,\eta. \tag{5.7}$$

Eq. 5.7 tells us that in principle the lensing spectra can be obtained (e.g. in a Boltzmann code) extracting the power spectra of the Bardeen potentials Φ and Ψ. In the following we will move our discussion from the level of the (gravitational and lensing) potentials down to the level of the lensing observables and in particular the shear. This will allow us to better consider and reflect on observational effects such as the distribution in redshift of the sources and the error associated to our photometric estimates of the redshift, and include these aspects in our formalism. The neat effect of these features on our spectra will be to add computational complexity since the η matrices will have to be multiplied by other matrices that contain the distribution of sources and the photometric error. Each of this additional matrix entails one further integration and introduces further correlations between the amplitudes of the signal on different scales.

5.1.2. 3D cosmic shear

We present here a general formalism for a fully 3D expansion of the shear field that does not perform any binning in redshift. This is based on a spherical Fourier-Bessel decomposition of the shear, first introduced in lensing studies by Heavens (2003). Here we follow the notation and conventions of Zieser & Merkel (2016) and extend the presentation given there to a general modified gravity scenario characterised by the Bardeen potentials Φ and Ψ defined in eq. 5.1.

Inserting Eq. 5.2 and the definition of the shear in terms of $\eth$ derivatives of the lensing potential given by Eq. 4.25 in the spherical Fourier-Bessel expansion for the shear field Eq. 4.41, and applying a spherical Fourier-Bessel expansion to the Bardeen potentials Φ and Ψ, we can rewrite γ as

$$\begin{aligned}\gamma(\chi,\hat{\mathbf{n}}) = &\sqrt{\frac{2}{\pi}}\frac{1}{c^2}\int_0^\chi \mathrm{d}\chi'\,\frac{\chi-\chi'}{\chi\chi'} \\ &\times\int k^2\mathrm{d}k\sum_{\ell m}\sqrt{\frac{(\ell+2)!}{(\ell-2)!}}\left[\frac{\Phi_{\ell m}(k,\chi')+\Psi_{\ell m}(k,\chi')}{2}\right] j_\ell(k\chi')\,{}_2Y_{\ell m}(\hat{\mathbf{n}}),\end{aligned} \tag{5.8}$$

where the division by 2 comes from the prefactor in Eq. 4.25.

Poisson's equation can be used to link the coefficients in the spherical Fourier-Bessel decomposition of the lensing potential to those of the overdensity field $\delta_{\ell m}(k,\chi)$,

$$\frac{\Phi_{\ell m}(k,\chi)}{c^2} = -\frac{3}{2}\frac{\Omega_m}{(k\chi_H)^2}\frac{\delta_{\ell m}(k,\chi)}{a(\chi)}\mu(k,a(\chi)), \tag{5.9}$$

with the Hubble radius $\chi_H \equiv c/H_0$. Here the function $\mu(k,a(\chi))$ describes the mapping from the potential fluctuations to the density fluctuations. Eq. (5.9) can also be used as a parametrization of modified gravity theories (e.g. Planck Collaboration et al., 2016b). The latter approach, however, only holds in the quasi-static regime, where one neglects terms involving time derivatives in the Einstein equations for perturbations and keeps only spatial derivatives (Sawicki & Bellini, 2015; Baker & Bull, 2015). For the Euclid survey, it could be questionable if this approximation holds, given the large scales in principle accessible by the survey. The validity of the quasi-static approximation depends also on the single modified gravity model considered and its predictions for the sound speed of the additional scalar degree of freedom. That said, we stress that we are not using this parametrization, but rather take the potential and density statistics directly from `hi_class`, which does not use the quasi-static approximation.

z_m	$n_0[\text{arcmin}^{-2}]$	σ_z	$\Omega_{\text{survey}}[\text{deg}^2]$	ℓ_{min}	ℓ_{max}	k_{max}	n_{bins}
0.9	30	0.05	15000	10	1000	1.0	10

Table 5.1.: Specifications used in the Fisher matrix analysis for the Euclid survey: the median redshift z_m; the source density n_0; the error in photometric redshifts, $\sigma(z) = \sigma_z(1+z)$; the field size Ω_{sky}. ℓ_{min}, ℓ_{max} and k_{max} describe instead the minimum and maximum radial modes and the maximum angular mode, respectively, considered in the computation of the shear covariances (5.11) and the Fisher matrix (5.23). n_{bins} is the number of bins considered in the tomographic analysis.

The density field is statistically homogeneous and isotropic, characterised by a power spectrum which is diagonal in harmonic space

$$\left\langle \delta_{lm}(k,z)\delta^*_{\ell' m'}(k',z')\right\rangle = \frac{P_\delta(k,z,z')}{k^2}\delta^D(k-k')\delta^K_{\ell\ell'}\delta^K_{mm'}. \tag{5.10}$$

Using this, we can relate the covariance of shear modes to the matter power spectrum by

$$\boxed{\left\langle \bar{\gamma}_{lm}(k)\bar{\gamma}^*_{\ell' m'}(k')\right\rangle = \frac{9\Omega_m^2}{16\pi^4\chi_H^4}\frac{(\ell+2)!}{(\ell-2)!}\times\int\frac{\mathrm{d}\tilde{k}}{\tilde{k}^2}\,G_\ell(k,\tilde{k})\,G_\ell(k',\tilde{k})\,\delta^K_{\ell\ell'}\,\delta^K_{mm'}} \tag{5.11}$$

where

$$G_\ell(k,k') = \int \mathrm{d}z\, n_z(z)\, F_\ell(z,k)\, U_\ell(z,k'), \tag{5.12}$$

$$F_\ell(z,k) = \int \mathrm{d}z_p\, p(z_p|z)\, j_\ell[k\chi^0(z_p)], \tag{5.13}$$

$$U_\ell(z,k) = \frac{1}{2}\int_0^{\chi(z)}\frac{\mathrm{d}\chi'}{a(\chi')}\frac{\chi-\chi'}{\chi\chi'}\,j_\ell(k\chi')\,P_\delta^{1/2}\,(k,z(\chi)) \times\mu(k,a(\chi))\left[1+\frac{1}{\eta(k,a(\chi'))}\right]. \tag{5.14}$$

The bar over the $\gamma_{\ell m}$ coefficients stands for estimates of these quantities. This is done to keep into account practical considerations arising while trying to estimate the shear modes $\gamma_{\ell m}$ from real cosmic shear data, which consist of the shear $\gamma(z,\hat{\mathbf{n}})$ measured at discrete positions labelled by redshift z and angular coordinates $\hat{\mathbf{n}} = (\theta,\phi)$. An estimate for the amplitude $\gamma_{\ell m}(k)$ of a mode may then by obtained by dividing the three-dimensional space into cells and approximating the volume integral Eq. 4.41 by a sum. In the continuum, this estimate depends on the spatial distribution of sources in radial direction. In addition, redshifts are typically measured with photometry and therefore their associated errors need to be taken into account:

$$\bar{\gamma}_{\ell m}(k) = \sqrt{\frac{2}{\pi}}\int \mathrm{d}z\frac{n_z(z)}{4\pi}\int \mathrm{d}z_p p(z_p|z) j_\ell\left[k\chi(z_p)\right]\int \mathrm{d}\Omega\gamma(z,\hat{\mathrm{n}})\,{}_2Y^*_{\ell m}(\hat{\mathbf{n}}). \tag{5.15}$$

$\bar{\gamma}$ are estimates of the shear modes that, in addition to the pure lensing effect, keep into account the redshift distribution of galaxies and the redshift estimation error (see Sec. 5.2.2 for details on observational effects), as evidenced by the definition of the quantities in Eqs. 5.12, 5.13, 5.14, which contain the redshift distribution $n_z(z)$ of the lensed galaxies and the conditional probability $p(z_p|z)$ of estimating the redshift z_p given the true redshift z. These two elements and the lensing kernel contained in the function $U_\ell(z,k)$ introduce correlations between the amplitudes of the signal on different scales; the covariance matrix then acquires off-diagonal terms, the calculation of which is numerically involved. The basis of spherical Bessel functions leads to integrals with rapidly oscillatory kernels, which have

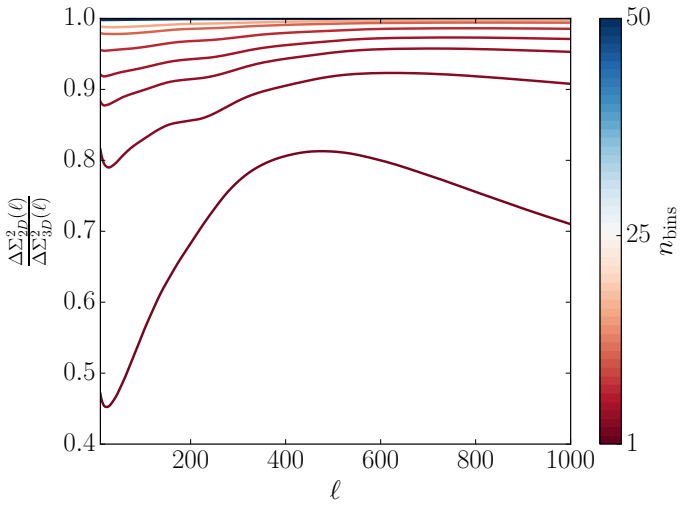

Figure 5.1.: Differential signal-to-noise ratio of a tomographic analysis relative to a 3D analysis. The number of tomographic bins is shown in the colour bar.

to be solved for a large number of parameter combinations. $\eta(k, a(\chi'))$ is defined as the ratio between the Bardeen potentials

$$\eta(k, a(\chi)) = \frac{\Phi(k, a(\chi))}{\Psi(k, a(\chi))}. \tag{5.16}$$

The $P_\delta^{1/2}(k, z(\chi))$ term comes from an approximation, introduced and justified in Castro et al. (2005), to calculate unequal-time correlators appearing in the comoving distance integrations by means of a geometric mean $P(k, z, z') \simeq \sqrt{P(k, z)\, P(k, z')}$ (see also Kitching & Heavens, 2017). This expression simplifies considerably in the linear regime of structure formation, retrieving the one presented in the seminal paper of Heavens (2003) where a product of the linear growth factors at different redshifts is present, acting on the matter power spectrum evaluated at the present time.

The noise term for the covariance matrix of the shear modes is given by the intrinsic ellipticity dispersion of source galaxies, as a result of the fact that the observed ellipticity ϵ is assumed to be the sum of the shear γ and the intrinsic ellipticity ϵ_S. The intrinsic ellipticity dispersion is given by $\left\langle \epsilon_S^2 \right\rangle = \sigma_\epsilon^2$. In the spherical Fourier-Bessel formalism (see Appendix A.1 for an explicit derivation), this gives

$$\boxed{\left\langle \gamma_{\ell m}(k)\, \gamma_{\ell' m'}(k') \right\rangle_{\rm SN} = \frac{\sigma_\epsilon^2}{2\pi^2} \int \mathrm{d}z\, n_z(z) j_\ell\left[k\chi_0(z)\right] j_{\ell'}\left[k'\chi_0(z)\right] \delta^K_{\ell\ell'} \delta^K_{mm'}} \tag{5.17}$$

and we set $\sigma_\epsilon = 0.3$. This expression for the noise holds only in absence of intrinsic alignments, i.e. assuming that the intrinsic ellipticities of galaxies are uncorrelated (see Merkel & Schäfer, 2013, for a study of intrinsic alignments in 3D weak lensing).

5.1.3. Tomography

Instead of keeping track of the photometric redshift error, as done in the 3D approach, by means of the probability $p(z_p|z)$ of estimating the redshift z_p conditional on the true redshift z, another possibility is to assign every galaxy to a redshift bin. In this case, as opposed to Eq. 5.11, the flat sky tomographic cosmic shear power spectrum in tomographic bins i and j is given by

$$C^\kappa_{ij}(\ell) = \int \frac{\mathrm{d}\chi}{\chi^2} W_i(\ell/\chi, \chi) W_j(\ell/\chi, \chi)\, P_\delta(\ell/\chi, \chi), \tag{5.18}$$

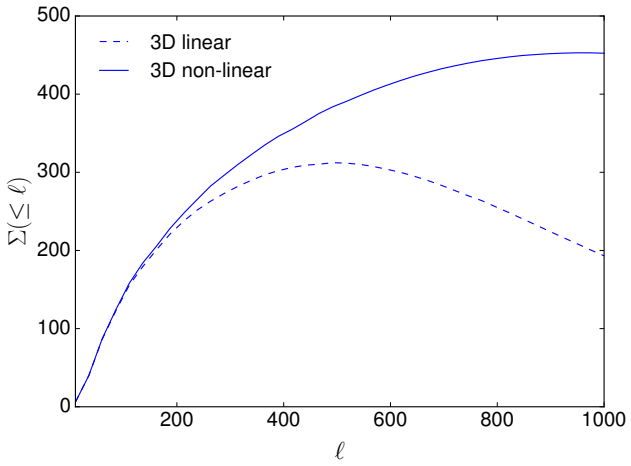

Figure 5.2.: Differential signal-to-noise ratio (Eq. 5.28), i.e. Σ^2 gained at each multipole ℓ: impact of the non-linear power on the 3D lensing signal. The solid curve shows the differential signal-to-noise with a non-linear power spectrum, while the dashed curve refers to a linear one.

where we used the Limber projection. The lensing efficiency function $W_i(\ell/\chi,\chi)$ is defined as

$$W_i(\ell/\chi,\chi) = \frac{3\Omega_m}{4\chi_H^2}\int_\chi^\infty \mathrm{d}\chi' \frac{\mathrm{d}z}{\mathrm{d}\chi'}\frac{n_i(z(\chi'))}{a(\chi')}\frac{\chi-\chi'}{\chi\chi'} \times\left(1+\frac{1}{\eta(\ell/\chi,\chi')}\right)\mu(\ell/\chi,\chi'), \tag{5.19}$$

$n_i(z(\chi))$ being the distance distribution of sources in the i-th bin normalised to one, $\int \mathrm{d}\chi\, n_i(z(\chi)) = 1$.

Observed spectra suffer from Poissonian noise due to the intrinsic ellipticity dispersion of galaxies σ_ϵ and their finite number n_0. Choosing our tomographic bins so as to have equal number of galaxies in each of them, the observed tomographic weak lensing spectrum is given by

$$\hat{C}_{ij}^\kappa(\ell) = C_{ij}^\kappa(\ell) + \frac{\sigma_\epsilon^2 n_{\mathrm{bins}}}{n_0}\delta_{ij}. \tag{5.20}$$

It should be noticed that comparing tomographic and 3D lensing must be done with some care since some approximations enter in Eq. (5.18). In particular, we made use of the flat sky approximation and the Limber projection (Kaiser, 1992, 1998; Loverde & Afshordi, 2008), neither of which is included in the 3D formalism. For a detailed discussion on this we refer to Kitching et al. (2017); Kilbinger et al. (2017); Lemos et al. (2017) for an excellent discussion of various approximations performed in cosmic shear analyses.

5.2. Forecasts methodology

5.2.1. Fisher matrix forecasts

In order to present forecasts for a Euclid-like experiment on the parameters considered, we perform a Fisher matrix analysis (Tegmark et al., 1997). Provided that the likelihood surface near the maximum is well approximated by a multivariate Gaussian, the Fisher matrix gives a realistic expectation of the foreseen error for a given experimental setting. The Fisher matrix is defined as the expectation value of the derivative of the logarithmic likelihood $\mathcal{L}$ with respect to the parameters θ_α:

$$F_{\alpha\beta} \equiv -\left\langle \frac{\partial^2 \ln\mathcal{L}}{\partial\theta^\alpha \partial\theta^\beta}\right\rangle, \tag{5.21}$$

evaluated at the maximum of the log-likelihood $\mathcal{L}$, which in a forecast analysis coincides with the reference fiducial model. Once we have the Fisher matrix, the Cramer-Rao bound $\Delta\theta_\alpha^2 \geq (\boldsymbol{F}^{-1})_{\alpha\alpha}$ gives a lower limit on the expected marginal error on the parameter θ_α. If the data are Gaussian distributed and the mean values vanish, the Fisher matrix can be calculated from the covariance matrix and its derivatives with respect to the parameters (Tegmark et al., 1997)

$$F_{\alpha\beta} = \frac{1}{2}\mathrm{Tr}\left[\boldsymbol{C}^{-1}\boldsymbol{C}_{,\alpha}\boldsymbol{C}^{-1}\boldsymbol{C}_{,\beta}\right]. \tag{5.22}$$

where derivatives have been denoted with a comma. Assuming full-sky coverage this expression can be simplified for a 3D weak lensing survey, as modes with different ℓ and m are uncorrelated ($\delta^K_{\ell\ell'}\delta^K_{mm'}$ in Eq. (5.11)), leading to

$$F_{\alpha\beta} = \frac{1}{2}\sum_\ell (2\ell+1)\mathrm{Tr}\left[\boldsymbol{C}_\ell^{-1}\boldsymbol{C}_{\ell,\alpha}\boldsymbol{C}_\ell\boldsymbol{C}_{\ell,\beta}\right], \tag{5.23}$$

as there are $2\ell + 1$ statistically independent $m-$ modes for each ℓ. Note that expressions 5.22 and 5.23 are only exact if the data, in this case the modes $\gamma_{\ell m}(k)$, follow a Gaussian distribution. This is not the case for high ℓ values, where structures due to non-linear clustering dominate the lensing signal (for a discussion on non-Gaussian statistics of the weak lensing field see e.g. Taruya et al., 2002; Joachimi et al., 2011; Clerkin et al., 2017). However, for the purpose of this Section this assumption will not be of any harm since the basic parameter dependencies are captured well enough within this approximation.

5.2.2. Observational effects and specifications

A 3D weak lensing analysis depends crucially on redshift estimation of the source galaxies, which for next generation surveys like Euclid (Laureijs et al., 2011) will be achieved using photometry, being the number of sources prohibitively high for spectroscopy. The estimated shear modes in Eq. 5.11 keep into account two observational effects, which are inherent in a redshift survey. The first one is described by the quantity G in Eq. 5.12 and represents the distribution in redshift of the galaxies, mainly due to the fact that they become fainter as redshift increases. For the source distribution we follow (Amendola et al., 2016) and choose in Eq. 5.12

$$n_z(z) \propto n_0 \left(\frac{\sqrt{2}}{z_m}\right)^3 z^2 \exp\left[-\left(\frac{\sqrt{2}z}{z_m}\right)^{3/2}\right], \tag{5.24}$$

where z_m is the median redshift of the survey and n_0 is the observed redshift-integrated source density n_0. The second observational effect, kept into account in the quantity F in Eq. 5.13, is the error associated to redshift estimation. This is described by the probability of estimating the reshift z_p given the measured redshift z. We take this probability distribution to be a Gaussian

$$p(z_p|z) = \frac{1}{\sqrt{2\pi}\sigma(z)}\exp\left[-\frac{(z_p - z)^2}{2\sigma^2(z)}\right], \tag{5.25}$$

with a redshift-dependent dispersion

$$\sigma(z) = \sigma_z(1+z). \tag{5.26}$$

If the sky coverage is not complete Eq. 5.23 is not completely correct, since the spherical basis is no longer orthogonal. Nonetheless, Eq. 5.23 is a good approximation by just multiplying the right hand side with the sky-fraction $f_{\rm sky} = \Omega_{\rm survey}/\Omega_{\rm sky}$. The choice we make for the specifications used in the analysis is summarized in Table 5.1.

5.2.3. Scales considered and non-linear corrections

The cuts in angular and radial scales that we perform, $\ell_{\rm max} = 1000$ and $k_{\rm max} = 1.0\,h/{\rm Mpc}$, are such that we avoid the deeply non-linear regime of structure growth. We demonstrate this point in Eq. 5.4, showing the signal and noise parts of the covariance matrices (Eq. 5.11 and Eq. 5.17) for two ℓ-modes, 10 and 1000, which correspond to the mininum and maximum multipole considered in our analysis, respectively. We notice how even for the higher ℓ case, the range of k scales considered justifies our choice to use the linear power spectrum for our analysis, since the higher-k part of the spectrum is dominated by the noise (notice the different orders of magnitude between the signal and noise contibutions). In Fig. 5.5 we plot only the diagonal contributions to the covariance matrices, distinguishing between the signal and noise parts, for $\ell = 10$ and $\ell = 1000$, our minimum and maximum angular multipoles. One can see how the orders of magnitude of the covariance matrices between different multipoles change and also how the dominance of the signal over the noise part gets inverted going from low to high ℓ values.

We compare forecasts obtained with a linear matter power spectrum in the calculation of the shear covariances to those obtained with a non-linear power spectrum. The current lack of solid understanding for non-linear corrections, in ΛCDM and even more in a modified gravity context, implies that any non-linear prescription should be employed with caution. The matter power spectra are produced using the `hi_class` code (Zumalacárregui et al., 2017), a modification of the CLASS Boltzmann solver (Lesgourgues, 2011) for Horndeski theories of gravity. In particular, `hi_class` allows the user to choose the parameterization for the $\alpha(\tau)$ functions which traces the evolution of the dark energy component, and the code takes then as input the proportionality coefficients $\hat{\alpha}$ (Eq. 2.39). The choice for the fiducial values of $\hat{\alpha}_B$ and $\hat{\alpha}_M$, reported in 5.2, is close enough to ΛCDM to represent General Relativity with an additional cosmological constant, without incurring in numerical difficulties in `hi_class` if the $\hat{\alpha}$ coefficients are all set to zero exactly. Since `hi_class` is a linear code it produces linearly evolved power spectra only. Non-linear corrections can however be incorporated by applying a non-linear transfer function using `halofit` (Smith et al., 2003; Takahashi et al., 2012; Bird et al., 2012) as implemented in `hi_class` or a more state of the art version, HMCODE, developed by Mead et al. (2015), which we employ for our non-linear forecasts. Both `halofit` and HMCODE however deal with non-linearities only in a setting where standard General Relativity is true. In order to get consistent constraints we therefore follow Alonso et al. (2016) and introduce a screening mechanism to recover General Relativity on small scales by a phenomenological modification of the $\alpha(\tau)$ functions, employing a Gaussian kernel in Fourier space with a characteristic scale k_V:

$$\alpha(\tau) \rightarrow \alpha(\tau, k) = \alpha(\tau) \exp\left(-\frac{1}{2}\left(\frac{k}{k_V}\right)^2\right) . \tag{5.27}$$

We marginalise over the scale $k_V = 0.1\,h/{\rm Mpc}$ at which the screening mechanism becomes effective. The fiducial choice for the screening is important in the sense that non-linear effects become important at scales smaller than $0.1\,h/{\rm Mpc}$. Additionally, Barreira et al. (2013) showed that the typical scale of Vainshtein screening is roughly at $0.1\,h/{\rm Mpc}$. A plot with two different choices of k_V can be found in Appendix A.3.

5.2.4. Implementation

Here we describe the structure of the C++ code that we used to produce our Fisher matrix forecasts, obtained tackling the 3D cosmic shear integrations by means of the Levin integration method discussed in the next Chapter. A diagram summarising the structure of the code is reported in Fig. 5.3:

- As a first step, the user specifies the fiducial cosmology in the `Cosmology` class and the parameters describing the survey considered, i.e. the redshift distribution $n(z)$ and the redshift

error probability distribution $p(z|z_p)$ (cf. Eqs. 5.24 and 5.25) in the `Survey` class. Specifically, the user should specify the n_0 parameter representing the integrated source density, the median redshift z_m and the width of the redshift error distribution σ_z. Given these values, the `Survey` class computes and normalises the source distribution $n(z) \propto n_0 \left(\frac{\sqrt{2}}{z_m}\right)^3 z^2 \exp\left[-\left(\frac{\sqrt{2}z}{z_m}\right)^{3/2}\right]$ and a photometry distribution $p(z_p|z)$, assumed to be Gaussian as in Eq. 5.25;

- Subsequently, the user specifies parameters relevant for the calculation of the covariance matrices and the Fisher matrix, i.e. the range of multipoles ℓ, from $\ell_{\rm min}$ to $\ell_{\rm max}$, for which the covariance matrices $C_\ell(k, k')$ need to be calculated and summed over in the expression for the Fisher matrix (cf. Eq. 5.23). The number of ℓ modes defines the spacing $\Delta\ell$ that appears in the expression for the Fisher matrix. We verified that the code is very stable against the choice of the number of ℓ modes considered. This is due to the smoothness of the signal-to-noise curve (see Fig. 5.1), which results from the high precision achieved thanks to the Levin integration;

- The number of Fourier modes for the covariance matrix, ranging from $k_{\rm min}$ to $k_{\rm max}$, has to be set. For the range $[5 \cdot 10^{-3}h/\text{Mpc}, 1.0h/\text{Mpc}]$ that we used in our forecasts, we recommend to use at least 500 k-values linearly spaced between those extremes, in order to make sure that the 3D cosmic shear signal is sufficiently sampled;

- A `Covariance` class and subsequent inherited classes will account for:
  - the distance-redshift relation, calculated by the class `Cosmometry`;
  - classes producing quantities relevant to the clustering of structure, i.e.

    - `Growth`: this produces the linear growth factor $D_+(a)$ (in GR, or $D_+(a, k)$ in modified gravity), sourcing it directly from `Class/HiClass` and storing it in tables after interpolation in a and k. The `Growth` class also stores interpolated table for μ, η and their screened versions;

    - `CDMspectrum`, standing for 'cold dark matter power spectrum', connects the `C++` code to `Class/HiClass`(written in `C`) via a `C/C++` wrapper;

    - the `Covariance` class itself produces the covariance matrix calculations, i.e. the integrations in Eqs. 5.12, 5.13, 5.14 by means of the Levin integration scheme, implemented in three classes `levinFunctions, levinBase, levinIteration`(not shown in Fig.5.3);

- For the statistical analysis, a `FisherMatrix` class accounts for
  - calling the covariance matrix class methods to

    - compute signal and noise part of the C_ℓ matrices;

    - evaluate the C_ℓ's at slightly bigger and slightly smaller values of each cosmological parameters, to be able to perform the differentiation of the covariance matrices by means of central finite differences;

    - compute the inverses C_ℓ^{-1};

    - assemble the cosmic shear covariance matrices, their derivatives and inverses to produce the Fisher matrix elements $F_{\alpha\beta}$ following Eq. 5.23.

5.3. Constraints on Horndeski gravity

Here we propose for the first time 3D cosmic shear as a probe of Horndeski theories of modified gravity. We analyse in detail the expected performance of a Euclid-like experiment, with the aim of

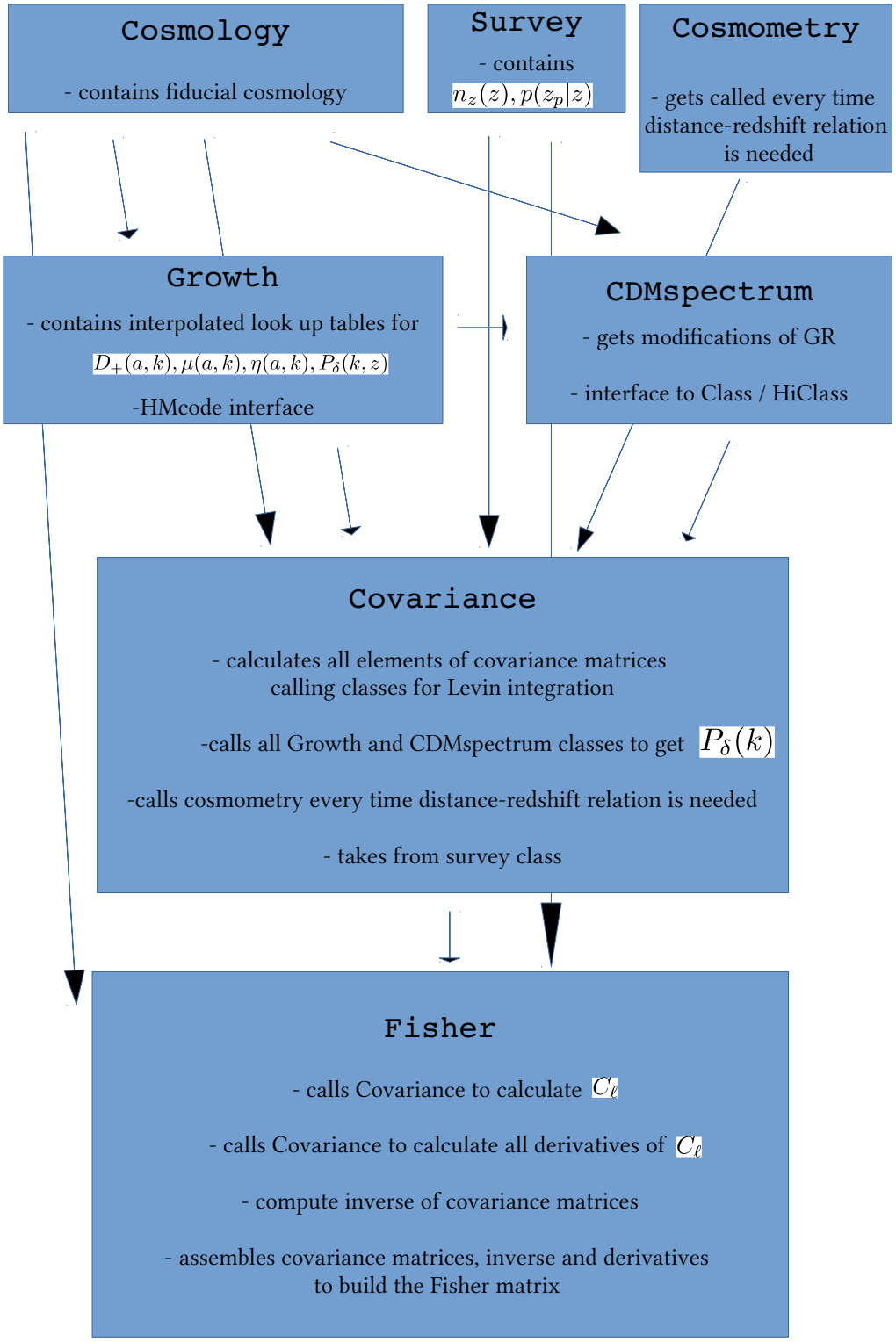

Figure 5.3.: Schematic map describing the mdular structure of the *C++* implementation for our Fisher forecasts.

	$\hat{\alpha}_K$	$\hat{\alpha}_\mathrm{B}$	$\hat{\alpha}_\mathrm{M}$	$\hat{\alpha}_T$	Ω_m	σ_8	h	n_s	Ω_b	$\sum m_\nu$
fiducial value	0.01	0.05	0.05	0.05	0.314	0.834	0.678	0.968	0.0486	0.05
error 3DWL linear, varying $\hat{\alpha}_K$ (fiducial value 0.01), varying $\hat{\alpha}_T$ (fiducial value 0.05)	137 (13564%)	0.82 (78%)	2.24 (213%)	5.62 (535%)	0.015 (4.78%)	0.045 (5.39%)	0.307 (45.28%)	0.101 (10.43%)	0.027 (55.55%)	0.484 (968%)
error tomography linear, $\hat{\alpha}_K = 0.01$, $\hat{\alpha}_T = 0$	/	0.57 (54%)	1.66 (158%)	/	0.017 (5.41%)	0.056 (6.71%)	0.257 (37.90%)	0.088 (9.09%)	0.022 (45.27%)	0.472 (944%)
error 3DWL linear, $\hat{\alpha}_K = 0.01$, $\hat{\alpha}_T = 0$	/	0.43 (41%)	1.32 (126%)	/	0.013 (4.14%)	0.042 (5.03%)	0.239 (35.25%)	0.080 (8.26%)	0.021 (43.21%)	0.408 (816%)
error 3DWL linear, $\hat{\alpha}_K = 10$, $\hat{\alpha}_T = 0$	/	0.48 (46%)	1.40 (133%)	/	0.014 (4.46%)	0.043 (5.15%)	0.238 (35.10%)	0.080 (8.26%)	0.021 (43.21%)	0.400 (800%)
error 3DWL non-linear, $\hat{\alpha}_K = 0.01$, $\hat{\alpha}_T = 0$	/	0.25 (24%)	0.55 (52%)	/	0.011 (3.50%)	0.010 (1.20%)	0.134 (19.76%)	0.038 (3.92%)	0.016 (32.92%)	0.229 (458%)

Table 5.2.: Marginalized errors $\sigma_i = \left[(\boldsymbol{F}^{-1})_{ii}\right]^{1/2}$, for the survey characteristics from 5.1. We compare a tomographic analysis with the 3D analysis using a linear power spectrum. Furthermore the influence of the value $\hat{\alpha}_K$ is investigated. Lastly the impact of non-linear clustering is quantified. When reporting the relative percentage error for the $\hat{\alpha}$ coefficients, we calculate it with respect to their fiducial values increased by one, as they are indeed expected to be $O(1)$ if one assumes the parameterization tracing the dark energy density fraction to model the time evolution of the $\hat{\alpha}(\tau)$ functions (Bellini et al., 2016).

forecasting the precision with which future stage IV surveys will be able to constrain this class of alternatives to General Relativity using cosmic shear datasets. We choose the parametrization of linear perturbations in Horndeski gravity that we already described in Sec. 2.6.2, first proposed by Gleyzes et al. (2013); Bellini & Sawicki (2014) and based on four functions of time only, which completely describe the evolution of linear perturbations once the background evolution is fixed. We model their time evolution assuming proportionality to the density fraction of dark energy and constrain the proportionality coefficients under the assumption of Gaussian likelihood. We simultaneously also place constraints on a set of standard cosmological parameters describing the evolution of the background, including the sum of the neutrino masses.

We produce our forecasts for both a fully 3D and a tomographic analysis of the measurements, with the aim of comparing the performances of the two methods on both modified gravity and standard cosmological parameters. Kitching et al. (2011) showed the relationship between weak lensing tomography and the 3D cosmic shear field, connected by the Limber approximation, a harmonic-space transform and a discretization in wavenumber. Our work presents for the first time a quantitative comparison on an equal footing between 3D and tomographic techniques for cosmic shear in terms of Fisher forecasts, showing that the 3D approach has more sensitivity than tomography to both standard and modified gravity cosmological parameters. We vary both the background cosmological parameters and those describing Horndeski theories, and consider only weak gravitational lensing as a cosmological observable, to test its power in constraining modified gravity theories without other probes and compare 3D and tomographic methodologies.

In this section we will investigate the signal strength of a weak lensing analysis carried out using the full photometric redshift information via the 3D method, as well as by using a tomographic technique. As already mentioned we calculate the tomographic lensing power spectrum using the Limber approximation; for a more detailed discussion we refer to Kitching et al. (2017). We then show the possible constraints on Horndeski cosmological models with survey specifications given in Tab. 5.1.

5.3.1. Signal to noise for 3D and tomographic weak lensing

Fig. 5.1 shows the differential signal-to-noise (SNR) curve for a tomographic survey relative to a 3D analysis as a function of the number of tomographic bins. The total signal-to-noise ratio is calculated as

$$\Sigma^2(\leq \ell) = f_{\rm sky} \sum_{\ell'=\ell_{\rm min}}^{\ell} \frac{2\ell'+1}{2} {\rm Tr}\left[\boldsymbol{C}_{\ell'}^{-1}\boldsymbol{S}_{\ell'}\boldsymbol{C}_{\ell'}^{-1}\boldsymbol{S}_{\ell'}\right] \equiv \sum_{\ell'=\ell_{\rm min}}^{\ell} \Delta\Sigma^2(\ell') \,, \tag{5.28}$$

where $\boldsymbol{S}$ is the signal covariance (5.11) or (5.18) only, while $\boldsymbol{C}$ refers to the sum of signal and noise, i.e. (5.11)+(5.17) or (5.20). The number of tomographic redshift bins is shown in the colour bar. Clearly an increase of the number of bins used increases the SNR, however, the gain in signal saturates for $n_{\rm bins} \approx 15$ due to the non-vanishing cross-correlation between the different bins (caused by the overlap of the redshift distributions of different bins). It should be noted that there is in principle an additional effect due to the finite width of the photometric redshift estimation. If the average bin width in the tomographic case is of the order of the width of the distribution of redshift estimation error, the correlation between neighbouring bins will be underestimated, thus producing artificially signal. This effect is, however, very small as long as σ_z is sufficiently small. In the 3D case this correlation is represented by the covariance (5.11).

Fig. 5.2 displays the impact of non-linear clustering: if one only considers linear structure growth (dashed line), the shot noise starts dominating the signal at $\ell \approx 450$. For the non-linear power spectrum instead the differential signal-to-noise rises until $\ell \approx 1000$ (solid line) due to the enhancement of small scale structure by non-linear clustering. This shows the importance of the inclusion of high multipoles into the analysis. More specifically we see that the non-linear effects become important already at a relatively low $\ell \lesssim 200$.

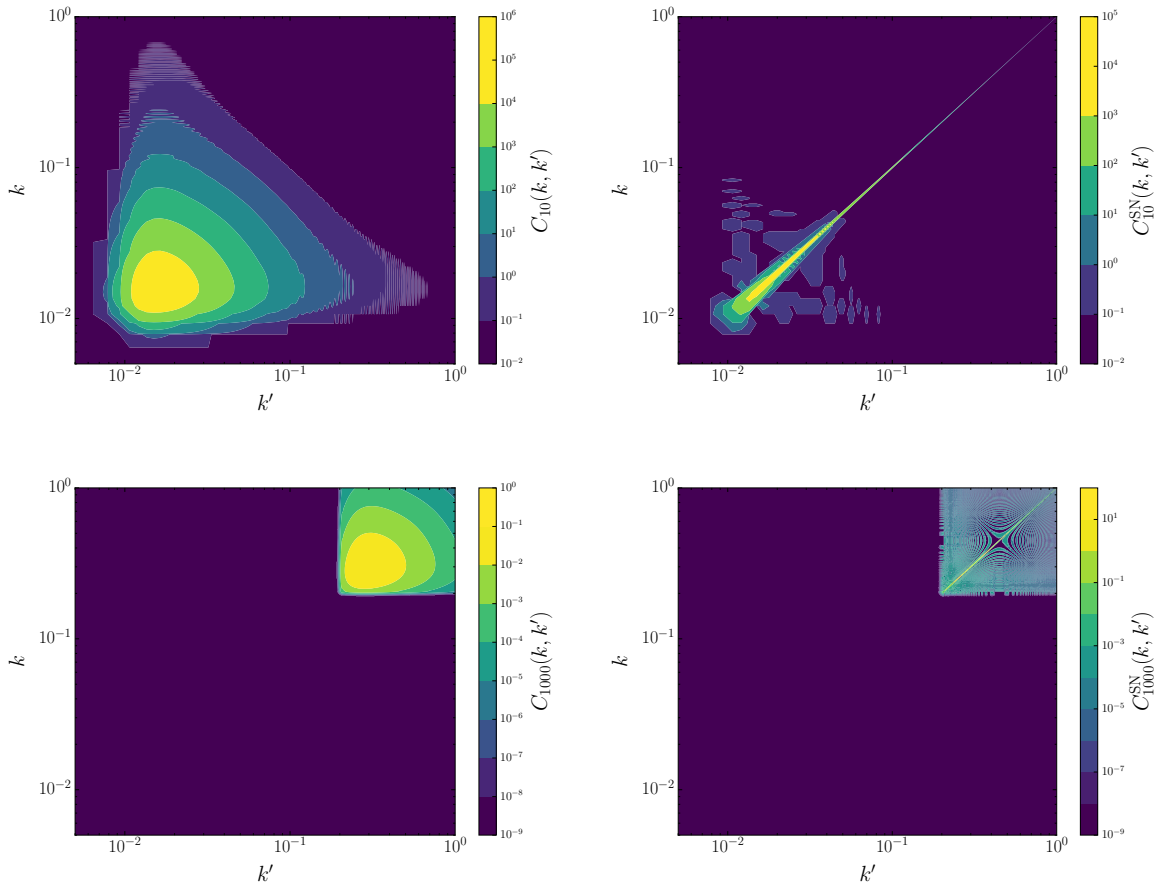

Figure 5.4.: Signal (*left*, labelled C_ℓ) and noise (*right*, labelled $C_\ell^{\rm SN}$, where the subscript stands for shot noise) parts of the covariance matrix (Eqs. 5.11 and 5.17, respectively) for the minimum and maximum ℓ-mode considered in the analysis, $\ell = 10$ (*upper panels*) and $\ell = 1000$ (*bottom panels*), respectively. Note the different ranges of the colour bars, in logarithmic scale. See also Fig. 5.5 for a comparison between the diagonal elements of the matrices, highlighting how different multipoles have contributions with different orders of magnitude and how the signal and noise part become dominant for low and high ℓ values, respectively.

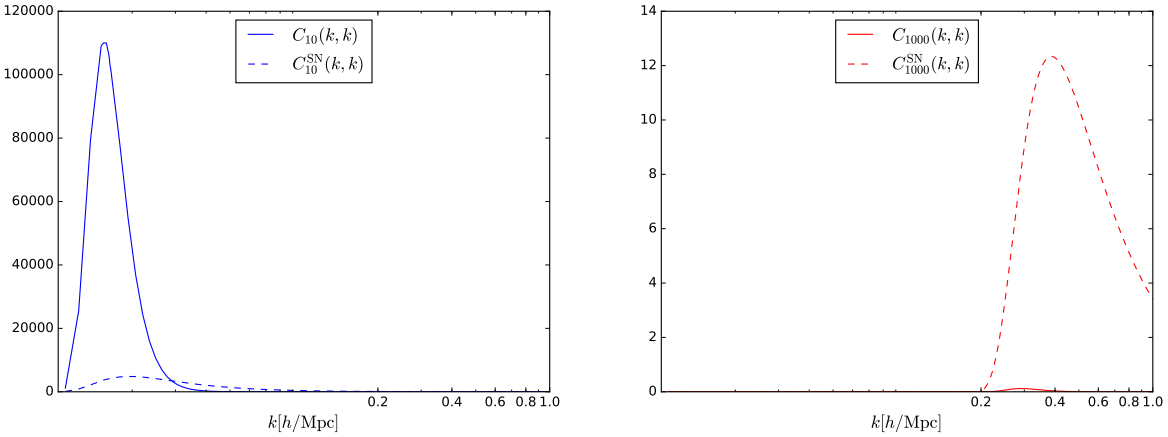

Figure 5.5.: Comparison between the diagonal elements of the signal (*solid line*, as given by eq. 5.11 and labelled C_ℓ) and noise (*dashed line*, as given by eq. 5.17 and labelled $C_\ell^{\rm SN}$, where the subscript stands for shot noise) contributions to the covariance matrices of the shear modes, for the minimum and maximum angular multipole considered in this analysis, i.e. $\ell = 10$ (*blue*) and $\ell = 1000$ (*red*), respectively. Note the different orders of magnitude for the different multipoles, and how the signal prevails on the noise for low multipoles, while the noise dominates for higher ℓ values. Note also the log-scale on the x-axis, to help identify the different k-regions where most of the contributions come from, for different multipoles.

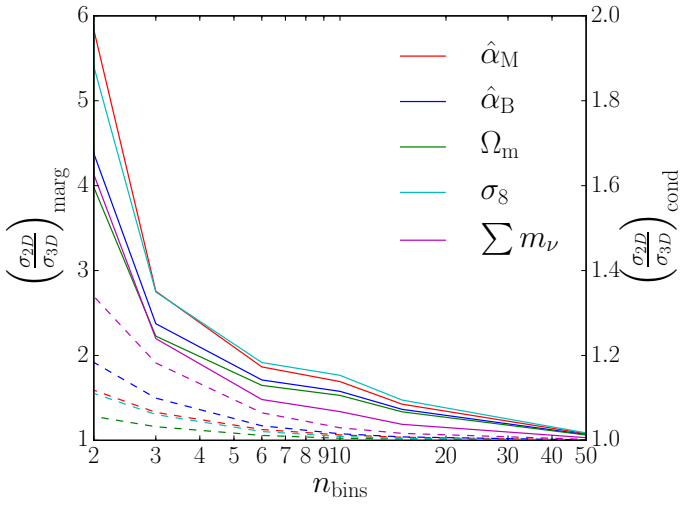

Figure 5.6.: Errors on different parameters from a tomographic analysis relative to a 3D analysis as a function of the number of tomographic bins, $n_{\rm bins}$ (in log-scale on the x-axis). The solid lines show the ratio of the marginal errors belonging to the left y-axis, while the dashed lines show the conditional errors belonging to the right y-axis.

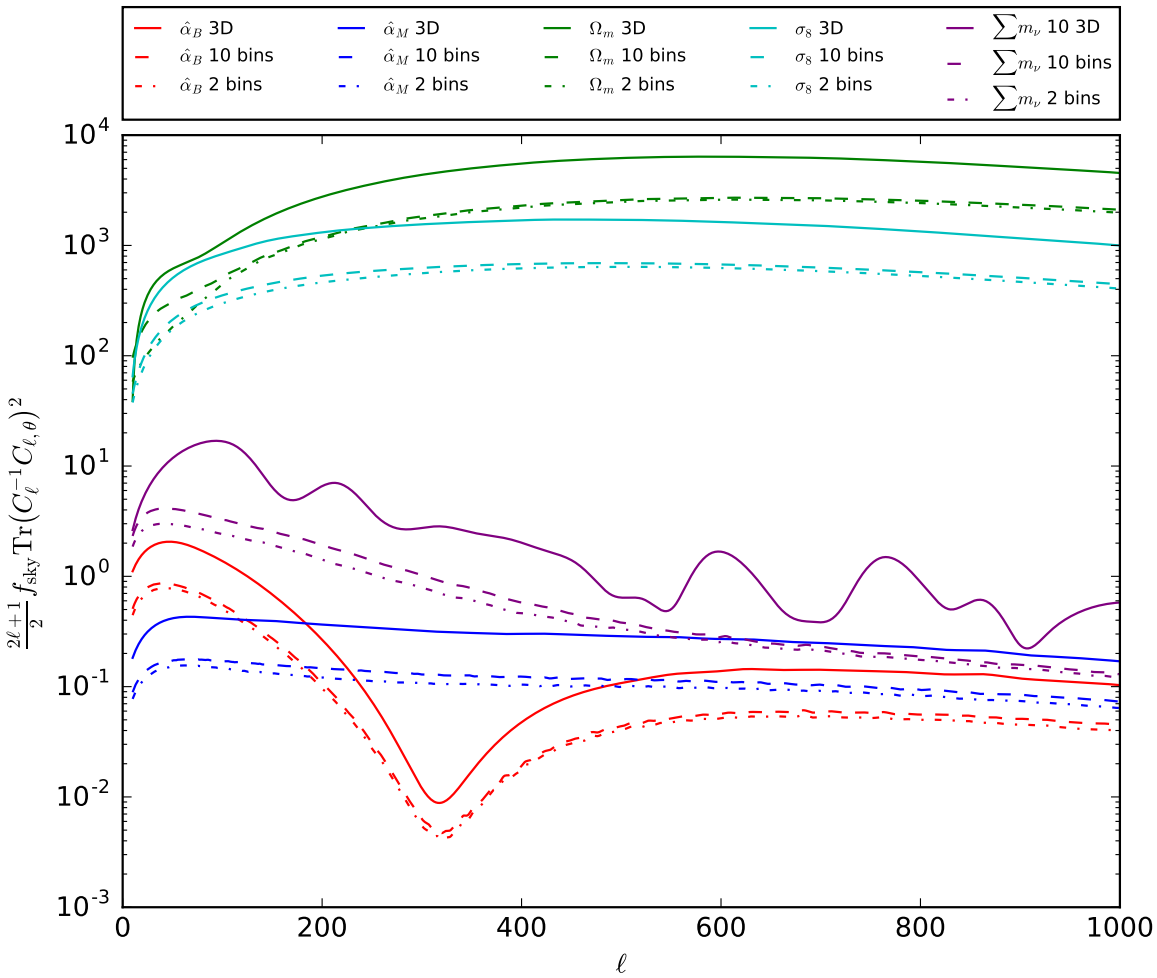

Figure 5.7.: Plots of the "sensitivity" curves to some selected cosmological parameters (they same chosen for Fig.5.6), where the sensitivity is defined by $\frac{2\ell+1}{2} f_{\rm sky} \mathrm{Tr}\left(C_\ell^{-1} C_{\ell,\theta}\right)^2$, i.e. the value of the summand in the formula 5.23 for the Fisher matrix, for different parameters θ (the comma denotes derivative with respect to θ). We notice that while increasing the number of bins (we plot in different *dashed* lines the curves for 2 and 10 bins) also increases the sensitivity, even considering 10 bins the curve does not reach the one achieved by 3D cosmic shear (plotted in *solid* line).

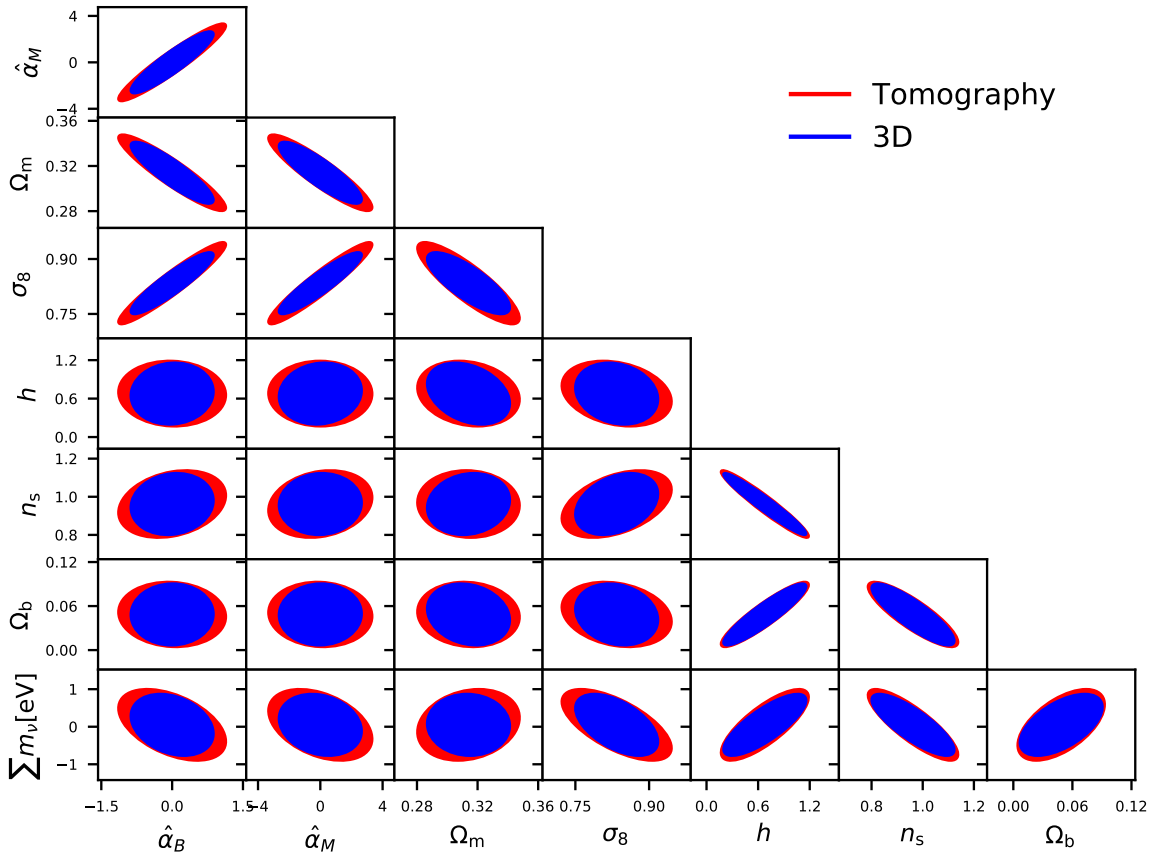

Figure 5.8.: 1-σ forecast contours for a Euclid-like survey, showing a comparison between a fully 3D (*blue*) and a tomographic cosmic shear analysis (*red*). The fiducial values can be found in Tab. 5.2, while $\hat{\alpha}_K = 0.01$ and the survey specifications are given in Tab. 5.1. We used the linear matter power spectrum for both.

In Fig. 5.7 the contribution of individual angular modes to the full Fisher information is illustrated: we plot the value of the summand in the formula Eq. 5.23 for the Fisher matrix, for different parameters (the same considered in Fig. 5.6), i.e. $\hat{\alpha}_B, \hat{\alpha}_M, \Omega_m, \sigma_8, \sum m_\nu$), as a function of the multipole order ℓ, which is identical to the Fisher matrix if only a single angular mode is considered. We plot these curves for the 3D cosmic shear case in solid lines and for tomography with 2 and 10 bins in different styles of dashed lines. We can interpret these curves as the "sensitivity" curves to the parameters considered. While we notice an increase of the sensitivity when increasing the number of bins, it still remains well below the 3D cosmic shear level.

5.3.2. Cosmological constraints on Horndeski functions

In our forecasts we fix α_T very close to zero and do not consider it as a parameter in our Fisher matrix analysis, reflecting the recent very strong constraints on the gravitational waves speed set by the detection of the binary neutron star merger GW170817 and the gamma ray burst GRB170817A (Abbott et al., 2017a,b; Baker et al., 2017; Creminelli & Vernizzi, 2017; Ezquiaga & Zumalacárregui, 2017; Sakstein & Jain, 2017). Furthermore, the kineticity α_K is largely unconstrained by cosmological observables (Bellini et al., 2016; Alonso et al., 2016), therefore we fix the coefficient $\hat{\alpha}_K$ to its fiducial value. However, we study the impact of the choice of $\hat{\alpha}_K$ on the constraints on the other parameters by choosing two values which differ by three orders of magnitude.

As already seen in Fig. 5.1 a lot of signal comes from non-linear scales, in fact for Euclid one

expects that about two thirds of the total signal to noise, $\Sigma(< \ell_{\rm max})$ with $\ell_{\rm max} \approx 2000$, originate from non-linear scales. It is therefore evident that one has to include non-linear clustering in the analysis in order to get the necessary statistical power to constrain a high-dimensional parameter space.

As seen before, increasing the number of tomographic bins yields more signal. In the inference process, however, the sensitivity to the model parameters plays an important role. For linear model parameters one expects the sensitivity to be a rescaled version of the SNR curve. In Fig. 5.6 we show the marginal and conditional errors of a tomographic analysis relative to the 3D analysis for a few parameters. The marginal errors are more strongly affected, since the contributions from the conditional errors add up during the marginalization procedure. Furthermore, we see the same trend as for the SNR: the expected errors tend towards the errors of a 3D analysis for $n_{\rm bins} \gg 1$.

Fig. 5.8 shows a comparison between cosmological constraints obtained with 3D cosmic shear and tomography with specifications from Tab. 5.1. Constraints from a 3D analysis are tighter than those from tomography, due to the increased redshift information. Furthermore, the degeneracies are in all cases very similar for the two methods, which is expected since the two methods probe the same quantity. In particular we find the usual degeneracy in $\Omega_{\rm m}$ and σ_8, which is slightly reduced in the 3D case. Generally the biggest improvement can be seen for parameters carrying information about the background evolution and the growth of structures; in contrast, parameters such as the spectral index $n_{\rm s}$ are not that much influenced.

In Fig. 5.9, instead, we study the impact of the choice of $\hat{\alpha}_K$ and compare constraints obtained only with 3D cosmic shear using a linear matter power spectrum, but with two different choices of fixed $\hat{\alpha}_K$, namely $\hat{\alpha}_K = 0.01$ and $\hat{\alpha}_K = 10$. We find, in agreement with Alonso et al. (2016), that the choice of $\hat{\alpha}_K$ does not affect the large-scale structure observables significantly. Since the largest effect on structure formation of $\hat{\alpha}_K$ comes from very large scales beyond those considered in this work, we do not expect any significant dependence of the Fisher matrix on this parameter.

Finally we investigate the impact of non-linear clustering in Fig. 5.10 as outlined before. Constraints are, as expected, tighter with the addition of the non-linear corrections. In particular we find a significant gain in $\Omega_{\rm m}$, σ_8 and $\sum m_\nu$[eV]. Other parameters such as the spectral index $n_{\rm s}$ and the Hubble constant h are not that much affected, since the main characteristics are already captured in the linear power spectrum. Furthermore, we find a gain in sensitivity in $\hat{\alpha}_B$ and $\hat{\alpha}_M$. This reduction of the error for the modified gravity parameters is mainly due to the marginalisation process and the degeneracies with the other parameters such as $\Omega_{\rm m}$, which are better constrained now. In fact, it should be noted that the conditional constraints on $\hat{\alpha}_{\rm M}$ and $\hat{\alpha}_{\rm B}$ become slightly worse then in the linear case, which has a subtle reason: the screening scale is chosen such that modified gravity effects are suppressed as soon as non-linear effects set in, on the other hand however, there is loss of power on intermediate scales. This effectively yields a loss in sensitivity, since the full effect of modified gravity is only present up to intermediate scales, whereas the signal gain on small scales does not contribute to the Fisher matrix. Finally, as a consequence of the screening mechanism, the orientation of the ellipses, for example in the case $\Omega_m - \hat{\alpha}_M$ and $\sigma_8 - \hat{\alpha}_M$, can change; this is due to the increase in signal at high ℓ values and the change of the sensitivity to cosmological parameters, especially because the sensitivity to the modified gravity parameters on those scales vanishes by construction.

In order to exploit the full potential of non-linear clustering one would need to have a reliable model for non-linear structure formation in a modified gravity setting. The way it is presented here effectively assumes that modifications of the gravitational field equation only play a role at linear order, while higher orders are treated in the usual framework of perturbation theory in a ΛCDM cosmology.

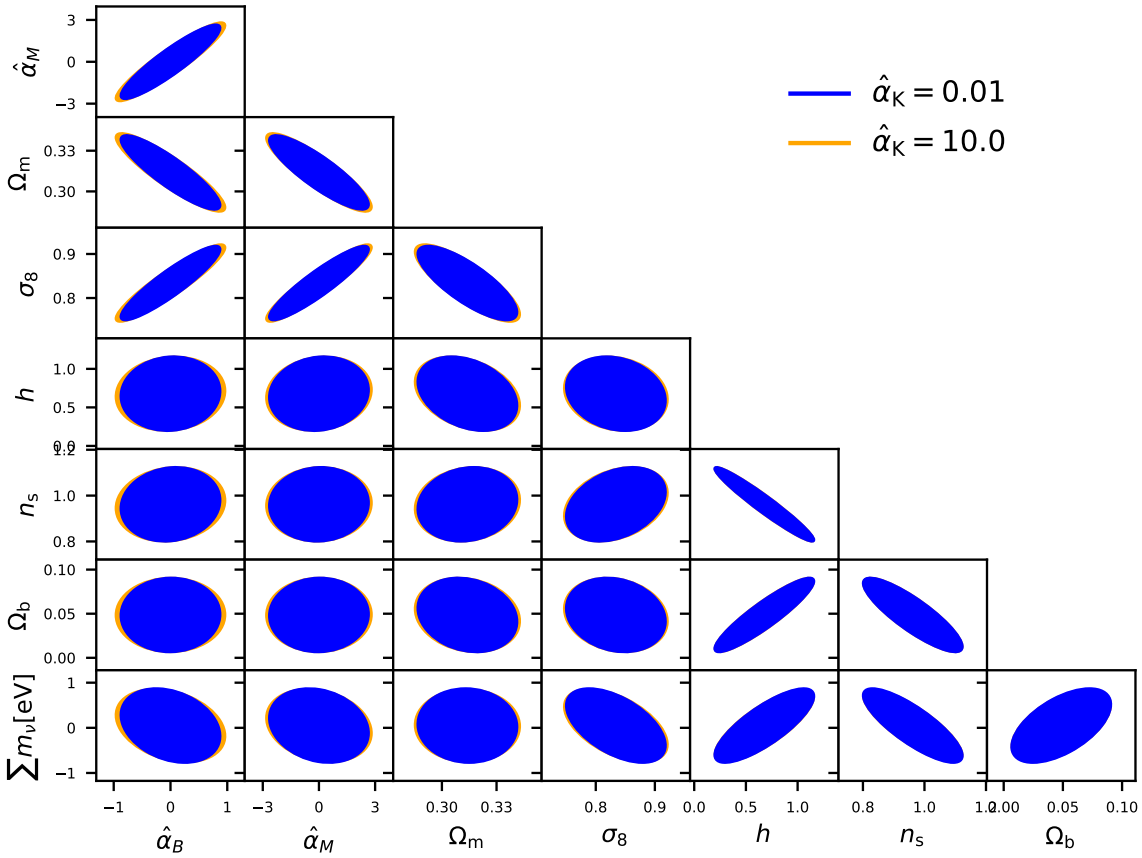

Figure 5.9.: Impact of the choice of $\hat{\alpha}_K$ for a 3D cosmic shear analysis with fiducial values from Tab. 5.2 and survey specifications given in Tab. 5.1. We used a linear power spectrum for the analysis and show the difference in the 1-σ contours when fixing $\hat{\alpha}_K = 0.01$ (*blue*) or $\hat{\alpha}_K = 10$ (*orange*).

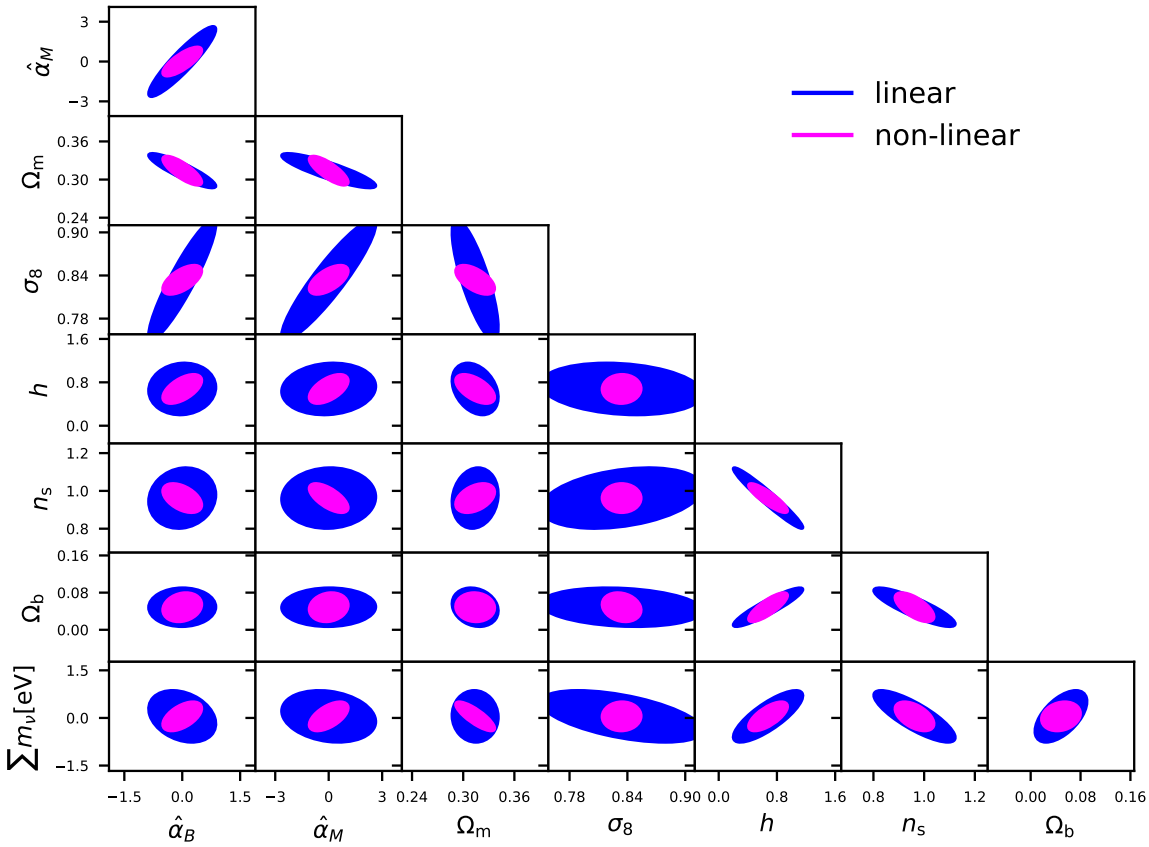

Figure 5.10.: Impact of non-linear clustering with fiducial values from Tab. 5.2 and survey specifications are given in Tab. 5.1. We show the constraints obtained with the linear power spectrum in *blue* and with the non-linear one in *magenta*.

5.4. Discussion and conclusions

In this Chapter we investigated the performance of a 3D analysis of cosmic shear measurements as a probe of Horndeski theories of modified gravity. We set constraints by means of a Fisher matrix analysis on a set of parameters that completely describe the evolution of linear perturbations in Horndeski gravity, using the specifications of a future Euclid-like experiment. We placed simultaneously our constraints on both the modified gravity parameters and on a set of standard cosmological parameters, including the sum of neutrino masses. Analogous forecasts for a tomographic analysis with six bins were produced given the same specifications of the cosmic shear experiment, with the aim of comparing the two methods. Our analysis was restricted to angular modes $\ell \leq 1000$ and $k \leq 1\,h/\mathrm{Mpc}$, to avoid the deeply non-linear regime of structure growth. We summarize our results as follows.

The signal-to-noise ratio of both a 3D analysis and a tomographic one is very similar, since it is mainly driven by the amplitude of the lensing signal and a tomographic method effectively agrees with a decomposition into spherical harmonics and radial Bessel functions if the bin width gets as small as the width of the photometric redshift errors.

3D cosmic shear provides tighter constraints than 10 bins tomography. Even with our conservative cut in angular and radial scales and using a linear matter power spectrum for the calculation of the covariance of the shear modes, 3D weak lensing performs better than tomography for all cosmological parameters, with both methods showing very similar degeneracies. For the parameters of the Bellini & Sawicki (2014) parametrization describing Horndeski theories, the gain is of the order of roughly 20 % in the errors.

We investigated the impact of the fiducial value chosen for the kineticity and found that the constraints are largely unaffected by the choice of $\hat{\alpha}_K$. In particular we used $\hat{\alpha}_K = 0.01$ and $\hat{\alpha}_K = 10$.

To illustrate the importance of non-linear corrections, we showed the expected improvement in the size of the constraints obtained employing a non-linear matter power spectrum: the results obtained in this case serve as an illustrative example of the constraining power of non-linear scales. In order to obtain a complete and self-consistent picture, one would need a formalism to construct the non-linear corrections in a general modified gravity setting (see e.g. Lombriser, 2016; Fasiello & Vlah, 2017). Here we introduced an artificial screening scale, which pushes the deviations from General Relativity to zero below its value. This is however not a fully exhaustive ansatz and many more investigations in this direction are required. The gain in signal if non-linear clustering is considered clearly shows the importance and calls for the development of analytic or semi-analytic prescriptions for the treatment of non-linear scales in ΛCDM and modified gravity. These will play a crucial role in allowing cosmic shear measurements to set strong constraints on parameters describing deviations from General Relativity. Due to the screening the constraints on modified gravity parameters are only improved because of the marginalization over the remaining parameters.

Compared to the analysis of Pratten et al. (2016), who considered 3D weak lensing only for a restricted number of modified gravity theories, our study extends the scope to the full Horndeski class and we do not fix all the parameters describing the background to their ΛCDM values.

This large parameter space and the fact that we only considered weak gravitational lensing as our observable makes our constraints less tight than the ones presented in Alonso et al. (2016), given also that our range in scales is less extended. Additionally, while Alonso et al. (2016) considered only a tomographic analysis, we present a 3D analysis along with a tomographic one, showing the increase in sensitivity of the former.

In our analysis we did not consider spurious contributions to the pure lensing signal coming from systematics such as the intrinsic alignments of source galaxies (Joachimi & Bridle, 2010; Mandelbaum, 2017). These are expected to dominate the error budget for future cosmic shear surveys and need therefore to be carefully accounted for. These contributions are also expected to influence mostly

the lensing signal on small, non-linear scales. The scales we considered were also chosen with the purpose of avoiding the regime of domination of these effects, which we considered neither in our fully 3D approach nor in the tomographic one, so that the comparison could remain fair. However, it has been shown that this kind of systematics can be carefully accounted for in 3D analyses (Merkel & Schäfer, 2013) and we plan to investigate their impact in future work, together with cross-correlations with other probes which we envisage as one of the most powerful tools to test gravity on cosmological scales.

6 Chapter 6

Testing dark energy and modified gravity with current and future cross-correlations of weak gravitational lensing

In this Chapter we investigate the possibility of obtaining constraints on Horndeski gravity with cross-correlations of cosmic shear with other probes, either using currently available datasets or making predictions for future Stage IV surveys.

First, we present cosmological constraints on Horndeski theories of gravity obtained from the joint analysis of tomographic cosmic shear power spectra from $\sim 450\text{deg}^2$ of data from the Kilo Degree Survey (KiDS), galaxy-matter cross-correlation power spectra of foreground galaxies from the Galaxies And Mass Assembly (GAMA) survey and background KiDS sources, and the angular clustering power spectra of the same GAMA galaxies.

As expected, the data from the KiDS and GAMA surveys available to us while performing this analysis did not allow for very tight constraints on the Horndeski parameters. However, our analysis represents the first example in the literature of a complete inference pipeline producing weak lensing constraints on Horndeski gravity: the methodology and numerical implementation that we have developed could be used in the future to either repeat the same analysis with larger data releases from the KiDS survey or, in the next decade, to fully exploit the statistical richness of even larger datasets coming from Stage IV surveys such as Euclid (Laureijs et al., 2011) or LSST (LSST Science Collaboration et al., 2009).

These next-generation surveys are the focus of attention in the second part of the analysis developed in this Chapter. We forecast constraints on a large parameter space that includes standard cosmological parameters as well as Horndeski parameters, considering the combination of cosmic shear, galaxy clustering, Cosmic Microwave Background anisotropies and lensing as will be available to us in the next decade from surveys such as Euclid and CMB-S4 (Abazajian et al., 2016). We produce our forecasts with both a Fisher matrix approach and a full MCMC sampling of the posterior distribution. We find that datasets such as the ones provided by those surveys will effectively be able to constrain Horndeski theories with relative errors on their gravitational parameters below 10%.

The structure of this Chapter is as follows: we begin in Sec. 6.1 with an introduction to the possible cross-correlations that can be obtained from optical surveys between estimates of the ellipticity and of the number density contrast. This allows us to introduce the probes mentioned subsequently in the Chapter, such as galaxy-galaxy lensing and galaxy clustering. We proceed in Sec. 6.2 to describe our analysis of the KiDS and GAMA surveys. In Sec. 6.3 we present our forecasts on cross-correlations of Stage IV surveys. Finally, we draw our conclusions for this Chapter in Sec. 6.4.

6.1. Cross-correlations shear-density field

In the previous Sections we have considered tomography and 3D cosmic shear as the two main techniques to analyse a cosmic shear survey: despite being based on different formalisms, both methods share in common the need of galaxy ellipticity measurements in order to build estimates of the shear modes. Such a gravitational lensing analysis requires therefore information on the shapes of the observed galaxies; an estimate of their ellipticity quantifies the shear distortion and subsequently allows for the study of the correlation of these distorted shapes on the sky. Additionally, redshift information can be incorporated in the analysis (in different ways as we have seen) to increase the amount of information available.

Assuming the gravitational shear to be weak (see e.g. Bartelmann & Schneider, 2001), as in the case of cosmic shear, the ellipticity can be inferred from the galaxy shapes in a given region of space as

$$\epsilon^{(i)}(\boldsymbol{\theta}) = \gamma_{\mathrm{G}}^{(i)}(\boldsymbol{\theta}) + \gamma_{\mathrm{I}}^{(i)}(\boldsymbol{\theta}) + \epsilon_{\mathrm{rnd}}^{(i)}(\boldsymbol{\theta}) \,, \tag{6.1}$$

where the superscript in parentheses assigns a photo-z bin i, while the vector $\boldsymbol{\theta}$ denotes angular coordinates (θ, ϕ). The observed ellipticity ϵ has contributions from the gravitational shear γ_{G} and an intrinsic shear γ_{I}, which is caused by the alignment of a galaxy in its surrounding gravitational field. Moreover, ϵ is assumed to have an uncorrelated component ϵ_{rnd}, which accounts for the purely random part of the intrinsic orientations and shapes of galaxies.

Here we would like to widen the context of our study and review how shape information can be combined with information on the positions of galaxies; the latter can be used to construct an estimate of the number density contrast

$$n^{(i)}(\boldsymbol{\theta}) = n_{\mathrm{m}}^{(i)}(\boldsymbol{\theta}) + n_{\mathrm{g}}^{(i)}(\boldsymbol{\theta}) + n_{\mathrm{rnd}}^{(i)}(\boldsymbol{\theta}) \,, \tag{6.2}$$

determined by the intrinsic number density contrast of galaxies n_{g} and the alteration of galaxy counts due to lensing magnification n_{m}. An uncorrelated shot noise contribution is added via n_{rnd}.

In contrast to $\epsilon^{(i)}(\boldsymbol{\theta})$ the number density contrast $n^{(i)}(\boldsymbol{\theta})$ cannot be estimated from individual galaxies. One can understand $n^{(i)}(\boldsymbol{\theta})$ as the ensemble average over a hypothetical, Poisson-distributed random field of which the observed galaxy distribution is one particular representation. The formal relation between the projected number density contrast as used in Eq. 6.2 and the three-dimensional galaxy number density fluctuations will be provided below, see Eq. 6.12.

As noted e.g. in Joachimi & Bridle (2010), Eqs. 6.1 and 6.2 are symmetric in that they both contain an intrinsic contribution and a term caused by gravitational lensing effects. The intrinsic contributions are both normally dominant over the lensing parts, when considered singularly. In fact, the intrinsic ellipticity of galaxies modifies their shape from circular even in absence of lensing, and is about two orders of magnitude larger than the gravitational shear. As a (very rough) first approximation one can assume that the intrinsic ellipticities of galaxies are randomly distributed on the sky, so that they are correlated neither with the intrinsic ellipticities nor with the shears of other galaxies. As a consequence, the correlation of ellipticities becomes dominated by the gravitational shear; in contrast, the largest term in Eq. 6.2 is due to the intrinsic number density contrast, which is also responsible for the major contribution in the correlation of the number density constrast.

Both ellipticity and number over-density vanish if averaged over sufficiently large scales. Thus, one considers to lowest order two-point statistics of these quantities. Since all real-space two-point measures are related to the power spectrum (see e.g. Kaiser, 1992), one can work in terms of power spectra insted of correlation functions, which is desirable in particular due to a simpler structure of the signal covariances in Fourier space. Denoting the Fourier transform by a tilde, the power spectrum $C_{\mathrm{ab}}^{(ij)}(\ell)$ between redshift bins i and j can then be defined by

$$\left\langle \tilde{x}_{\mathrm{a}}^{(i)}(\boldsymbol{\ell})\, \tilde{x}_{\mathrm{b}}^{(j)}(\boldsymbol{\ell}') \right\rangle = (2\pi)^2\, \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell} - \boldsymbol{\ell}')\, C_{\mathrm{ab}}^{(ij)}(\ell) \,, \tag{6.3}$$

measured correlation	2D PS	3D PS
shear	C_{GG}	$P_{\delta\delta}$
intrinsic-shear	C_{IG}	$P_{\delta\mathrm{I}}$
intrinsic	C_{II}	P_{II}
galaxy clustering	C_{gg}	P_{gg}
clustering-magnification	C_{gm}	$P_{\mathrm{g}\delta}$
magnification	C_{mm}	$P_{\delta\delta}$
clustering-shear	C_{gG}	$P_{\mathrm{g}\delta}$
clustering-intrinsic	C_{gI}	P_{gI}
magnification-shear	C_{mG}	$P_{\delta\delta}$
magnification-intrinsic	C_{mI}	$P_{\delta\mathrm{I}}$
galaxy ellipticity (observable)	$C_{\epsilon\epsilon}$	
galaxy number density (observable)	C_{nn}	
number density-ellipticity (observable)	$C_{n\epsilon}$	

Table 6.1.: Overview on the two-point correlations considered in this Section 6.1. Listed are the symbols used for the two-dimensional projected power spectra and the underlying three-dimensional power spectra.

where $\delta_{\mathrm{D}}^{(2)}$ is the two-dimensional Dirac delta-distribution, and where $\boldsymbol{\ell}$ denotes the angular frequency, i.e. the Fourier variable on the sky. The measures x_{a} and x_{b} can correspond to any of the set $\left\{\gamma_{\mathrm{G}}, \gamma_{\mathrm{I}}, n_{\mathrm{g}}, n_{\mathrm{m}}\right\}$. The random contributions in Eqs. 6.1 and 6.2 are not correlated with any of the other measures and only yield a contribution to the noise.

Following e.g. Joachimi & Bridle (2010), we can insert Eqs. 6.1 and 6.2 into 6.3, thus obtaining the complete set of tomographic two-point observables which are available from shape and number density information

$$C_{\epsilon\epsilon}^{(ij)}(\ell) = C_{\mathrm{GG}}^{(ij)}(\ell) + C_{\mathrm{IG}}^{(ij)}(\ell) + C_{\mathrm{IG}}^{(ji)}(\ell) + C_{\mathrm{II}}^{(ij)}(\ell) \tag{6.4}$$

$$C_{nn}^{(ij)}(\ell) = C_{\mathrm{gg}}^{(ij)}(\ell) + C_{\mathrm{gm}}^{(ij)}(\ell) + C_{\mathrm{gm}}^{(ji)}(\ell) + C_{\mathrm{mm}}^{(ij)}(\ell) \tag{6.5}$$

$$C_{n\epsilon}^{(ij)}(\ell) = C_{\mathrm{gG}}^{(ij)}(\ell) + C_{\mathrm{gI}}^{(ij)}(\ell) + C_{\mathrm{mG}}^{(ij)}(\ell) + C_{\mathrm{mI}}^{(ij)}(\ell)\,, \tag{6.6}$$

Signals originating from galaxy shape information are denoted by capital letters ('G' for gravitational shear, 'I' for intrinsic shear), while signals related to galaxy number densities by small letters ('g' for intrinsic number density fluctuations, 'm' for lensing magnification). Table 6.1 summarizes the nomenclature used for the correlations in Eqs. 6.4, 6.5, 6.6. Note that Eqs. 6.4 and 6.5 are symmetric with respect to their photo-z bin arguments. Hence, if N_{zbin} denotes the number of available photo-z bins, one has $N_{\mathrm{zbin}}(N_{\mathrm{zbin}}+1)/2$ observables for every considered angular frequency. In contrast, one can exploit N_{zbin}^2 ellipticity-number density cross-correlation power spectra (Eq. 6.6) for each ℓ.

The set of observables in Eq. 6.4 is the one that cosmic shear analyses are based on. The shear correlation signal (GG) is a clean probe of the underlying matter power spectrum and is thus powerful in constraining cosmological parameters. However, shape measurements incorporate further terms stemming from correlations of intrinsic ellipticities (II) and shear-intrinsic cross-correlations (IG, or equivalently GI) whose contribution can be substantial, but is to date poorly known. These terms exist because the shapes and orientations of galaxies are influenced via the tidal gravitational fields of the matter structures in their surrounding, which firstly induce correlations between neighbouring galaxies, and secondly cause correlations by determining the intrinsic shape of a foreground object and adding to the shear signal of a background galaxy (see Joachimi et al., 2015; Troxel & Ishak, 2015, for reviews on the topic). The matter structure around galaxies can modify their intrinsic shape and their orientation. Firstly, this can result in correlations between the intrinsic shapes of galaxies which are close both on the sky and in redshift (intrinsic ellipticity correlations, or II correlations).

Moreover, a dark matter halo can intrinsically align a physically close galaxy in the foreground and at the same time contribute to the lensing signal of a background object, which induces gravitational shear-intrinsic ellipticity correlation (GI contribution).

Intrinsic *galaxy clustering* (gg) adds a strong signal to the correlations of galaxy number densities (Eq. 6.5), but its use to obtain cosmological parameter estimates is faced by poor knowledge of galaxy bias i.e. the relation between the distribution of galaxies and the underlying dark matter density field. Gravitational lensing modifies the flux of objects and thus reduces or increases number counts of galaxies above a certain limiting magnitude. This produces magnification correlations (mm) and intrinsic number density-magnification cross-correlations (gm). The gm correlations occur when a foreground mass overdensity (underdensity) contains an overdensity (underdensity) of galaxies and (de)-magnifies background objects along the same line of sight causing an apparent over- or underdensity of galaxies at higher redshift.

Cross-correlations between galaxy number densities and ellipticities (Eq. 6.6) contain contributions from cross terms between intrinsic clustering and shear (gG), intrinsic clustering and intrinsic shear (gI), magnification and shear (mG), and magnification and intrinsic shear (mI). For instance, one expects to find gI and gG signals when a mass structure leads to an overdensity in the local galaxy distribution and influences the intrinsic shape of galaxies at the same redshift or contributes to the shear of background objects. The latter case, where an overdensity in the foreground galaxy distribution contributes to the shear of background sources, is typically referred to as *galaxy-galaxy lensing*. Because a foreground overdensity can in addition enhance galaxy counts due to lensing magnification, the mG and mI signals will also be non-vanishing. In our analysis in next Section we will ignore the magnification contribution, since it was not considered in the first place in the original ΛCDM analysis of the Kilo-Degree Survey (van Uitert et al., 2018), of which our work represents an extension to modified gravity.

6.1.1. Connection to the matter power spectrum

All non-random terms in Es. 6.1 and 6.2, given for a photometric redshift bin i, can be related to a source term $\mathcal{S}$, which is a function of spatial coordinates, i.e.

$$x_{\rm a}^{(i)}(\boldsymbol{\theta}) = \int_0^{\chi_{\rm hor}} {\rm d}\chi \; w^{(i)}(\chi) \; \mathcal{S}_{\rm a}\left(\chi\boldsymbol{\theta},\chi\right) \; , \tag{6.7}$$

where we defined a weight function w that depends on the photo-z bin i (for a similar approach see Hu & Jain, 2004). Here and throghout the Section we consider a flat Universe (expressions for the non-flat case can be obtained generalising χ to a non-flat Universe $f_K(\chi)$). If Eq. 6.7 holds for two quantities $x_{\rm a}^{(i)}$ and $x_{\rm b}^{(j)}$, their projected power spectrum is given by the line-of-sight integral of the three-dimensional source power spectrum $P_{\mathcal{S}_a\mathcal{S}_b}$ via Limber's equation in Fourier space (Kaiser, 1992),

$$C_{\rm ab}^{(ij)}(\ell) = \int_0^{\chi_{\rm hor}} {\rm d}\chi \; \frac{w^{(i)}(\chi)\; w^{(j)}(\chi)}{\chi^2} \; P_{\mathcal{S}_a\mathcal{S}_b}\left(\frac{\ell}{\chi},\chi\right) \; . \tag{6.8}$$

By identifying weights and source terms for gravitational and intrinsic shear, as well as intrinsic clustering and magnification, we can derive Limber equations for all power spectra entering Eqs. 6.4-6.6.

To compute the equivalent of Eq. 6.7 for the cosmic shear case, we first note that in Fourier space the shear and the convergence are related by the simple equation $\tilde{\kappa}_{\rm G}(\boldsymbol{\ell}) = \tilde{\gamma}_{\rm G}(\boldsymbol{\ell})\, {\rm e}^{-2{\rm i}\varphi_\ell}$, where φ_ℓ is the polar angle of $\boldsymbol{\ell}$. As a consequence, we can use for simplicity the convergence as the cosmic shear observable in our calculations. The convergence is related to the three-dimensional matter density contrast δ via

$$\kappa_{\rm G}^{(i)}(\boldsymbol{\theta}) = \int_0^{\chi_{\rm hor}} {\rm d}\chi \; q^{(i)}(\chi) \; \delta\left(\chi\boldsymbol{\theta},\chi\right) \; , \tag{6.9}$$

where the weight is given by

$$q^{(i)}(\chi) = \frac{3H_0^2\Omega_\mathrm{m}}{2\,c^2}\frac{\chi}{a(\chi)}\int_\chi^{\chi_\mathrm{hor}} \mathrm{d}\chi'\; p^{(i)}(\chi')\,\frac{\chi'-\chi}{\chi'}\;, \tag{6.10}$$

see Bartelmann & Schneider (2001) for details. Here a denotes the scale factor and $p^{(i)}(\chi)$ the comoving distance probability distribution of those galaxies in bin i for which shape information is available.

Analogously to the lensing case, one can define a convergence of the intrinsic shear field $\kappa_\mathrm{I}^{(i)}(\boldsymbol{\theta})$, which is directly related to the intrinsic shear via $\tilde{\kappa}_\mathrm{I}(\boldsymbol{\ell}) = \tilde{\gamma}_\mathrm{I}(\boldsymbol{\ell})\,\mathrm{e}^{-2\mathrm{i}\varphi_\ell}$. This intrinsic convergence is a projection of the three-dimensional intrinsic shear field $\bar{\kappa}_\mathrm{I}$, which can be written as

$$\kappa_\mathrm{I}^{(i)}(\boldsymbol{\theta}) = \int_0^{\chi_\mathrm{hor}} \mathrm{d}\chi\; p^{(i)}(\chi)\;\bar{\kappa}_\mathrm{I}\left(\chi\boldsymbol{\theta},\chi\right)\;, \tag{6.11}$$

and analogously in terms of intrinsic shear.

Likewise, angular galaxy number density fluctuations $n_\mathrm{g}^{(i)}(\boldsymbol{\theta})$ are given by the line-of-sight projection of three-dimensional number density fluctuations δ_g as

$$n_\mathrm{g}^{(i)}(\boldsymbol{\theta}) = \int_0^{\chi_\mathrm{hor}} \mathrm{d}\chi\; p^{(i)}(\chi)\;\delta_\mathrm{g}\left(\chi\boldsymbol{\theta},\chi\right)\;. \tag{6.12}$$

The distribution of galaxies is expected to follow the distribution of dark matter, so that the galaxy clustering power spectra should be related to $P_{\delta\delta}$, the (cold dark) matter power spectrum. However, to date it is unknown how much the galaxy clustering deviates from dark matter clustering, in particular on small scales. This is usually expressed in terms of the galaxy bias b_g, which is a function of both angular scale k and redshift or line-of-sight distance χ. Hence, one can write

$$\begin{aligned} P_\mathrm{gg}(k,\chi) &= b_\mathrm{g}^2(k,\chi)\;P_{\delta\delta}(k,\chi) \\ P_{\mathrm{g}\delta}(k,\chi) &= b_\mathrm{g}(k,\chi)\;r_\mathrm{g}(k,\chi)\;P_{\delta\delta}(k,\chi)\;, \end{aligned} \tag{6.13}$$

where to describe the cross-correlation between matter and galaxy clustering, we introduced a correlation coefficient r_g in the second equality.

The intrinsic alignment power spectra depend on the intricacies of galaxy formation and evolution within their dark matter environment. Again, precise models of the intrinsic alignment have to rely on baryonic physics and are currently not available. For symmetry reasons we parameterize our lack of knowledge about the intrinsic alignment power spectra similarly to the galaxy bias as

$$\begin{aligned} P_\mathrm{II}(k,\chi) &= b_\mathrm{I}^2(k,\chi)\;P_{\delta\delta}(k,\chi) \\ P_{\delta\mathrm{I}}(k,\chi) &= b_\mathrm{I}(k,\chi)\;r_\mathrm{I}(k,\chi)\;P_{\delta\delta}(k,\chi)\;, \end{aligned} \tag{6.14}$$

with the intrinsic alignment bias b_I and correlation coefficient r_I. Although the power spectrum P_gI could in principle contain a third, independent correlation coefficient, we assume that it is sufficient to write

$$P_\mathrm{gI}(k,\chi) = b_\mathrm{I}(k,\chi)\;r_\mathrm{I}(k,\chi)\;b_\mathrm{g}(k,\chi)\;r_\mathrm{g}(k,\chi)\;P_{\delta\delta}(k,\chi)\;, \tag{6.15}$$

i.e. we hypothesize that correlations between intrinsic number density fluctuations and intrinsic alignments can entirely be traced back to the effects of the intrinsic alignment bias and the galaxy bias.

We insert the parameterizations Eqs. 6.13, 6.14 and 6.15 into the set of Limber equations and can this way relate all power spectra entering Eqs. (6.4-6.6) to the three-dimensional matter power spectrum:

$$C^{(ij)}_{\rm GG}(\ell) = \int_0^{\chi_{\rm hor}} {\rm d}\chi \, \frac{q^{(i)}(\chi)\, q^{(j)}(\chi)}{\chi^2} \, P_{\delta\delta}\left(\frac{\ell}{\chi},\chi\right) \tag{6.16}$$

$$C^{(ij)}_{\rm IG}(\ell) = \int_0^{\chi_{\rm hor}} {\rm d}\chi \, \frac{p^{(i)}(\chi)\, q^{(j)}(\chi)}{\chi^2} \times b_{\rm I}\left(\frac{\ell}{\chi},\chi\right) \, r_{\rm I}\left(\frac{\ell}{\chi},\chi\right) \, P_{\delta\delta}\left(\frac{\ell}{\chi},\chi\right) \tag{6.17}$$

$$C^{(ij)}_{\rm II}(\ell) = \int_0^{\chi_{\rm hor}} {\rm d}\chi \, \frac{p^{(i)}(\chi)\, p^{(j)}(\chi)}{\chi^2} \times b^2_{\rm I}\left(\frac{\ell}{f_{\rm K}(\chi)},\chi\right) \, P_{\delta\delta}\left(\frac{\ell}{\chi},\chi\right) \tag{6.18}$$

$$C^{(ij)}_{\rm gg}(\ell) = \int_0^{\chi_{\rm hor}} {\rm d}\chi \, \frac{p^{(i)}(\chi)\, p^{(j)}(\chi)}{\chi^2} \times b^2_{\rm g}\left(\frac{\ell}{\chi},\chi\right) \, P_{\delta\delta}\left(\frac{\ell}{\chi},\chi\right) \tag{6.19}$$

$$C^{(ij)}_{\rm gG}(\ell) = \int_0^{\chi_{\rm hor}} {\rm d}\chi \, \frac{p^{(i)}(\chi)\, q^{(j)}(\chi)}{\chi^2} \times b_{\rm g}\left(\frac{\ell}{\chi},\chi\right) \, r_{\rm g}\left(\frac{\ell}{\chi},\chi\right) \, P_{\delta\delta}\left(\frac{\ell}{\chi},\chi\right) \tag{6.20}$$

$$C^{(ij)}_{\rm gI}(\ell) = \int_0^{\chi_{\rm hor}} {\rm d}\chi \, \frac{p^{(i)}(\chi)\, p^{(j)}(\chi)}{\chi^2} \, b_{\rm g}\left(\frac{\ell}{\chi},\chi\right) \, r_{\rm g}\left(\frac{\ell}{\chi},\chi\right) \times \, b_{\rm I}\left(\frac{\ell}{\chi},\chi\right) \, r_{\rm I}\left(\frac{\ell}{\chi},\chi\right) \, P_{\delta\delta}\left(\frac{\ell}{\chi},\chi\right) . \tag{6.21}$$

6.2. KiDS+GAMA: constraints on Horndeski gravity from cross-correlations of current surveys

In this Section we present cosmological constraints on Horndeski theories of gravity obtained from the joint analysis of tomographic cosmic shear power spectra from $\sim 450{\rm deg}^2$ of data from the KiDS survey, galaxy-matter cross-correlation power spectra of foreground galaxies from the GAMA survey and background KiDS sources, and the angular clustering power spectra of the same GAMA galaxies.

This work extends the analysis of van Uitert et al. (2018), carried out in a ΛCDM scenario, to the parameters that fully describe the evolution of linear perturbations in Horndeski gravity, which we introduced in Sec. 2.6.2; however, our results are obtained with a numerical implementation that is completely independent from the one used in van Uitert et al. (2018). To validate our implementation, before producing constraints on Horndeski gravity we demonstrate that our ΛCDM results show excellent agreement with those of van Uitert et al. (2018). Our likelihood module, developed for the sampler MONTE PYTHON (Audren et al., 2013), will be publicly released, with the goal of providing the cosmological community with a benchmarked weak lensing inference pipeline that is able to produce cosmological constraints both within the concordance model and in an extended modified gravity scenario.

Our analysis also represents an important cross-check of the robustness of the cosmological constraints derived by van Uitert et al. (2018), which alleviated the existing tension between the fiducial cosmological analysis of the KiDS survey (Hildebrandt et al., 2017) and the Planck results (Planck Collaboration et al., 2016a).

The cosmological inference module developed in our analysis is completely independent from the pipeline used in van Uitert et al. (2018), having been written in a different programming language and for a different likelihood sampler; only the data input in the form of power spectra catalogues is shared between the two analyses. With our new likelihood module the user has the freedom to choose all possible combinations of the three probes to analyse, allowing them to either consider the different probes singularly, in pairs or all together. The user can also choose whether to perform the analysis in a ΛCDM scenario or in a Horndeski gravity context. For all the possible probe combinations we find excellent agreement with van Uitert et al. (2018) and we present constraints on the Horndeski functions in each of these configurations.

We start in Sec. 6.2.1 describing the motivation for a joint analysis that considers cosmic shear, galaxy-galaxy lensing and angular clustering; in Sec.6.2.2 we briefly describe the theoretical back-

ground for the three probes in terms of their projected power spectra; in Sec.6.2.3 we provide details concerning our methodology, such as the calculation of the covariance, the intrinsic alignment and baryonic feedback models employed in our analysis; in Sec.6.2.4 we show our results, first in terms of comparison with the constraints obtained by van Uitert et al. (2018), and subsequently focusing on the constraints on Horndeski parameters. The conclusions from this analysis will be drawn in Sec. 6.4 and compared with those reached in the Stage IV-analysis of Sec. 6.3.

6.2.1. Introduction: multi-probe analyses

The statistic describing at lowest order the galaxy spatial distribution is the two-point correlation function, which in the past provided early evidence for the ΛCDM model (Baugh, 1996; Eisenstein & Zaldarriaga, 2001; Saunders et al., 2000; Huterer et al., 2001; Hamilton & Tegmark, 2002; Cole et al., 2005; Tegmark et al., 2006). Today, studying the spatial distribution of galaxies and its evolution in time is crucial to analyse possible extensions to the cosmological concordance model, in particular to shed light on the true nature of cosmic acceleration. However, the interpretation of galaxy clustering is complicated by galaxy bias, the relation between the galaxy spatial distribution and the theoretically predicted matter distribution.

In addition to galaxy clustering, weak gravitational lensing has become one of the principal probes of cosmology, providing a direct measurement of the total mass distribution including dark and baryonic matter. Distortions in the galaxy shapes of background galaxies are caused by fluctuations of the gravitational field of the large-scale structure in the foreground (Bartelmann & Schneider, 2001). Information on these distortions can be extracted by correlating the observed shapes of galaxies, i.e. studying the so-called cosmic shear field (for a review, see Kilbinger, 2015), which we introduced in the previous Chapters. Alternatively, one can correlate the positions of lens galaxies in the foreground (tracing the large-scale structure) with the shapes of source galaxies in the background, i.e. considering the galaxy-matter cross-correlation, often referred to as galaxy-galaxy lensing, introduced in Sec. 6.1.

The shape distortions produced by gravitational lensing, while cosmologically informative, are extremely difficult to measure, since the induced source galaxy ellipticities are at the percent level, and a number of systematic effects can obscure the signal. For this reasons, only recently have cosmic shear measurements begun to provide competitive constraints on cosmological parameters. Galaxy–galaxy lensing measurements have also matured to the point where their combination with galaxy clustering breaks degeneracies between the cosmological parameters and bias, thereby helping to constrain dark energy (see e.g. Kwan et al., 2017).

Most cosmic shear studies to date have been using the shear correlation functions (e.g. Heymans et al., 2013; Abbott et al., 2016; Hildebrandt et al., 2017) or its equivalent in Fourier space, the shear power spectrum (e.g. Brown et al., 2003; Heymans et al., 2005; Kitching et al., 2007; Lin et al., 2012; Kitching et al., 2014; Köhlinger et al., 2016; Abbott et al., 2016; Alsing et al., 2017; Köhlinger et al., 2017), to constrain cosmological parameters. An important result of the fiducial cosmic shear analyses of the Canada-France-Hawaii Lensing Survey (CFHTLenS; Heymans et al., 2013) and the Kilo Degree Survey (KiDS; Hildebrandt et al., 2017), two of the most constraining surveys to date, is that they prefer a cosmological model that is in mild tension with the best-fitting cosmological model from Planck Collaboration et al. (2016a). The first cosmological results from the Dark Energy Survey (DES; Abbott et al., 2016) are consistent with Planck, but their uncertainties are considerably larger. Also, the result from the Deep Lens Survey (DLS; Jee et al., 2016) agrees with Planck. This tension is currently being further investigated in order to understand its origin: if confirmed and not due to yet unexplored systematics, it could potentially lead to very interesting new cosmological scenarios (see e.g. Battye & Moss, 2014; MacCrann et al., 2014; Kitching et al., 2016; Joudaki et al., 2017).

Modern optical imaging surveys measure the positions and gravitational lensing-induced shears of

millions of galaxies. From these measurements, one can compute two fields on the sky: the spin-0 galaxy overdensity field, δ_g, and the spin-2 weak lensing shear field, γ. Two-point cross-correlations between these fields are powerful cosmological probes, as they are sensitive to both the geometry of the Universe and the growth of structure. Joint fits to multiple two-point correlations offer the possibility of breaking degeneracies between cosmological and nuisance parameters, as well as significantly improving cosmological constraints.

In the past a number of studies have combined individual large-scale structure probes with SNIa or Cosmic Microwave Background (CMB) measurements (see e.g. Betoule et al., 2014; Abbott et al., 2016; Planck Collaboration et al., 2016a; Alam et al., 2017). In both of these cases, the information from the two sets of probes is largely uncorrelated. However, weak lensing and galaxy clustering, i.e. the major cosmological probes from large galaxy surveys, are highly correlated with each other in that they are tracers of the same underlying density field, and in that they share common systematic effects. Moreover, future surveys will cover large, overlapping regions of the observable Universe and will therefore not be statistically independent. In general, each cosmological probe provides a measurement of the cosmic structures through a different physical field, such as density, velocity, gravitational potentials, and temperature. A promising way to test for new physics, such as modified gravity, is to look directly for deviations from the expected relationships of the statistics of the different fields. Consequently, a multi-probe analysis based on correlated photometric probes can no longer simply combine the optimal versions of individual analyses. Instead, to take full advantage of the power of combining probes of large-scale structure, one must build a tailored analysis pipeline that can model cosmological observables and their correlated systematics consistently. In addition to this modeling framework, multi-probe analyses require the ability to compute joint covariance matrices that properly account for the cross-correlation of various observables. The integrated treatment of the probes from the early stages of the analysis will thus provide the cross-checks and the redundancy needed not only to achieve high-precision but also to challenge the different sectors of the cosmological model.

Several earlier studies have considered joint analyses of various cosmological probes. Mandelbaum et al. (2013); Cacciato et al. (2013) and Kwan et al. (2017) for example derived cosmological constraints from a joint analysis of galaxy-galaxy lensing and galaxy clustering. Recently, Singh et al. (2017) performed a joint analysis of CMB lensing as well as galaxy clustering and weak lensing. Furthermore, Eifler et al. (2014) and Krause & Eifler (2017) have theoretically investigated joint analyses for photometric galaxy surveys by modelling the full non-Gaussian covariance matrix between cosmic shear, galaxy-galaxy lensing, galaxy clustering, photometric Baryon Acoustic Oscillations (BAO), galaxy cluster number counts and galaxy cluster weak lensing.

The combination of galaxy clustering, cosmic shear, and galaxy-galaxy lensing measurements, in particular, powerfully constrains structure formation in the late universe. It has been recognized for more than a decade that such a combination contains a tremendous amount of complementary information, as it is remarkably resilient to the presence of nuisance parameters that describe systematic errors and non-cosmological information (Joachimi & Bridle, 2010).

Such a combined analysis has recently been executed by van Uitert et al. (2018), who presented a joint analysis of cosmic shear, galaxy–galaxy lensing and galaxy clustering using power spectrum measurements, combining weak lensing from ~450 deg^2 of KiDS with a spectroscopic galaxy sample from the Galaxies And Mass Assembly (GAMA) survey in ~180 deg^2 of KiDS–GAMA overlap area.

In their analysis van Uitert et al. (2018) did not combine the different cosmological probes at the likelihood level; instead they followed a more optimal 'self-calibration' approach by modelling them within a single framework, as this enables a coherent treatment of systematic effects and a lifting of parameter degeneracies (Nicola et al., 2017). They adopted a formalism from Schneider et al. (2002) to estimate power spectra by performing simple integrals over the real-space correlation functions using appropriate weight functions. In a last step they computed the power spectrum covariance matrix and

combined it with theoretical predictions to derive constraints on cosmological parameters from a joint fit to the measured power spectra.

In the following, after briefly reviewing the theoretical background behind the power spectra for the three probes considered both in this work and in van Uitert et al. (2018), we describe the methodology we followed for our analysis. Since this shares some aspects with that of van Uitert et al. (2018), we will focus in particular those aspects that are different, in particular when altered in a modified gravity with respect to a standard ΛCDM scenario. We will not repropose here some of the details presented in van Uitert et al. (2018) (especially those concerning the computation of the power spectra from the correlation functions) that are totally unchanged with respect to their analysis. We refer the reader to the original paper van Uitert et al. (2018) for details regarding these common aspects.

6.2.2. Power spectra

The data vector for the analysis carried out in van Uitert et al. (2018) (and shared by our analysis) is represented by estimates of the projected power spectrum for each probe, as a function of the tomographic bin and the angular multipole ℓ.

Although the shear correlation functions may be easier to measure, power spectrum estimators have a number of advantages (Köhlinger et al., 2016). Firstly, they enable a clean separation of different ℓ-modes, which makes it easier to identify systematics that may affect only certain ℓ-modes. Furthermore, the covariance matrix of the power spectra is more diagonal than its real-space counterpart and this also helps towards a cleaner separation of scales. Finally, the power spectrum estimators can be readily modified to extract the B-mode part of the signal, which should be consistent with zero at lowest order in lensing if systematics are absent and hence serves as a systematic check.

For the estimation of the power spectra, the authors of van Uitert et al. (2018) followed the formalism, originally developed for cosmic shear only by Schneider et al. (2002), to calculate the power spectra by performing simple integrals over the real-space correlation functions using appropriate weight functions. van Uitert et al. (2018) extended the formalism to galaxy–galaxy lensing and galaxy clustering, integrating over the corresponding real-space correlation functions, which can be readily measured with existing public code.

There are several advantages from this type of approach to power spectra estimation. Computing power spectra directly from the data, for example using a quadratic estimator (Hu & White, 2001), is usually a complicated and CPU-intensive task (e.g. Köhlinger et al., 2016). This is particularly challenging for cosmic shear studies as the high signal-to-noise regime of the cosmological measurements is on relatively small scales, thus requiring high resolution measurements. Alternatively, pseudo-C_ℓ methods can be used (Hikage et al., 2011; Asgari et al., 2018), but they are sensitive to the details of the survey mask.

We conform to van Uitert et al. (2018) in calling the power spectra associated to cosmic shear, galaxy-galaxy lensing and galaxy clustering as $P^{\rm E}, P^{\rm gm}, P^{\rm gg}$, respectively. The cosmic shear label stands for E-mode, as in fact van Uitert et al. (2018) also measured the B-modes power spectra: they found that the inclusion of B-modes corrections shifted their main cosmological result by less than 0.5σ and therefore decided to ignore those B-modes corrections. Since P^E does not vary rapidly with ℓ, van Uitert et al. (2018) only needed a few ℓ-bins to capture most of the cosmological information. They used five logarithmically-spaced bins, whose logarithmic means range from $\ell = 200$ to $\ell = 1500$. $P^{\rm gm}$ and $P^{\rm gg}$ are estimated adopting the same ℓ ranges.

Modifications of Poisson equation and ratio of Bardeen potentials

In order to provide equations for the power spectra of the three probes in a general modified gravity scenario, we need to introduce the modifications to the Poisson equation and the ratio of the Bardeen

potentials that distinguish a modified gravity theory from General Relativity in absence of anisotropic stress. In Newtonian gauge the linear perturbation equations in Fourier space in a general modified gravity scenario are given by

$$\Phi = -\frac{3}{2}\frac{\Omega_m H_0^2}{k^2}\frac{\delta}{a}\,\mu(k,\chi) \tag{6.22}$$

$$\frac{\Psi}{\Phi} = \eta(k,\chi) \tag{6.23}$$

where μ and η are in general functions of both time and scale, and equal to 1 in General Relativity in absence of anisotropic stress. The equations presented in the following are for this generic modified gravity scenario, and the ones used in van Uitert et al. (2018) for ΛCDM can be deduced from ours by setting $\mu = \eta = 1$ identically. For all our power spectrum estimators we assume the Limber approximation, in the form given by Loverde & Afshordi (2008), which uses $\ell+1/2$ in the argument of the matter power spectrum but no additional prefactors. A number of recent papers have demonstrated for the case of cosmic shear that these approximations are very good on the scales that we consider (Kitching et al., 2017; Lemos et al., 2017; Kilbinger et al., 2017).

Cosmic shear

The weak lensing convergence power spectrum can be obtained from the 3-D matter power spectrum P_δ via

$$\boxed{P_{ij}^{E}(\ell) = \left(\frac{3H_0^2\Omega_m}{2c^2}\right)^2 \int_0^{\chi_H} \mathrm{d}\chi\, \frac{g_i(\chi)g_j(\chi)}{a(\chi)^2}\, P_\delta\left(\frac{\ell+1/2}{\chi};\chi\right) \mu\left(\frac{\ell+1/2}{\chi},\chi\right)^2 \left(\frac{1+\eta\left(\frac{\ell+1/2}{\chi},\chi\right)}{2}\right)^2} \tag{6.24}$$

where

$$g_i(\chi) = \int_\chi^{\chi_H} \mathrm{d}\chi' n_i(\chi')\frac{\chi'-\chi}{\chi'} \tag{6.25}$$

and with H_0 the Hubble constant, Ω_m the present-day matter density parameter, c the speed of light, χ the comoving distance, $a(\chi)$ the scale-factor, $f_K(\chi)$ the comoving angular diameter distance, χ_H the comoving horizon distance, and $g(\chi)$ a geometric weight factor, which depends on the source redshift distribution $p_z(z)\mathrm{d}z = p_\chi(\chi)\mathrm{d}\chi$:

$$g(\chi) = \int_\chi^{\chi_\mathrm{H}} \mathrm{d}\chi'\; p_\chi(\chi')\frac{f_K(\chi'-\chi)}{f_K(\chi')}\,. \tag{6.26}$$

Hence for a given theoretical matter power spectrum P_δ, we can predict the observed convergence power spectrum once the source redshift distribution is specified.

Galaxy-galaxy lensing

The projected galaxy-matter power spectrum is related to the matter power spectrum via:

$$\boxed{P_{ij}^{\mathrm{gm}}(\ell) = b_i\left(\frac{3H_0^2\Omega_m}{2c^2}\right) \int_0^{\chi_H} \mathrm{d}\chi\, \frac{p_{Fi}(\chi)g_j(\chi)}{a(\chi)\chi} P_\delta\left(\frac{\ell+1/2}{\chi};\chi\right) \mu\left(\frac{\ell+1/2}{\chi},\chi\right)^2 \left(\frac{1+\eta\left(\frac{\ell+1/2}{\chi},\chi\right)}{2}\right)} \tag{6.27}$$

with $p_\mathrm{F}(\chi)$ the redshift distribution of the foreground sample. We assume that the galaxy bias is linear and deterministic[11] such that b_i is the effective bias of the lens sample labelled by the index i.

[11] In other words, the cross-correlation coefficient r presented in e.g. Pen (1998); Dekel & Lahav (1999) is fixed to unity.

Galaxy clustering

The angular power spectrum can be determined from the matter power spectrum via:

$$\boxed{P_{ij}^{\rm gg}(\ell) = b_i b_j \int_0^{\chi_H} {\rm d}\chi \, \frac{p_{Fi}(\chi) p_{Fj}(\chi)}{\chi^2} P_\delta\left(\frac{\ell+1/2}{\chi}; \chi\right) \mu\left(\frac{\ell+1/2}{\chi}, \chi\right)^2} \tag{6.28}$$

where, as above, b corresponds to the effective bias of the sample. The 0^{th} order Limber approximation for the angular correlation function is accurate to less than a percent at scales $\ell > 5\chi(z_0)/\sigma_\chi$, with $\chi(z_0)$ the comoving distance of the mean redshift of the foreground sample and σ_χ the standard deviation of the galaxies' comoving distances around the mean (see Sect. IV-B of Loverde & Afshordi, 2008). For our low- and high-redshift foreground samples, we obtain scales of $\ell \gtrsim 15$ and $\ell \gtrsim 25$, respectively. Since the minimum ℓ scale entering the analysis is 150, the Limber approximation is valid here.

6.2.3. Cosmological inference

We derive cosmological constraints from the measured tomographic power spectra in a Bayesian framework. For the estimation of model parameters θ we sample the likelihood

$$-2\ln\mathcal{L}(\theta) = \sum_{\alpha,\beta} d_\alpha(\theta)(\mathbf{C}^{-1})_{\alpha\beta}\, d_\beta(\theta)\,, \tag{6.29}$$

where the indices α, β run over the probes considered, as well as the tomographic bins and the angular multipole ℓ. The analytical covariance matrix $\mathbf{C}$ is calculated as outlined below. Eq. 6.29 assumes that the estimated power spectra are Gaussian distributed around their mean. While this is a valid approximation at high ℓ as a consequence of the Central Limit Theorem, we notice that its validity may be questionable at lower angular multipoles.

The components of the data vector are calculated as

$$d_\alpha(\mathbf{p}) = P_\alpha - \langle P_\alpha(\theta)\rangle^{\rm model}\,, \tag{6.30}$$

where the dependence on cosmological parameters enters only in the calculation of the predicted power spectra, $\langle P_\alpha(\ell)\rangle^{\rm model}$.

For an efficient evaluation of the likelihood (in the ΛCDM case) we employ the Nested Sampling algorithm MULTINEST (Feroz et al., 2009) included in the framework of the cosmological likelihood sampling package MONTE PYTHON (Audren et al., 2013) with which we derive all cosmology-related results in this analysis.

Covariance

As Eq. 6.29 indicates, the covariance matrix, or more precisely its inverse, the precision matrix, is the decisive quantity that determines the errors on cosmological parameters. Obtaining precision matrices is an area of active research; methods can be broadly separated into 3 categories: estimation from numerical simulations, estimation from data directly, and analytical modeling/computation. We briefly summarize the current state of research, before detailing the prescription that was followed for the calculation of our covariance matrix, which we share with the analysis of van Uitert et al. (2018).

- **Estimation from simulations** The precision matrix can be estimated from numerical N-body simulations using a standard Maximum Likelihood estimator. However, already in the single-probe case this approach is computationally prohibitively expensive due to the number of simulations required, as well as the high particle resolution and box size needed in the simulations

to overcome the intrinsic noise properties of the estimator; the situation becomes even more complicated in the multi-probe case, where the covariance is substantially larger.

- **Estimation from data** Estimating covariance matrices from the data directly (through bootstrap or Jackknife estimators[12]) avoids any assumptions about cosmological or other model parameters that need to be specified in the numerical simulation approach (and in the theoretical modeling approach). However, given the limited survey area, it is difficult to obtain a sufficiently large number of regions of sky for the method to work, and it is unclear if these regions can be treated as independent.

- **Analytical modeling/computation** The analytic computation of weak lensing covariances was detailed in Krause et al. (2017), which derive straightforward expressions for Fourier and configuration space covariances under the assumption that the density field is Gaussian, so that the four-point correlation of the density field can be expressed as the product of two-point correlations. On small and intermediate scales this assumption is inaccurate; analytical expressions of non-Gaussian weak lensing covariances were derived in Takada & Jain (2009) and Sato et al. (2009). These expressions were generalized to a 3x2pt analysis[13] in Krause & Eifler (2017). The main advantage of an analytical (inverse) covariance matrix is the lack of a noisy estimation process, which substantially reduces the computational effort in creating a large number of survey realizations; the disadvantage is that the modeling of the non-Gaussian covariance terms, which employs a halo model is less precise compared to sophisticated numerical simulations.

Our covariance matrix is unchanged with respect to the one used by van Uitert et al. (2018), i.e. is determined *analytically*, following a similar formalism as in Hildebrandt et al. (2017). The assumption behind this is that the covariance does not depend on the model considered. In order to compute the covariance matrix, an initial fiducial cosmology as well as values for the effective galaxy bias are assumed, which are the same used in van Uitert et al. (2018): for the fiducial cosmology, the best-fit parameters from Planck Collaboration et al. (2016a) are used, while the effective galaxy biases are assumed to be unity for both bins. We refer to van Uitert et al. (2018) for the description of an iterative approach that is used to update the covariance with the parameter values of the best-fitting model after the initial cosmological inference; such an iterative approach benefits greatly by the use of an analytical covariance matrix, which is relatively fast and easy to compute. We also refer to van Uitert et al. (2018) for details on how the complication arising from the partial sky overlap of the different probes is accounted for.

The analytical covariance matrix consists of three terms: (i) a Gaussian term that combines the Gaussian contribution to sample variance, shape noise, and a mixed noise-sample variance term, estimated following Joachimi et al. (2008), (ii) an in-survey non-Gaussian term from the connected matter trispectrum, and (iii) a super-sample covariance term, which deals with the cosmic variance modes that are larger than the survey window and couple to smaller modes within. To compute the latter two terms, van Uitert et al. (2018) closely followed the formalism outlined in Takada & Hu (2013), which can be readily expanded to galaxy-galaxy lensing and clustering measurements (e.g. Krause & Eifler, 2017). Intrinsic alignments and baryonic feedback are not included in the covariance modelling, since given that the measurements are dominated by the cosmological signals, the impact of the astrophysical nuisances on sample variance is small.

The covariance matrix includes the analytic cross-covariance between the different probes. One advantage of this approach is that it properly accounts for super-sample covariance. This term is typically underestimated when the covariance matrix is estimated from the data itself, for example through

[12] These estimators (the latter being a more sophisticated version of the former) first divide the data set into subsamples, which consist either of individual objects or groups of objects, which are then resampled in a particular way (Efron, 1979).

[13] '3x2pt' is the technical name given to this type of cross-correlation analysis shear-shear, shear-density, density-density.

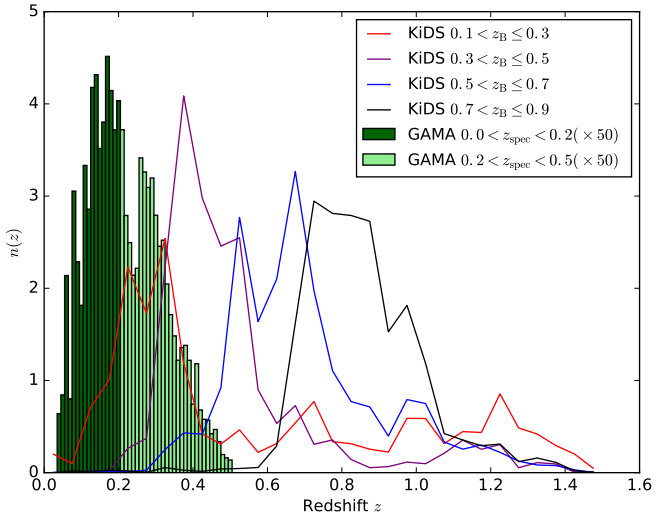

Figure 6.1.: Normalised redshift distribution of the four tomographic source bins of KiDS (*solid* lines), used to measure the weak gravitational lensing signal, and the normalised redshift distribution of the two spectroscopic samples of GAMA galaxies (*histograms*), that serve as the foreground sample in the galaxy-galaxy lensing analysis and that are used to determine the angular correlation function. For plotting purposes, the redshift distribution of GAMA galaxies has been multiplied by a factor 50.

jackknifing, or when it is estimated from simulations. Another advantage is that it is free of simulation sampling noise, which could significantly affect a joint probe analysis with large data vectors.

Data

Our power spectra are estimated from the KiDS-450 shape measurement catalogues (Hildebrandt et al., 2017), to measure the weak lensing signals, and the foreground galaxies from the GAMA survey (Driver et al., 2009, 2011; Liske et al., 2015) from the three equatorial patches that are completely covered by KiDS, to determine the galaxy-matter cross-correlation as well as the projected clustering signal.

The Kilo Degree Survey (KiDS; de Jong et al., 2013) is an optical imaging survey that aims to span 1500 deg^2 of the sky in four optical bands, *u*, *g*, *r* and *i*, complemented with observations in five infrared bands from the VISTA Kilo-degree Infrared Galaxy (VIKING) survey (Edge et al., 2013). The exceptional imaging quality particularly suits the main science objective of the survey, which is constraining cosmology using weak gravitational lensing. In this study, we use data from the most recent public data release, the KiDS-450 catalogues (Hildebrandt et al., 2017; de Jong et al., 2017), which contains the shape measurement and photometric redshifts of 450 deg^2 of data, split over five different patches on the sky, which include the three equatorial patches that completely overlap with GAMA.

In this work, we use the samples associated to the same four KiDS tomographic source redshift bins adopted in Hildebrandt et al. (2017) and van Uitert et al. (2018), spanning in redshift z_B the intervals $0.1 < z_B \leq 0.3$, $0.3 < z_B \leq 0.5$, $0.5 < z_B \leq 0.7$ and $0.7 < z_B \leq 0.9$. The redshift distribution of the four source samples is shown in Fig. 6.1: note that the redshift distributions are higher in correspondence of the relevant redshift bin, but their tails extend beyond that. The main properties of the source samples, such as their average redshift, number density and ellipticity dispersion, can be found in Table 1 of Hildebrandt et al. (2017).

GAMA (Driver et al., 2009, 2011; Liske et al., 2015) is a highly complete spectroscopic survey up to a Petrosian r-band magnitude of 19.8. In total, it targeted ∼240 000 galaxies. van Uitert et al. (2018) used a subset of ∼180 000 galaxies that reside in the three patches of 60 deg^2 each, as those patches fully overlap with KiDS. Due to the flux limit of the survey, GAMA galaxies have redshifts between 0 and 0.5. van Uitert et al. (2018) selected two GAMA samples, a low redshift sample with $z_{\rm spec} < 0.2$, and a high redshift sample with $0.2 < z_{\rm spec} < 0.5$. Their redshift distributions are also shown in Fig. 6.1.

Model fitting

To costrain the cosmological parameters, we used the sampler MontePython and built a likelihood module to analyse the three probes we considered. The likelihood module is based on the one developed in Köhlinger et al. (2017) for the KiDS-450 quadratic estimator analysis of cosmic shear data from the same 450deg^2 considered in our analysis. The likelihood module has been built having in mind as high priority the flexibility to choose different combinations of cosmological probes and allowing the user to perform the analysis in both a standard ΛCDM cosmological scenario and an extended modified gravity scenario within the Horndeski class of theories.

In contrast, the analysis of van Uitert et al. (2018) was carried out with a modified version of the sampler CosmoMC (Lewis & Bridle, 2002) for cosmological parameter estimation. The version used in van Uitert et al. (2018) is based on the one used in Joudaki et al. (2017) for the fiducial KiDS cosmic shear analysis, which includes prescriptions to deal with intrinsic alignment, the effect of baryons on the non-linear power spectrum, and systematic errors in the redshift distribution. van Uitert et al. (2018) extended the module to consider galaxy–galaxy lensing and angular clustering, and modified the code in order to fit the power spectra instead of the correlation functions.

Non-linear structure formation and baryonic feedback model

In order to derive accurate cosmological parameters from the cosmic shear power spectrum measurement it is important to account for a number of astrophysical systematics and model accurately the power on non linear scales.

The effect of non-linear structure formation and baryonic feedback can be now modelled in MontePython using a module called HMCODE (available to us in a beta, still unreleased version), which is based on the results of Mead et al. (2015). Feedback from AGN modifies the matter distribution at small scales (e.g. Semboloni et al., 2013), resulting in a modification of the dark matter power spectrum at high multipoles. Although the full physical description of baryon feedback is not established yet, hydrodynamical simulations offer one route to estimate its effect on the matter power spectrum. In general, the effect is quantified through a bias function with respect to the dark-matter only P_δ (e.g. Semboloni et al., 2013; Harnois-Déraps & van Waerbeke, 2015):

$$b^2(k,z) \equiv \frac{P_\delta^{\rm mod}(k,z)}{P_\delta^{\rm ref}(k,z)}\,, \tag{6.31}$$

where $P_\delta^{\rm mod}$ and $P_\delta^{\rm ref}$ denote the power spectra with and without baryon feedback, respectively. The KiDS cosmic shear analysis presented in Köhlinger et al. (2017) makes use of the results obtained from the OverWhelmingly Large Simulations (OWLS van Daalen et al., 2011) by implementing the fitting formula for baryon feedback from Harnois-Déraps & van Waerbeke (2015):

$$b^2(k,z) = 1 - A_{\rm bary}[A_z {\rm e}^{(B_z x - C_z)^3} - D_z x {\rm e}^{E_z x}]\,, \tag{6.32}$$

where $x = \log_{10}(k/1\,{\rm Mpc}^{-1})$ and the terms A_z, B_z, C_z, D_z, and E_z are functions of the scale factor $a = 1/(1+z)$. These terms also depend on the baryonic feedback model and we refer the reader to

Harnois-Déraps & van Waerbeke (2015) for the specific functional forms and constants. Additionally, Köhlinger et al. (2017) introduce a general free amplitude $A_{\rm bary}$ that they use as a free parameter to marginalise over while fitting for the cosmological parameters.

In the formalism of Mead et al. (2015), baryonic effects are accounted for by modifying the parameters that describe the shape of dark matter haloes. AGN and supernova feedback, for example, blow material out of the haloes, making them less concentrated. This is incorporated in HMCODE by choosing the following form for the mass-concentration relation,

$$c(M,z) = c_{\rm min}\frac{1+z_{\rm f}}{1+z}\ , \tag{6.33}$$

with $z_{\rm f}$ the formation redshift of a halo, which depends on halo mass. The free parameter in the fit, $c_{\rm min}$, modulates the amplitude of this mass-concentration relation. It also sets the amplitude of a 'halo bloating' parameter η_0 which changes the halo profile in a mass dependent way (see equation 26 of Mead et al., 2015). van Uitert et al. (2018) followed the recommendation of Mead et al. (2015) by fixing $\eta_0 = 1.03 - 0.11c_{\rm min}$. Setting $c_{\rm min} = 3.13$ corresponds to a dark-matter-only model. The resulting model is verified with power spectra measured on large hydrodynamical simulations, and found to be accurate to 5% for $k \leq 10h/{\rm Mpc}$. This is a relative uncertainty, not an absolute one (the absolute accuracy of any theoretical matter power spectrum prediction is not well established), and indicates the relative accuracy of their halo model fits with respect to hydrodynamical simulations, which are uncertain themselves. In addition, as Fig. 2 of Mead et al. (2015) shows, this accuracy is strongly k-dependent, and at small k ($k < 0.05\,h/{\rm Mpc}$), the agreement is much better than 5%. Therefore, putting a meaningful prior on the accuracy of the theory predictions is currently out of reach. However, the main source of theoretical uncertainty is caused by baryonic feedback, which mainly affects the small scales (high k). By marginalizing over $c_{\rm min}$, one can account for this main source of uncertainty. This will be done in the following every time we show the contour plots for pairs of parameters, marginalising over all the others.

Intrinsic alignments model

Intrinsic alignments affect both the cosmic shear power spectrum and the galaxy-matter power spectrum. For the cosmic shear power spectrum, there are two contributions, the intrinsic-intrinsic (II) and the shear-intrinsic (GI) terms (see Eqs. 5 and 6 of Joudaki et al., 2017). The galaxy-matter power spectrum has a galaxy-intrinsic contribution (e.g. Joachimi & Bridle, 2010). These three terms can be computed once the intrinsic alignment power spectrum is specified, which is assumed to follow the non-linear modification of the linear alignment model (Catelan et al., 2001; Hirata & Seljak, 2004; Bridle & King, 2007; Hirata & Seljak, 2010):

$$P_{\delta{\rm I}}(k,z) = -A_{\rm IA}C_1\rho_{\rm crit}\frac{\Omega_{\rm m}}{D(k,z)}P_\delta(k,z)\ , \tag{6.34}$$

with $P_\delta(k,z)$ the full non-linear matter power spectrum, $D(k,z)$ the linear growth factor, normalised to unity at $z = 0$ and in general dependent on scale and redshift in modified gravity, $\rho_{\rm crit}$ the critical density, $C_1 = 5\times10^{-14}h^{-2}M_\odot^{-1}{\rm Mpc}^3$ a normalization constant, and $A_{\rm IA}$ the overall amplitude, which is a free parameter in our model. Our intrinsic alignment model is minimally flexible with a single, global amplitude parameter.

To model $P^{\rm gm}$ and $P^{\rm gg}$, we assume that the galaxy bias is constant and scale-independent. Since we include non-linear scales in our fit, this bias should be interpreted as an effective bias. It is fitted separately for the low-redshift and high-redshift foreground sample. The scale dependence of the bias has been constrained in observations by combining galaxy-galaxy lensing and galaxy clustering measurements for various flux-limited samples and was found to be small (e.g. Hoekstra et al., 2002;

Table 6.2.: Priors on the fit parameters. Rows 1–5 contain the priors on cosmological parameters, rows 6–9 the priors on astrophysical 'nuisance' parameters, rows 10–13 the priors on modified gravity parameters. All priors are flat within their ranges.

Parameter	Description	Prior range
$\Omega_c h^2$	Cold dark matter density	$[0.01, 0.99]$
$\Omega_b h^2$	Baryon density	$[0.019, 0.026]$
$\ln(10^{10} A_s)$	Scalar spectrum amplitude	$[1.7, 5.0]$
n_s	Scalar spectral index	$[0.7, 1.3]$
h	Dimensionless Hubble parameter	$[0.64, 0.82]$
A_{IA}	Intrinsic alignment amplitude	$[-6, 6]$
c_{min}	Baryonic feedback amplitude	$[2, 4]$
b_{z1}	Galaxy bias of low-z lens sample	$[0.1, 5]$
b_{z2}	Galaxy bias of high-z lens sample	$[0.1, 5]$
$\hat{\alpha}_B$	Prop. coeff. $\alpha_B = \hat{\alpha_B}\Omega_{DE}$	$[-2, 2]$
$\hat{\alpha}_M$	Prop. coeff. $\alpha_M = \hat{\alpha}_M\Omega_{DE}$	$[-2, 2]$
Ω_0	Prop. coeff. $\Omega(\tau) = \Omega_0 a(\tau)$	$[0, 1]$
k_s	screening scale	$[0, 10]\mathrm{Mpc}^{-1}$

Simon et al., 2007; Jullo et al., 2012; Cacciato et al., 2012). We refer to van Uitert et al. (2018) for a study of how this approximation of scale-independent bias is benign in their analysis as well as in ours.

For cosmic shear, the intrinsic alignments contributions are given by:

$$P_{ij}^{GI}(\ell) = \left(\frac{3H_0^2\Omega_m}{2c^2}\right)\int_0^{\chi_H} \mathrm{d}\chi \, \frac{g_i(\chi) n_j(\chi) F_j\left(\frac{\ell+1/2}{\chi}, \chi\right)}{a(\chi)\chi} P_\delta\left(\frac{\ell+1/2}{\chi}; \chi\right) \times \mu\left(\frac{\ell+1/2}{\chi}, \chi\right)^2 \left(\frac{1+\eta\left(\frac{\ell+1/2}{\chi}, \chi\right)}{2}\right) \quad + \quad i \leftrightarrow j \tag{6.35}$$

$$P_{ij}^{II}(\ell) = \int_0^{\chi_H} \mathrm{d}\chi \, \frac{n_i(\chi) n_j(\chi) F_i\left(\frac{\ell+1/2}{\chi}, \chi\right) F_j\left(\frac{\ell+1/2}{\chi}, \chi\right)}{\chi^2} P_\delta\left(\frac{\ell+1/2}{\chi}; \chi\right) \times \mu\left(\frac{\ell+1/2}{\chi}, \chi\right)^2, \tag{6.36}$$

while for galaxy-galaxy lensing, the intrinsic alignments contribution is given by:

$$P_{ij}^{gI}(\ell) = \int_0^{\chi_H} \mathrm{d}\chi \, n_j(\chi) p_{Fi}(\chi) F_j\left(\frac{\ell+1/2}{\chi}, \chi\right) \frac{b_j(\chi)}{\chi^2} P_\delta\left(\frac{\ell+1/2}{\chi}; \chi\right) \mu\left(\frac{\ell+1/2}{\chi}, \chi\right)^2. \tag{6.37}$$

Priors

We adopt top-hat priors on the cosmological parameters, as well as the physical 'nuisance' parameters discussed earlier in this section. The prior ranges are listed in Table 6.2. Furthermore, we fix k_{pivot}, the pivot scale where the scalar spectrum has an amplitude of A_s, to 0.05/Mpc. Even though the sum of the neutrino masses is known to be non-zero, we adopt the same prior as Hildebrandt et al. (2017) and van Uitert et al. (2018) and fix it to zero.

Parameter	$P^{\rm E}$	$P^{\rm E} + P^{\rm gm} + P^{\rm gg}$	$P^{\rm E} + P^{\rm gm}$	$P^{\rm E} + P^{\rm gg}$	$P^{\rm gm} + P^{\rm gg}$
$\Omega_{\rm cdm}h^2$	$0.139^{+0.051}_{-0.041}$	$0.165^{+0.033}_{-0.045}$	$0.142^{+0.032}_{-0.043}$	$0.177^{+0.035}_{-0.038}$	$0.175^{+0.034}_{-0.051}$
$\ln10^{10}{\rm A_s}$	$2.65^{+0.21}_{-0.92}$	$2.47^{+0.44}_{-0.48}$	$2.63^{+0.43}_{-0.72}$	$2.22^{+0.14}_{-0.49}$	$2.53^{+0.41}_{-0.48}$
$\Omega_{\rm b}h^2$	$0.0225^{+0.0030}_{-0.0033}$	$0.0225^{+0.0030}_{-0.0033}$	$0.0225^{+0.0036}_{-0.0031}$	$0.0225^{+0.0034}_{-0.0029}$	$0.0226^{+0.0035}_{-0.0029}$
n_s	$1.09^{+0.21}_{-0.08}$	$0.95^{+0.10}_{-0.21}$	$1.03^{+0.20}_{-0.15}$	$0.93^{+0.11}_{-0.19}$	$0.89^{+0.09}_{-0.18}$
h	$0.74^{+0.07}_{-0.06}$	$0.73^{+0.08}_{-0.07}$	$0.74^{+0.08}_{-0.04}$	$0.73^{+0.07}_{-0.07}$	$0.74^{+0.08}_{-0.05}$
$c_{\rm min}$	$3.24^{+0.74}_{-0.54}$	$2.87^{+0.34}_{-0.79}$	$3.19^{+0.79}_{-0.36}$	$2.98^{+0.69}_{-0.66}$	$2.66^{+0.25}_{-0.65}$
A_{IA}	$0.91^{+0.75}_{-0.57}$	$1.24^{+0.37}_{-0.36}$	$1.36^{+0.36}_{-0.37}$	$0.88^{+0.65}_{-0.47}$	$1.38^{+0.42}_{-0.43}$
${\rm bias}_{z1}$	-	$1.17^{+0.16}_{-0.18}$	$0.84^{+0.15}_{-0.21}$	$1.25^{+0.16}_{-0.17}$	$1.18^{+0.18}_{-0.19}$
${\rm bias}_{z2}$	-	$1.29^{+0.18}_{-0.19}$	$1.55^{+0.28}_{-0.33}$	$1.49^{+0.20}_{-0.19}$	$1.27^{+0.19}_{-0.20}$
$\Omega_{\rm m}$	$0.30^{+0.08}_{-0.08}$	$0.35^{+0.06}_{-0.06}$	$0.30^{+0.06}_{-0.07}$	$0.38^{+0.06}_{-0.07}$	$0.36^{+0.06}_{-0.08}$
σ_8	$0.77^{+0.07}_{-0.16}$	$0.74^{+0.06}_{-0.09}$	$0.77^{+0.08}_{-0.12}$	$0.68^{+0.05}_{-0.08}$	$0.77^{+0.07}_{-0.10}$
S_8	$0.760^{+0.039}_{-0.038}$	$0.792^{+0.032}_{-0.031}$	$0.756^{+0.039}_{-0.035}$	$0.752^{+0.036}_{-0.036}$	$0.840^{+0.045}_{-0.040}$

Table 6.3.: Mean and marginalised 68% confidence interval on the parameters listed, obtained with our new likelihood in a ΛCDM scenario with the priors specified in Tab. 6.2.

6.2.4. Results

Our cosmological results have two primary goals:

- test the validity of the new pipeline in the ΛCDM scenario, with the goal of reproducing the results from van Uitert et al. (2018)
- produce constraints on the α functions belonging to the Bellini & Sawicki (2014) parametrization

Therefore the presentation of our results will follow these two main guidelines for the analysis of our work.

Comparison with van Uitert et al. (2018)

In Figs. 6.2, 6.3, 6.4, 6.5, 6.6 we show the comparison of our constraints with those produced by van Uitert et al. (2018). For all combinations of probes considered we find excellent agreement with the previous KiDSxGAMA analysis. Our results are obtained with a completely different implementation, which strengthens both the validity of the KiDS analysis and the consistency of our likelihood. In Fig. 6.7 the comparison is performed between constraints obtained only with our likelihood, and specifically comparing the cosmic shear only case with the three probes-case. We notice how the combination of probes tightens significantly the costraints, in particular on nuisance parameters such as the intrinsic alignment amplitude A.

Tabs. 6.3, 6.4 and 6.5 summarise the comparison between our results and the ones obtained by van Uitert et al. (2018): they show respectively the mean and 68% confidence intervals obtained from our analysis, the mean and 68% confidence intervals from van Uitert et al. (2018) and the percentage difference between the two analyses at the level of both the mean values and the 68% confidence intervals. We notice in particular the excellent agreement in the parameter $S_8 = \sigma_8 \sqrt{\Omega_{\rm m}/0.3}$, derived

Parameter	$P^{\rm E}$	$P^{\rm E}+P^{\rm gm}+P^{\rm gg}$	$P^{\rm E}+P^{\rm gm}$	$P^{\rm E}+P^{\rm gg}$	$P^{\rm gm}+P^{\rm gg}$
$\Omega_{\rm cdm}h^2$	$0.134^{+0.044}_{-0.046}$	$0.153^{+0.030}_{-0.040}$	$0.132^{+0.032}_{-0.044}$	$0.169^{+0.034}_{-0.039}$	$0.159^{+0.035}_{-0.041}$
$\ln10^{10}{\rm A_s}$	$2.76^{+0.30}_{-1.06}$	$2.63^{+0.36}_{-0.44}$	$2.81^{+0.52}_{-0.78}$	$2.30^{+0.21}_{-0.53}$	$2.67^{+0.37}_{-0.45}$
$\Omega_{\rm b}h^2$	$0.0224^{+0.0036}_{-0.0034}$	$0.0225^{+0.0035}_{-0.0035}$	$0.0224^{+0.0036}_{-0.0034}$	$0.0225^{+0.0034}_{-0.0035}$	$0.0225^{+0.0034}_{-0.0033}$
n_s	$1.11^{+0.19}_{-0.05}$	$0.97^{+0.15}_{-0.19}$	$1.08^{+0.22}_{-0.07}$	$0.97^{+0.14}_{-0.18}$	$0.93^{+0.08}_{-0.22}$
h	$0.74^{+0.08}_{-0.04}$	$0.73^{+0.09}_{-0.06}$	$0.74^{+0.08}_{-0.03}$	$0.73^{+0.04}_{-0.08}$	$0.73^{+0.09}_{-0.08}$
$c_{\rm min}$	$3.27^{+0.73}_{-0.23}$	$2.97^{+0.56}_{-0.71}$	$3.28^{+0.72}_{-0.22}$	$3.08^{+0.81}_{-0.39}$	$2.86^{+0.30}_{-0.84}$
A_{IA}	$0.92^{+0.78}_{-0.59}$	$1.27^{+0.39}_{-0.40}$	$1.46^{+0.41}_{-0.42}$	$0.88^{+0.70}_{-0.50}$	$1.38^{+0.46}_{-0.49}$
${\rm bias}_{z1}$	-	$1.12^{+0.14}_{-0.15}$	$0.78^{+0.14}_{-0.19}$	$1.21^{+0.14}_{-0.15}$	$1.13^{+0.15}_{-0.16}$
${\rm bias}_{z2}$	-	$1.25^{+0.16}_{-0.17}$	$1.45^{+0.27}_{-0.33}$	$1.45^{+0.18}_{-0.18}$	$1.23^{+0.16}_{-0.17}$
$\Omega_{\rm m}$	$0.29^{+0.07}_{-0.10}$	$0.33^{+0.05}_{-0.06}$	$0.29^{+0.06}_{-0.08}$	$0.36^{+0.06}_{-0.06}$	$0.34^{+0.05}_{-0.06}$
σ_8	$0.80^{+0.09}_{-0.18}$	$0.78^{+0.06}_{-0.08}$	$0.81^{+0.09}_{-0.14}$	$0.70^{+0.05}_{-0.08}$	$0.80^{+0.07}_{-0.09}$
S_8	$0.761^{+0.040}_{-0.038}$	$0.800^{+0.030}_{-0.026}$	$0.769^{+0.037}_{-0.032}$	$0.759^{+0.036}_{-0.032}$	$0.835^{+0.038}_{-0.037}$

Table 6.4.: Mean and marginalised 68% confidence interval on the parameters listed, obtained by van Uitert et al. (2018) in a ΛCDM scenario with the priors specified in Tab. 6.2.

Parameter	$P^{\rm E}$		$P^{\rm E}+P^{\rm gm}+P^{\rm gg}$		$P^{\rm E}+P^{\rm gm}$		$P^{\rm E}+P^{\rm gg}$		$P^{\rm gm}+P^{\rm gg}$	
	% diff. mean	% diff. 68% int.	% diff. mean	% diff. 68% int.	% diff. mean	% diff. 68% int.	% diff. mean	% diff. 68% int.	% diff. mean	% diff. 68% int.
$\Omega_{\rm cdm}h^2$	3.8	2.0	7.7	10.4	8.0	1.7	4.5	1.1	9.7	11.9
$\ln10^{10}{\rm A_s}$	4.0	16.8	5.8	14.3	6.1	11.3	3.2	15.1	5.3	8.7
$\Omega_{\rm b}h^2$	0.4	10.0	0	10.0	0.4	4.3	0	8.7	0.4	4.5
n_s	1.8	13.6	2.2	6.9	4.2	22.4	3.7	3.9	3.7	9.2
h	0.1	10.0	0.1	4.5	0.4	0.9	0.2	12.9	0.4	25.2
$c_{\rm min}$	0.8	32.8	3.5	11.0	2.9	22.5	3.2	12.0	6.9	20.7
A_{IA}	1.1	3.5	1.8	8.3	6.8	12.2	0.2	7.6	0.6	11.3
${\rm bias}_{z1}$	-	-	3.8	15.4	7.6	10.5	3.2	15.4	4.3	18.2
${\rm bias}_{z2}$	-	-	3.4	13.5	6.9	2.1	2.5	8.9	3.1	14.4
$\Omega_{\rm m}$	3.4	7.0	6.6	14.3	6.2	1.8	3.7	7.9	7.9	20.6
σ_8	3.6	15.8	4.0	3.9	4.9	13.2	2.6	0.8	3.1	0.10
S_8	1.5	0.6	1.0	13.5	1.6	6.1	0.9	5.9	0.6	14.9

Table 6.5.: Percentage difference on the mean values and on the 68% marginalised confidence intervals obtained from our analysis with respect to the results of van Uitert et al. (2018). The mean values and 68% marginalised confidence intervals of both analyses can be found in Tabs. 6.3 and 6.4.

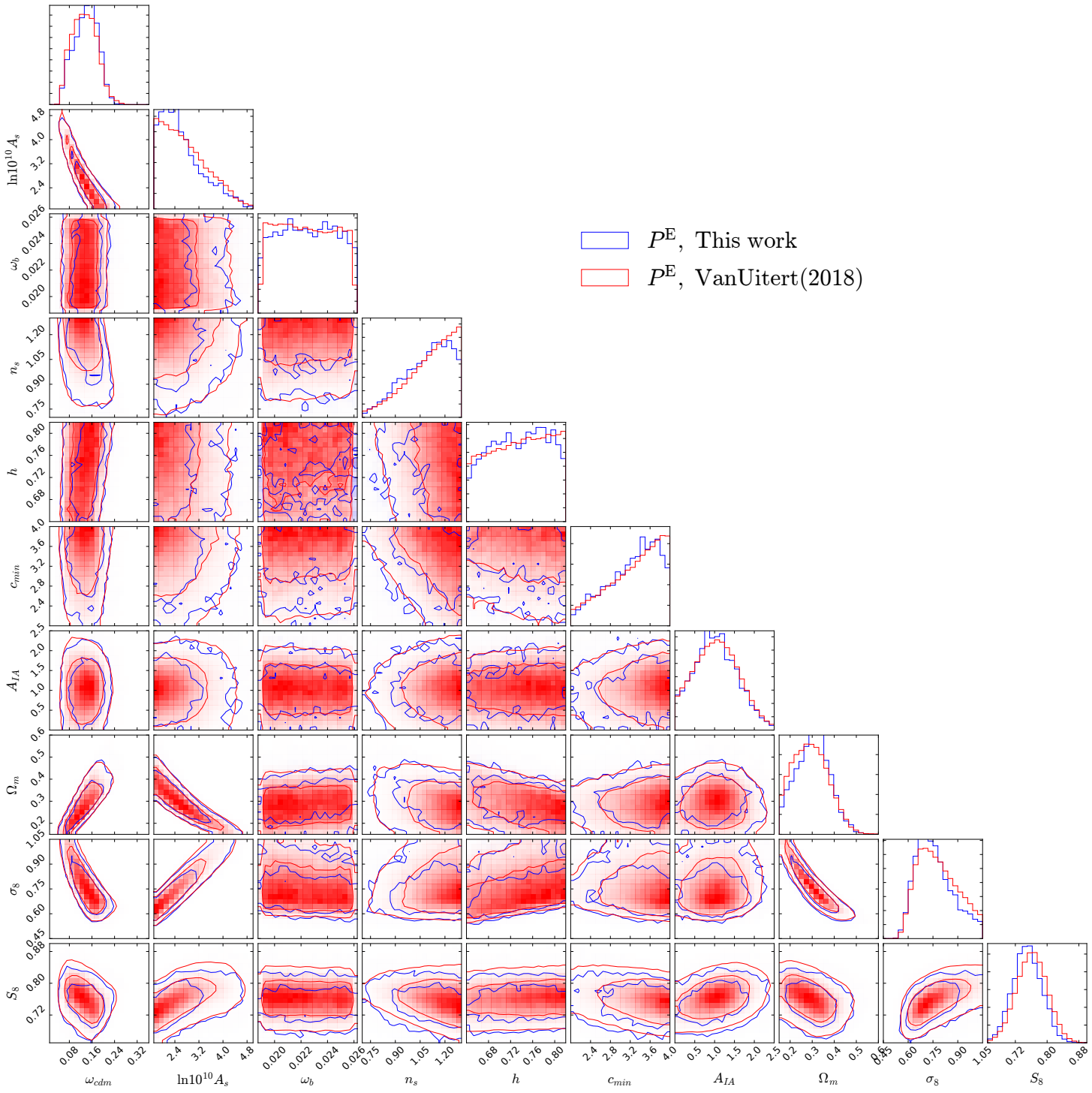

Figure 6.2.: Comparison of the marginalised 68 and 95% contours obtained with our new likelihood module for the KiDSxGAMA analysis and the results obtained by van Uitert et al. (2018). Here the probe considered is cosmic shear alone ($P^{\rm E}$).

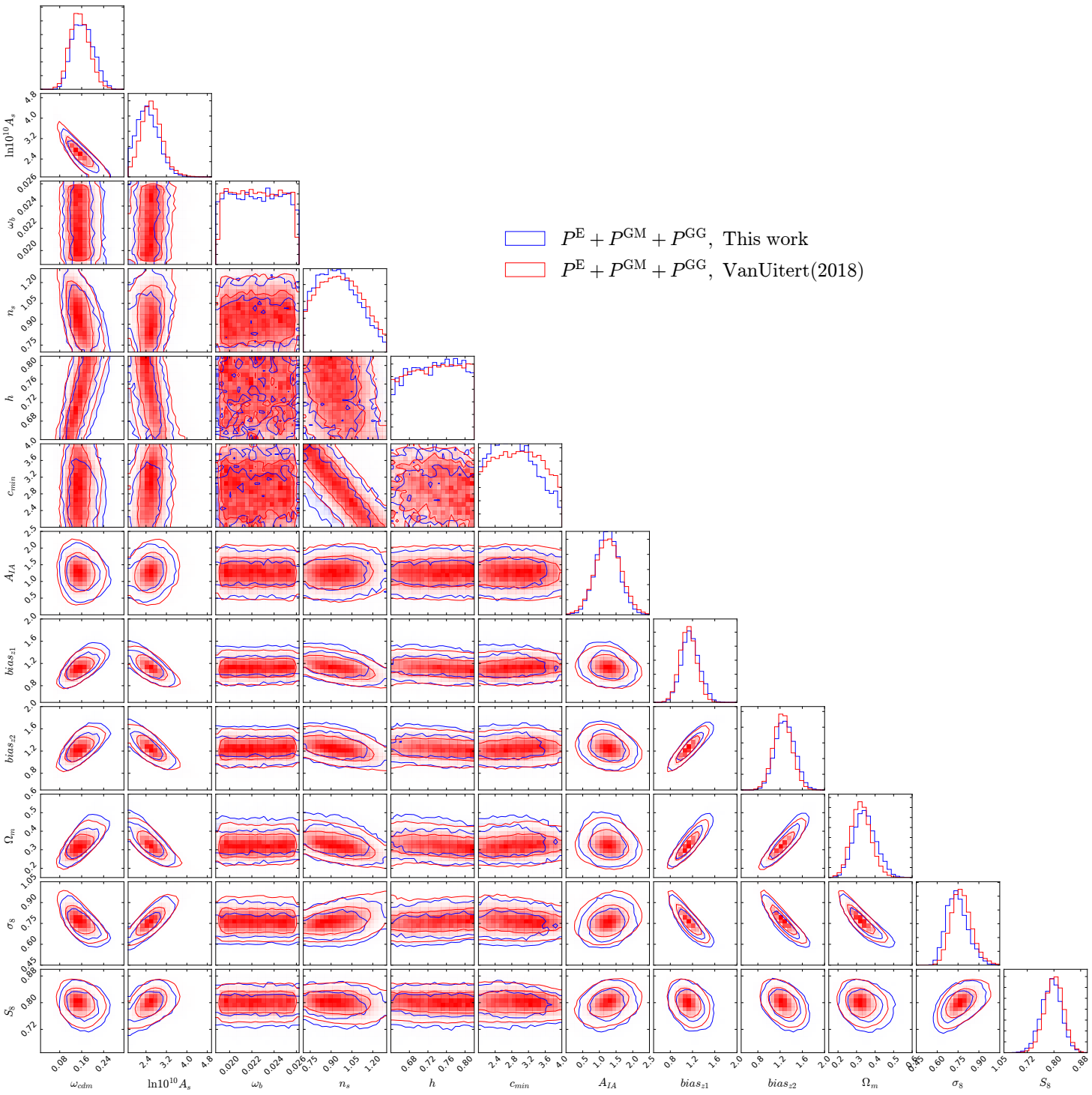

Figure 6.3.: Same as in Fig. 6.2, but here the combination of probes considered is given by all three probes in this analysis, i.e. cosmic shear, galaxy-galaxy lensing and angular clustering ($P^{\rm E} + P^{\rm GM} + P^{\rm GG}$).

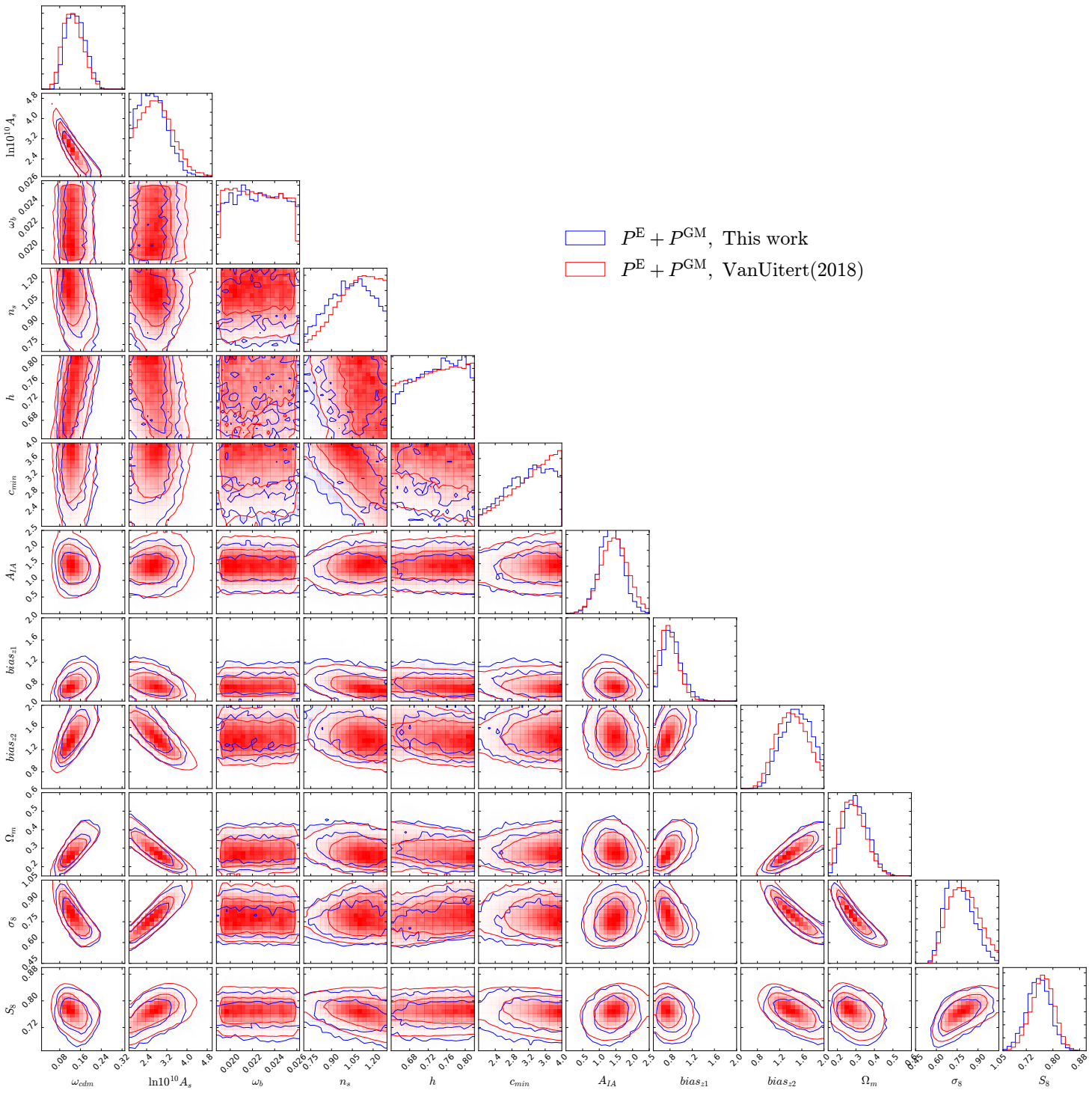

Figure 6.4.: Same as in Fig. 6.2, for the combination of cosmic shear and galaxy-galaxy lensing ($P^{\rm E}$ + $P^{\rm GM}$).

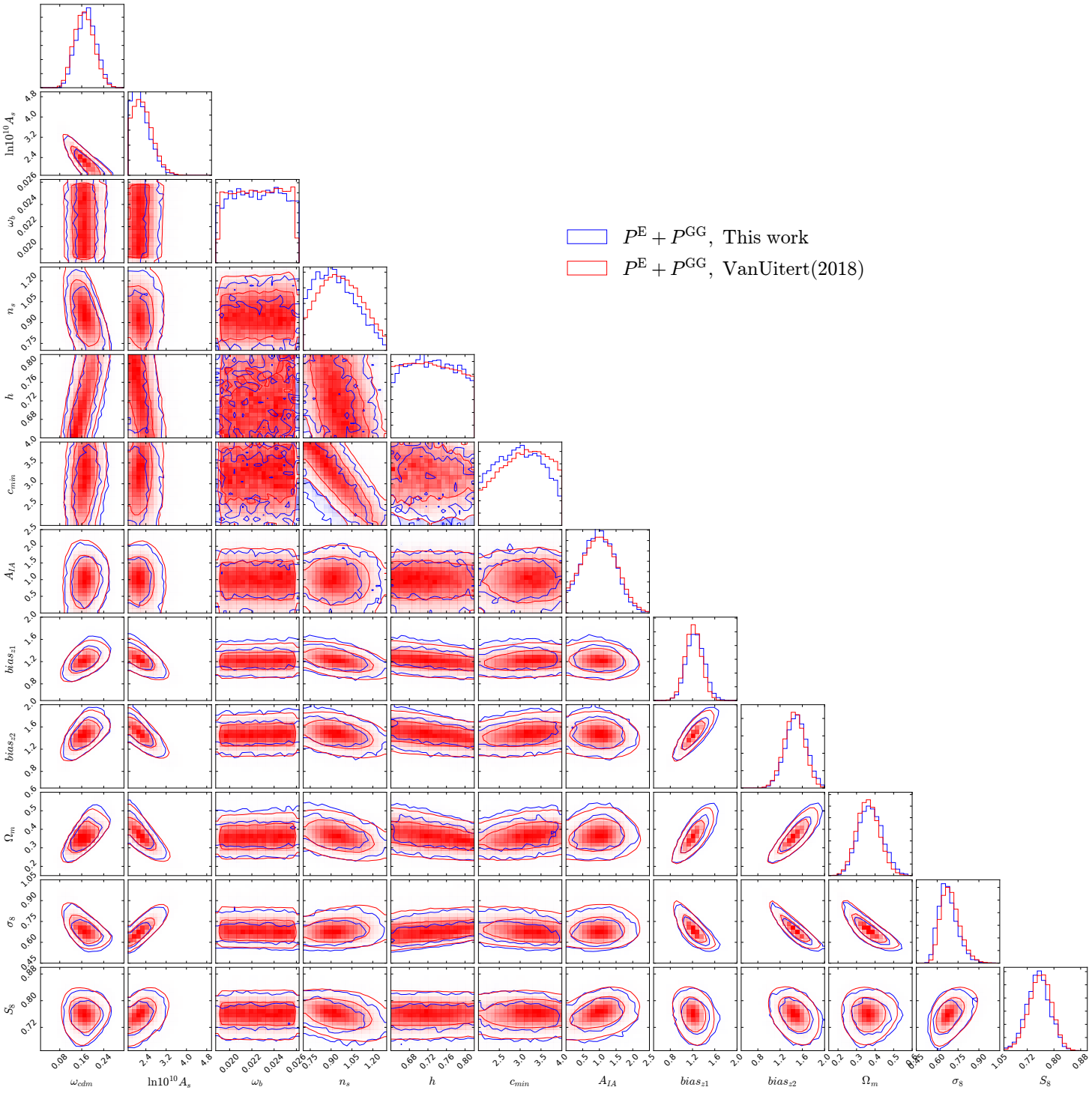

Figure 6.5.: Same as in Fig. 6.2, for the combination of cosmic shear and angular clustering ($P^{\rm E}$ + $P^{\rm GG}$).

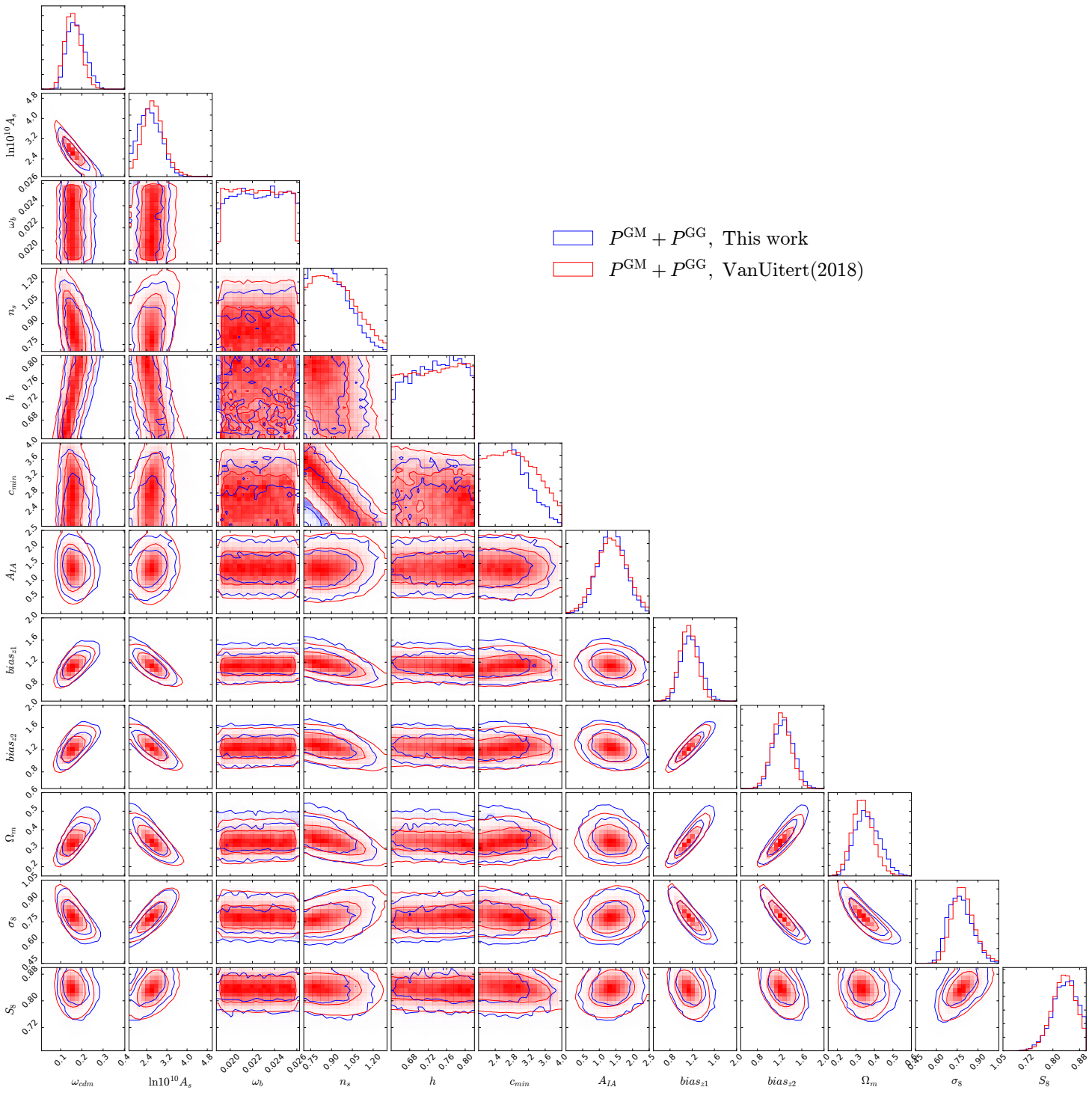

Figure 6.6.: Same as in Fig. 6.2, for the combination of galaxy-galaxy lensing and angular clustering ($P^{\rm GM} + P^{\rm GG}$).

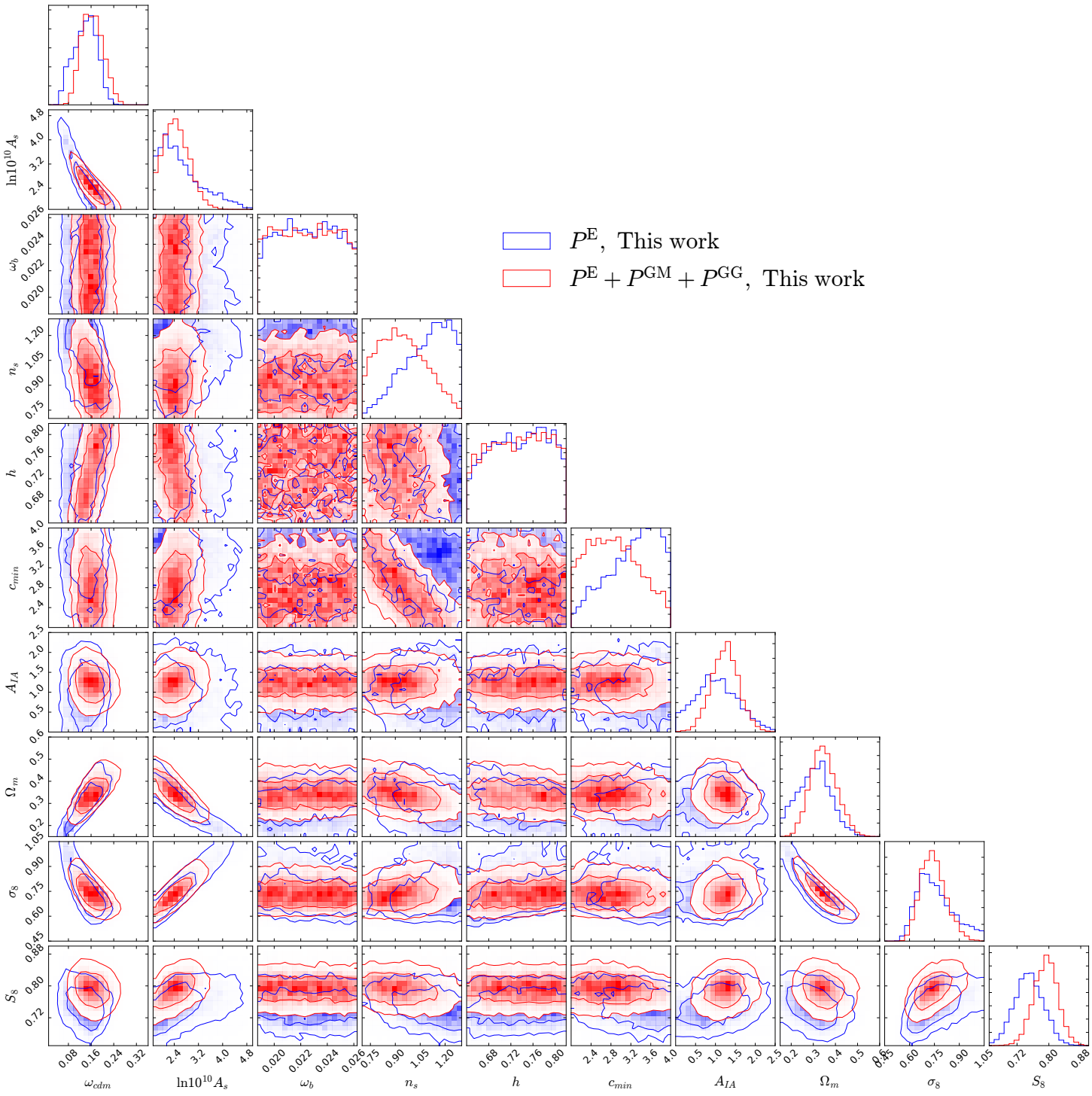

Figure 6.7.: Comparison between the constraints obtained with our 3 probes ($P^{\rm E}$ + $P^{\rm GM}$ + $P^{\rm GG}$) and with cosmic shear alone ($P^{\rm E}$).

Parameter	$P^{\rm E}$	$P^{\rm E} + P^{\rm gm} + P^{\rm gg}$	$P^{\rm E} + P^{\rm gm}$	$P^{\rm E} + P^{\rm gg}$	$P^{\rm gm} + P^{\rm gg}$
$\Omega_{\rm cdm}h^2$	$0.122^{+0.019}_{-0.029}$	$0.126^{+0.017}_{-0.024}$	$0.122^{+0.018}_{-0.024}$	$0.124^{+0.017}_{-0.026}$	$0.134^{+0.018}_{-0.023}$
$\ln10^{10}A_s$	$2.95^{+0.41}_{-0.35}$	$2.91^{+0.33}_{-0.23}$	$3.01^{+0.38}_{-0.28}$	$2.82^{+0.41}_{-0.31}$	$2.95^{+0.29}_{-0.22}$
$\Omega_b h^2$	$0.0225^{+0.0036}_{-0.0037}$	$0.0227^{+0.0035}_{-0.0040}$	$0.0222^{+0.0015}_{-0.0035}$	$0.0226^{+0.0037}_{-0.0038}$	$0.0228^{+0.0035}_{-0.0032}$
n_s	$1.07^{+0.23}_{-0.11}$	$1.14^{+0.16}_{-0.05}$	$1.13^{+0.17}_{-0.05}$	$1.15^{+0.15}_{-0.04}$	$1.09^{+0.21}_{-0.07}$
h	$0.73^{+0.09}_{-0.09}$	$0.74^{+0.08}_{-0.03}$	$0.76^{+0.06}_{-0.02}$	$0.74^{+0.08}_{-0.03}$	$0.74^{+0.08}_{-0.03}$
$c_{\rm min}$	$3.18^{+0.82}_{-0.34}$	$3.40^{+0.60}_{-0.23}$	$3.33^{+0.67}_{-0.20}$	$3.36^{+0.64}_{-0.24}$	$3.25^{+0.74}_{-0.26}$
$A_{\rm IA}$	$0.57^{+1.09}_{-0.66}$	$1.36^{+0.47}_{-0.43}$	$1.35^{+0.51}_{-0.45}$	$0.70^{+0.97}_{-0.60}$	$1.40^{+0.51}_{-0.52}$
bias_{z1}	-	$1.00^{+0.06}_{-0.08}$	$0.99^{+0.13}_{-0.14}$	$1.06^{+0.08}_{-0.14}$	$1.00^{+0.07}_{-0.09}$
bias_{z2}	-	$1.05^{+0.06}_{-0.09}$	$1.15^{+0.11}_{-0.21}$	$1.16^{+0.09}_{-0.18}$	$1.03^{+0.06}_{-0.09}$
σ_8	$0.85^{+0.09}_{-0.09}$	$0.88^{+0.06}_{-0.04}$	$0.91^{+0.09}_{-0.06}$	$0.84^{+0.09}_{-0.06}$	$0.92^{+0.06}_{-0.05}$
Ω_m	$0.27^{+0.04}_{-0.05}$	$0.27^{+0.02}_{-0.03}$	$0.25^{+0.03}_{-0.04}$	$0.27^{+0.03}_{-0.05}$	$0.29^{+0.02}_{-0.03}$
S_8	$0.800^{+0.050}_{-0.047}$	$0.834^{+0.037}_{-0.034}$	$0.832^{+0.041}_{-0.038}$	$0.790^{+0.043}_{-0.039}$	$0.894^{+0.046}_{-0.043}$
$\hat{\alpha}_{\rm B}$	$0.19^{+0.47}_{-0.53}$	$0.55^{+0.42}_{-0.62}$	$0.44^{+0.48}_{-0.73}$	$0.56^{+0.56}_{-0.67}$	$0.50^{+0.45}_{-0.64}$
$\hat{\alpha}_{\rm M}$	$0.56^{+0.44}_{-0.78}$	$0.12^{+0.31}_{-0.43}$	$0.53^{+0.54}_{-0.72}$	$0.34^{+0.38}_{-0.73}$	$0.23^{+0.35}_{-0.58}$
k_s	$0.17^{+0.02}_{-0.17}$	$0.12^{+0.05}_{-0.09}$	$0.14^{+0.01}_{-0.14}$	$0.10^{+0.02}_{-0.10}$	$0.11^{+0.03}_{-0.10}$

Table 6.6.: Mean and marginalised 68% confidence interval on the parameters listed obtained with all different combinations of probes. The parameterization for the α functions is given by proportionality to $\Omega_{\rm DE}$ so that the free parameters become the proportionality coefficients (in this case $\hat{\alpha}_{\rm B}$ and $\hat{\alpha}_{\rm B}$, since we fix $\alpha_{\rm T} = 0$ and $\hat{\alpha}_{\rm K} = 0.01$).

from $\Omega_{\rm m}$ and σ_8, i.e. the two parameters whose degenerate combination cosmic shear is most sensitive to (Hildebrandt et al., 2017).

Interpolations between the redshift distributions of the two foreground GAMA samples and the four KiDS cosmic shear background distributions (shown in Fig. 6.1) are needed, in order to calculate the galaxy-galaxy lensing signal (cf. Eq. 6.27), because the cosmic shear bins have redshift distributions more extended in z-range than the foreground samples, and at the same time the foreground samples have a finer resolution in z. In our implementation we decided to interpolate both background and foreground distributions. All interpolations are B-splines obtained with `scipy` in Python, specifically with the combination of the methods `splrep` and `splev`. The interpolation of the redshift distributions is a very delicate step in the implementation of the likelihood module: the lack of smoothness in the lensing source distributions makes the result strongly dependent on the methodology used for the interpolation.

The key to the success in the comparison with van Uitert et al. (2018) results has been the use of the Mead et al. (2015) recipe for the non-linear prescription. This has been made possible only very recently, thanks to a yet unpublished version of HMcode developed for the Boltzmann code Class (Lesgourgues, 2011).

The constraints on Modified Gravity have benefited greatly from this last point.

Parameter	$P^{\rm E}$	$P^{\rm E} + P^{\rm gm} + P^{\rm gg}$
$\Omega_{\rm cdm}h^2$	$0.130^{+0.015}_{-0.021}$	$0.134^{+0.016}_{-0.022}$
$\ln10^{10}{\rm A_s}$	$3.05^{+0.19}_{-0.18}$	$3.07^{+0.22}_{-0.20}$
$\Omega_b h^2$	$0.0223^{+0.0037}_{-0.0036}$	$0.0223^{+0.0040}_{-0.0035}$
n_s	$1.03^{+0.27}_{-0.21}$	$1.13^{+0.17}_{-0.05}$
h	$0.72^{+0.10}_{-0.08}$	$0.76^{+0.06}_{-0.02}$
$c_{\rm min}$	$3.15^{+0.85}_{-0.39}$	$3.27^{+0.73}_{-0.37}$
$A_{\rm IA}$	$0.34^{+1.12}_{-0.75}$	$1.30^{+0.46}_{-0.40}$
${\rm bias}_{z1}$	-	$1.00^{+0.05}_{-0.06}$
${\rm bias}_{z2}$	-	$1.04^{+0.05}_{-0.07}$
σ_8	$0.78^{+0.06}_{-0.08}$	$0.87^{+0.05}_{-0.04}$
$\Omega_{\rm m}$	$0.30^{+0.04}_{-0.05}$	$0.27^{+0.02}_{-0.03}$
S_8	$0.769^{+0.051}_{-0.046}$	$0.827^{+0.037}_{-0.028}$
Ω_0	$0.55^{+0.45}_{-0.20}$	$0.46^{+0.50}_{-0.46}$
k_s	$0.10^{+0.04}_{-0.05}$	$0.08^{+0.03}_{-0.07}$

Table 6.7.: Mean and marginalised 68% confidence interval on the parameters listed, obtained considering only cosmic shear ($P^{\rm E}$) and the combination of cosmic shear, galaxy-galaxy lensing and galaxy clustering ($P^{\rm E} + P^{\rm gm} + P^{\rm gg}$). The parameterization for the α functions is the `planck_linear` parameterization implemented in `hi_class`, where the α functions depend on only one function of time $\Omega(\tau)$ (see Eqs. 6.38), which is assumed to be proportional to the scale factor, $\Omega(\tau) = \Omega_0 a(\tau)$. The only free parameter as far as the α functions are concerned becomes then Ω_0.

Modified gravity constraints

In order to produce our constraints on the functions describing Horndeski theories at linear level, we used `hi_class` code (Zumalacárregui et al., 2017) as our Boltzmann code to extract our linear power spectrum in a Horndeski scenario.

The problem of non-linear contributions in this modified gravity context is even more exacerbated than in ΛCDM, since clear prescriptions for modified gravity models on nonlinear scales are even less abundant and reliable than for the standard cosmological scenario.

We decided to produce our nonlinear corrections using HMcode, which take into account baryonic effects. The main obstacle in this sense has been the absence of an implementation of HMcode in `hi_class`. We developed ourselves our version of HMcode for `hi_class`, using the one provided to us for Class as a starting point, but modifying substantially the code, not easily transferable to a modified gravity context.

With this at hand, we decided to investigate the strength of our likelihood by testing different time parameterizations for the α functions. We started with the usual proportionality to Ω_{DE}, and set constraints on the $\hat{\alpha}$ coefficients. Specifically, we considered $\hat{\alpha}_B$ and $\hat{\alpha}_M$ on top of our usual cosmological parameters. For both of them we took flat priors spanning $[-2, +2]$, however as we can see from the very sharp cutoff in the plane $\hat{\alpha}_B$-$\hat{\alpha}_M$ in Fig. 6.8, `hi_class` applies some stability checks that prevents the chains to end up in regions of parameter space where both of them are significantly negative. The investigation of these stability conditions is of the highest priority for extensions of this analysis.

Rather than plotting results for many different combinations of probes, we show in Fig.6.8 and 6.9 the comparison of the constraints obtained when considering cosmic shear only and the three probes, and we choose two different parameterization in time for the EFT functions. The first one is the standard proportionality to $\Omega_{\rm DE}$; the second parameterization is called `planck_linear` in `hi_class`, since it has been studied in Planck Collaboration et al. (2016b). It represents a subclass of models, specifically k-*essence* conformally coupled to gravity. In this parameterization the four α functions can specified through a single function of time $\Omega(\tau)$ affecting three of the α functions:

$$\begin{cases} \alpha_{\rm K} = & \frac{3(\rho_{\rm DE}+p_{\rm DE})}{H^2} + \frac{3\Omega(\rho_{\rm m}+p_{\rm m})}{H^2(1+\Omega)} - \frac{\Omega''-2aH\Omega'}{a^2H^2(1+\Omega)} \\ \alpha_{\rm M} = & \frac{\Omega'}{aH(1+\Omega)} \\ \alpha_{\rm B} = & -\alpha_{\rm M} \\ \alpha_{\rm T} = & 0\,. \end{cases} \tag{6.38}$$

The parameterization for $\Omega(\tau)$ is given by $\Omega_0 a(\tau)$: the only free modified gravity parameter becomes then Ω_0, for which we take the same prior range $\Omega_0 \in [0, 1]$ as in Planck Collaboration et al. (2016b).

In Figs. 6.8 and 6.9 we present marginalised 68% and 95% contours from the full likelihood analysis carried out on the parameter space. Our numerical results, always compatible with ΛCDM, are summarised in Tabs. 6.6 and 6.7. In this modified gravity case, since the chains are slower, we employed highly parallelised affine invariant sampling through the CosmoHammer (Akeret et al., 2012) suite, embedded in MontePython. CosmoHammer embeds in turn `emcee`, an implementation by Foreman-Mackey et al. (2013) of the affine invariant ensemble sampler by Goodman & Weare (2010).

Screening mechanisms were modelled in a phenomenological way, by applying a scale-dependent filter to the effective Newtonian coupling, the gravitational slip and the linear growth factor:

$$\begin{aligned} \mu(k,a) &\to \mu_{\rm GR} + \mu_{\rm MG}(a,k)\exp(-(k/k_{\rm s})^2) \\ \eta(k,a) &\to \eta_{\rm GR} + \eta_{\rm MG}(a,k)\exp(-(k/k_{\rm s})^2) \\ D_+(k,a) &\to D_{+\rm GR}(a) + D_{+\rm MG}(a,k)\exp(-(k/k_{\rm s})^2). \end{aligned} \tag{6.39}$$

This achieves the recovery of General Relativity predictions on scales $k > k_s$. A typical value for k_s would be $0.1h$Mpc^{-1}; we vary over this scale with a flat prior spanning $[0, 10]$Mpc^{-1}.

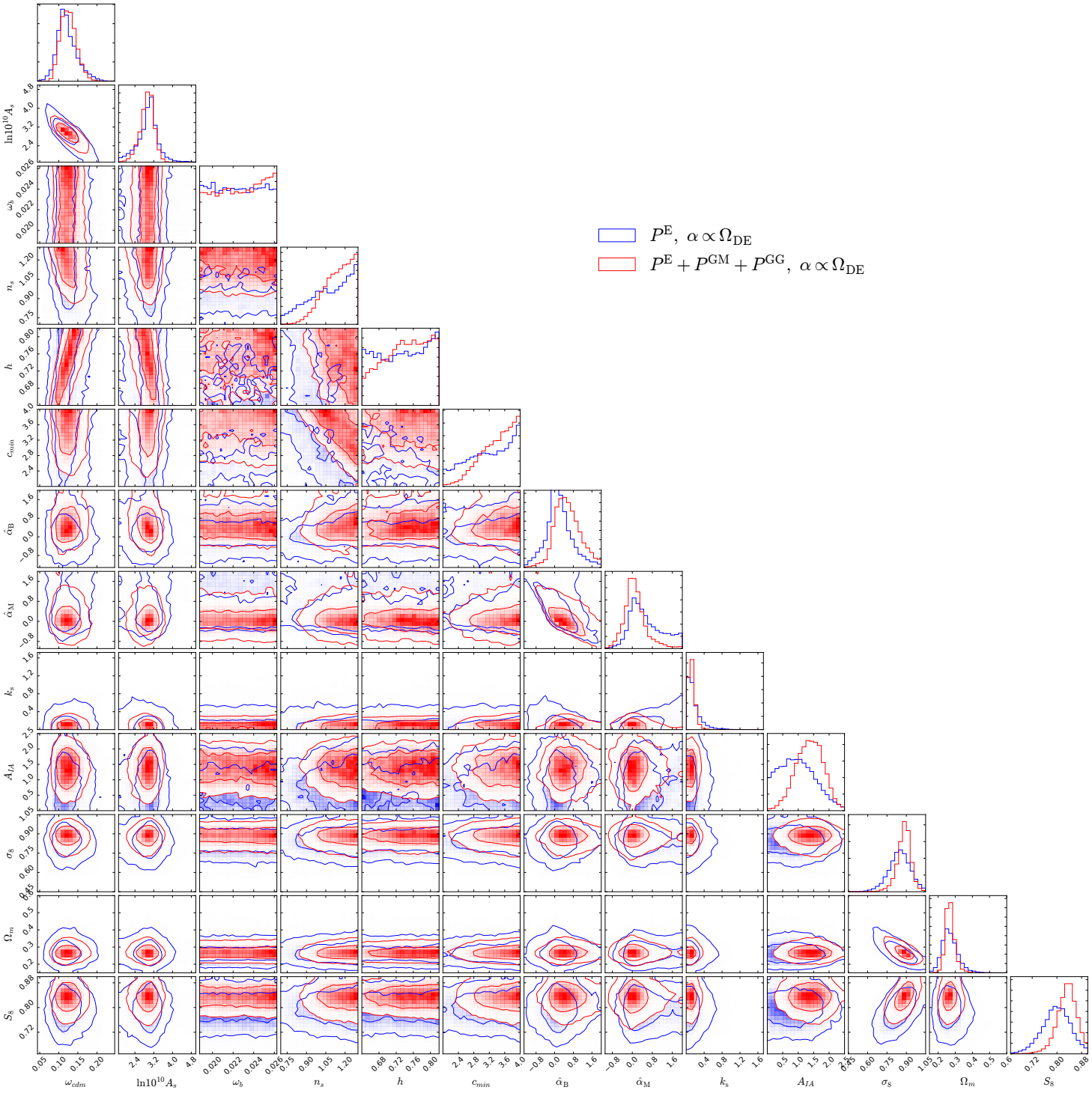

Figure 6.8.: Comparison of the marginalised 68 and 95% contours obtained with the three probes ($P^{\rm E}$ + $P^{\rm GM} + P^{\rm GG}$) and cosmic shear alone ($P^{\rm E}$) on the parameter space described earlier and augmented with the Horndeski parameters $\hat{\alpha}_B$ and $\hat{\alpha}_M$.

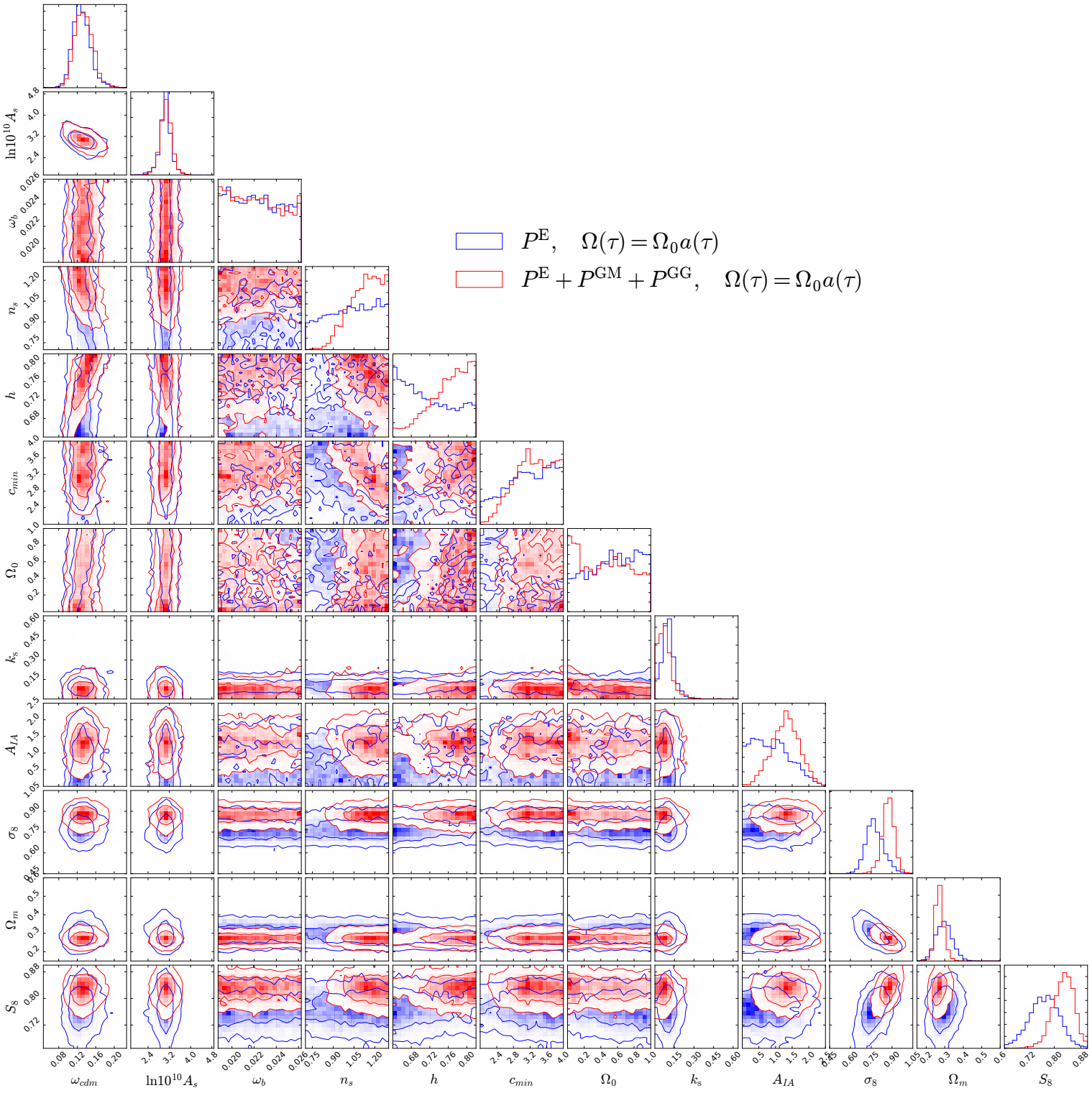

Figure 6.9.: Same as in Fig. 6.8, this time considering however the time dependence $\Omega(\tau) = \Omega_0 a(\tau)$ for Ω, the only free function in this parameterization, related to the α functions by Eqs. 6.38.

6.3. Cosmological constraints on Horndeski gravity from future cross-correlations

In this Section we take a look at future constraints on Horndeski gravity that may come from the synergy of different probes, which include cosmic shear. We produce forecasts considering cosmic shear, galaxy clustering, CMB primary anisotropies and CMB lensing. The forecasts are produced for future Stage IV surveys, such as Euclid (Laureijs et al., 2011) for galaxy clustering and cosmic shear, and the next generation ground-based Cosmic Microwave Background experiment CMB-S4 (Abazajian et al., 2016). Both CMB and optical surveys are expected to provide a dramatic leap forward in our understanding of the fundamental nature of space and time and the evolution of the Universe and in unravelling the origin of its accelerated expansion. The constraints we produce include numerical estimates for the expected sensitivity on the parameters that fully describe the evolution of linear perturbations in Horndeski gravity, introduced in Sec. 2.6.2.

A dataset as big as the one provided by the combination of stage IV surveys has the statistical power to achieve the significance required to constrain parameters belonging to a very large parameter space (including gravitational, cosmological, astrophysical and nuisance parameters). Additionally, the inclusion of CMB data in the analysis can help tighten the constraints coming from optical surveys on neutrino masses; this in turn will prove crucial in breaking degeneracies between neutrino masses and modifications of gravity (Baldi et al., 2014; Baldi & Villaescusa-Navarro, 2016).

However, such a big parameter space calls for sophisticated statistical techniques for the inference process. They will be certainly needed in the real data analysis phase, but are also important now, in preparation for these surveys. It would be highly desirable to have a realistic expectation of the constraining power of these surveys, which is as unbiased as possible in terms of assumptions entering the statistical analysis.

In particular, the Gaussian approximation for the posterior distribution, underlying the common use of the Fisher matrix as a method to produce forecasts for future experiments, needs to be carefully assessed in order to avoid erroneous conclusions on the expected sensitivity of future surves. This motivated the comparison, in this analysis, of results obtained with both a Fisher matrix approach and a full MCMC sampling of the posterior.

We give more details concerning the statistical tools employed in this analysis in Sec. 6.3.1. In Sec. 6.3.2 we present the probes considered and, very importantly, describe carefully all their possible correlations. In Sec. 6.3.3 we present our results. We draw our conclusions in Sec.6.4, jointly with those from the analysis presented in the previous Section.

This Section contains some of the results presented in the paper "Investigating scalar-tensor gravity with statistics of the cosmic large-scale structure" (Reischke et al., 2018), co-authored by the author of this thesis. The paper, submitted for publication to MNRAS and available online at arXiv:1804.02441, was authored by R. Reischke, A. Spurio Mancini, B. M. Schäfer and Ph. M. Merkel.

6.3.1. MCMC and Fisher forecasts

All probes considered (see Sec. 6.3.2 for a detailed description of each of them) are decomposed in spherical harmonic modes $\{a^X_{\ell m}\}$, where the superscript X labels the specific probe. Given statistical isotropy and homogeneity, the likelihood of a set of modes to be represented by its covariance is given by a product over the ℓ modes (and m modes, as evident by the exponent of each term that takes into account the multiplicity of the $2\ell+1$ m modes for each ℓ, since the covariance does not depend on m)

$$\mathcal{L}(\{\boldsymbol{a}_{\ell m}\}) = \prod_{\ell} p\left[\{\boldsymbol{a}_{\ell m}\}|\boldsymbol{C}_{\boldsymbol{a}}(\ell)\right]^{2\ell+1} \quad . \tag{6.40}$$

n_0	$\bar{z}$	$n_{\rm bin}$	$f_{\rm sky}$	$\ell_{\rm min}$	$\ell_{\rm max}$
$30\,{\rm arcmin}^{-2}$	0.9	6	$15.000\,{\rm deg}^2$	10	2000, 2500

Table 6.8.: Basic parameters of the experimental setup for a stage IV experiment. While the noise properties of galaxy clustering and cosmic shear is controlled by n_0, $n_{\rm bin}$ and $\sigma_\epsilon \approx 0.3$, the noise of CMB anisotropies depends on various parameters, which are summarised in Thornton et al. (2016). The two values for $\ell_{\rm max}$ refer to LSS and CMB observations respectively.

We assume that the modes are Gaussian distributed, i.e.

$$p\left(\{\boldsymbol{a}_{\ell m}\}\,|\boldsymbol{C}_{\boldsymbol{a}}(\ell)\right) = \frac{1}{\sqrt{(2\pi)^N {\rm det}\boldsymbol{C}(\ell)}} \exp\left[-\frac{1}{2}\boldsymbol{a}^{\dagger}_{\ell m}\boldsymbol{C}^{-1}_{\boldsymbol{a}}(\ell)\boldsymbol{a}_{\ell m}\right] . \tag{6.41}$$

The covariance matrix evaluated at the fiducial model $\hat{\mathbf{C}}_a$ has, for each ℓ and independently of m, the form given by

$$\hat{\mathbf{C}}_a = \begin{pmatrix}
TT & TE & 0 & TD & Tg_1 & \dots & Tg_n & T\gamma_1 & \dots & T\gamma_n \\
ET & EE & 0 & ED & 0 & \dots & 0 & 0 & \dots & 0 \\
0 & 0 & BB & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\
DT & DE & 0 & DD & Dg_1 & \dots & Dg_n & D\gamma_1 & \dots & D\gamma_n \\
g_1T & 0 & 0 & g_1D & g_1g_1 & \dots & g_1g_n & g_1\gamma_1 & \dots & g_1\gamma_n \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \dots & \vdots & \ddots & \vdots \\
g_nT & 0 & 0 & g_nD & g_ng_n & \dots & g_ng_1 & g_n\gamma_n & \dots & g_n\gamma_n \\
\gamma_1T & 0 & 0 & \gamma_1D & \gamma_1g_1 & \dots & \gamma_1g_n & \gamma_1\gamma_1 & \dots & \gamma_1\gamma_n \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \dots & \vdots & \ddots & \vdots \\
\gamma_nT & 0 & 0 & \gamma_nD & \gamma_ng_1 & \dots & \gamma_ng_n & \gamma_n\gamma_1 & \dots & \gamma_n\gamma_n
\end{pmatrix} .$$

The expression for the average of the logarithmic likelihood averaged over the data is

$$\langle L\rangle = \sum_{\ell}(2\ell+1)\left[\ln\left({\rm det}\boldsymbol{C}_{\boldsymbol{a}}\right) + {\rm tr}(\boldsymbol{C}^{-1}_{\boldsymbol{a}}\hat{\boldsymbol{C}}_{\boldsymbol{a}})\right] , \tag{6.42}$$

If the prior is uniform, the posterior is proportional to the likelihood, allowing us to sample the likelihood directly to reconstruct the shape of the posterior distribution. The sampling is done with an affine invariant sampler based on the work of Goodman & Weare (2010). After initial burn in, we let the random walkers of the sampling method run for a time period equal to multiple times the autocorrelation time; this way we check for convergence of the chains, in combination with running many chains.

Our forecasts with the Fisher matrix method are a straightforward implementation of the expression for the Fisher matrix for data that are Gaussian distributed (Tegmark et al., 1997)

$$F_{\mu\nu} = \sum_{\ell=\ell_{\rm min}}^{\ell_{\rm max}} \frac{2\ell+1}{2}{\rm tr}\left(\partial_\mu {\rm ln}[\boldsymbol{C}_{\boldsymbol{a}}(\ell)]\,\partial_\nu {\rm ln}[\boldsymbol{C}_{\boldsymbol{a}}(\ell)]\right) . \tag{6.43}$$

6.3.2. Cosmological probes considered and their correlations

A crucial part of this analysis resides in the correct evaluation of the inter-probe correlations. From a statistical point of view, considering the different probes as correlated rather than independent leads to a decreasing signal with respect to the independent case. However, precious cosmological information is encoded in the cross-correlations and a valid cosmological model has to explain this, which may

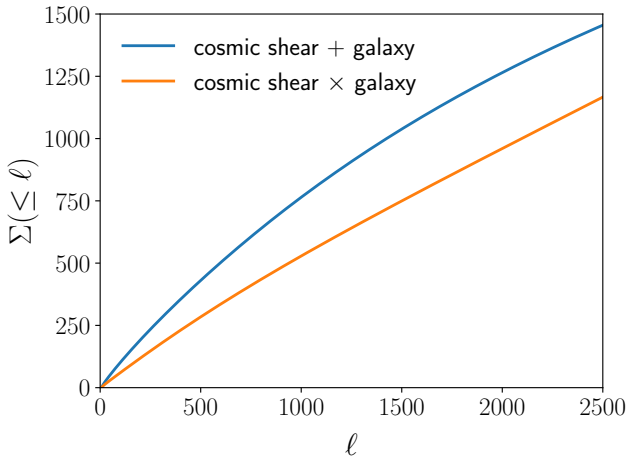

Figure 6.10.: *Right*: Cumulative signal-to-noise ratio (see Eq. 6.44). For a survey with settings as in Tab. 6.8 adding galaxy clustering and cosmic shear with (×) and without (+) cross-correlation.

finally lead to increased constraining power. We can appreciate this concept by looking at Fig.6.10, depicting the signal-to-noise ratio for the two probes galaxy clustering and cosmic shear, as a function of the multipole ℓ. The signal-to-noise ratio is given by

$$\Sigma^2(\leq \ell) = \sum_{\ell=\ell_{\rm min}}^{\ell_{\rm max}} \frac{2\ell+1}{2} {\rm tr}\left(\boldsymbol{C}_{\rm a}^{-1}(\ell)\boldsymbol{S}_{\rm a}(\ell)\boldsymbol{C}_{\rm a}^{-1}(\ell)\boldsymbol{S}_{\rm a}(\ell)\right) , \tag{6.44}$$

If we were to consider the two probes as uncorrelated we would overestimate the signal-to-noise ratio of $\simeq 16\%$. However, considering their cross-correlation in the analysis does not automatically imply that our constraints are less stringent than the uncorrelated case by the same amount; on the contrary, the inter-probe correlation contains valuable cosmological information, which tightens our constraints as long as our model is able to account for it.

We proceed here to list the probes considered in our analysis and for each of them we describe their power spectra:

- **Cosmic shear**. For this analysis we consider tomographic cosmic shear, since we need to perform fast evaluations of the covariance matrices in our MCMC chains, which would be extremely computationally challenging in the 3D cosmic shear formalism. We have already derived the expressions for the cosmic shear power spectrum in terms of the spectra of the Bardeen potentials Φ and Ψ, which we report here for simplicity; the angular power spectrum of the lensing potential ψ is related to that of the Bardeen potentials by

$$C_{\psi_i\psi_j}(\ell) = \int_0^{\chi_H} \frac{{\rm d}\chi}{\chi^2} W_{\psi_i}(\chi) W_{\psi_j}(\chi) P_{\Phi+\Psi}(\ell'/\chi, \chi) . \tag{6.45}$$

  where the tomographic efficiency function is

$$W_{\psi_i}(\chi) = \frac{1}{a\chi} \int_{\min(\chi,\chi_i)}^{\chi_{i+1}} {\rm d}\chi' p(\chi') \frac{{\rm d}z}{{\rm d}\chi'} \left(1 - \frac{\chi}{\chi'}\right) , \tag{6.46}$$

  The redshift distribution is given by the usual expression

$$p(z){\rm d}z \propto z^2 \exp\left[-\left(\frac{z}{z_0}\right)^{3/2}\right] . \tag{6.47}$$

The noise contribution to the observed lensing spectra is given by the Poissonian shape noise arising from the finite number of galaxies in each bin that contribute to the lensing signal

$$\hat{C}_{\psi_i\psi_j} = C_{\psi_i\psi_j} + \sigma_\epsilon^2 \frac{n_{\rm bin}}{\bar{n}} \delta_{ij}, \tag{6.48}$$

In this work we did not consider intrinsic alignments of galaxies and we stress that in the future this could be a potentially interesting feature to investigate in order to quantify their impact on cosmological inference.

- **Galaxy clustering**. The main issue affecting galaxy clustering analyses is the missing knowledge of completely reliable models for the bias relating the distribution of matter to that of galaxies. If the bias is known or assumed, the angular clustering power spectrum reads

$$C_{g_ig_j}(\ell) = \int_0^{\chi_H} \frac{{\rm d}\chi}{\chi^2} W_{g_i}(\ell'/\chi,\chi) W_{g_j}(\ell'/\chi,\chi) P_\delta(\ell'/\chi,\chi), \tag{6.49}$$

and in addition we will have also in this case a Poissonian shot noise term,

$$\hat{C}_{g_ig_j} = C_{g_ig_j} + \frac{n_{\rm bin}}{\bar{n}} \delta_{ij}. \tag{6.50}$$

The galaxy weight function is defined as

$$W_{g_i}(\ell/\chi,\chi) = \frac{H(\chi)}{c} b(\ell/\chi,\chi) p(\chi) \text{ if } \chi \in [\chi_i,\chi_{i+1}). \tag{6.51}$$

The model we assume for the galaxy bias is linear in redshift and without scale dependence Ferraro et al. (2015)

$$b(\chi) = b_0(1 + z(\chi)) \ . \tag{6.52}$$

In our forecasts we do not consider correlations between the six bins; if we were to consider a higher number of bins, following Bailoni et al. (2017) we would have to include their correlations in the forecacst to avoid biases up to 30% in the estimated errors.

- **CMB primary anisotropies**. Maps of CMB temperature (T) and polarisation (E, B) data are given in spherical harmonics with some instrumental noise whose root mean square is σ_P. We characterize the noise contributions to the observed CMB spectra by the (Gaussian) beam size $\theta_{\rm beam}$ of the experiment and its instrumental sensitivity (white noise) with respect to temperature and polarization (Knox, 1995)

$$N^P(\ell) \equiv \langle n^{P*}_{\ell m} n^{P'}_{\ell m} \rangle = \theta^2_{\rm beam} \sigma_P^2 \exp\left[\ell(\ell+1)\frac{\theta^2_{\rm beam}}{8\ln 2}\right] \delta_{PP'} \ . \tag{6.53}$$

diagonal in P as the noise of different maps is uncorrelated. The angular power spectrum is then given by

$$\langle a^{P*}_{\ell m} a^{P'}_{\ell' m'} \rangle \equiv \hat{C}^{PP'}(\ell) = \left(C^{PP'}(\ell) + N^P(\ell)\right) \delta_{\ell\ell'} \delta_{mm'}. \tag{6.54}$$

and is diagonal in ℓ and m due to statistical isotropy and homogeneity.

- **CMB lensing**. The lensing potential for the CMB lensing case is

$$\psi_{\rm CMB} = \int_0^{\chi_H} {\rm d}\chi W_{\rm CMB}(\Phi + \Psi), \tag{6.55}$$

where the lensing efficiency function for the CMB, similarly to Eq. 6.46, is given by

$$W_{\rm CMB}(\chi) = \frac{\chi_* - \chi}{\chi_* \chi} \frac{H(\chi)}{ca}. \tag{6.56}$$

and where χ_* is the comoving distance to the last scattering surface The CMB lensing signal then has a similar expression to Eq.6.45, but Eq. 6.56 replaces the weight function of cosmic shear. The noise properties of the lensing signal are completely determined by the reconstruction of the lensing signal and its expression is given in Hu & Okamoto (2002); Okamoto & Hu (2003).

- **Cross-correlations of cosmic shear, galaxy clustering and CMB primary anisotropies and lensing**. The cross-spectra can be derived assigning weight functions and power spectra according to the combination of probes considered. For example galaxy clustering and cosmic shear will have a cross-spectrum given by

$$C_{g_i\psi_j}(\ell) = \int_0^{\chi_H} \frac{\mathrm{d}\chi}{\chi^2} W_{g_i}(\ell'/\chi, \chi) W_{\psi_j}(\ell'/\chi) P_{\delta,\Phi+\Psi}(\ell'/\chi, \chi). \tag{6.57}$$

to which a Poissonian noise term must be added, again due to the finite sample of the same galaxies used to calculate the lensing and the clustering signal. The correlation of cosmic shear the CMB lensing signal is described e.g. in Kitching et al. (2014); Merkel & Schäfer (2017) and similarly we can calculate the correlation between galaxy clustering and CMB lensing. Due to the integrated Sachs Wolfe effect (iSW, Sachs & Wolfe, 1967), CMB temperature fluctations are also correlated with the clustering, shear and CMB temperature signal. The iSW has a characteristic weight function

$$W_{\mathrm{iSW}}(k, a) = \frac{3}{2\chi_H^3} a^2 E(a) F'(k, a), \tag{6.58}$$

where the derivative is taken with respect to a and

$$F(k, a) = \mu(k, a) \frac{D_+(k, a)}{a} \left(1 + \frac{1}{\gamma(k, a)}\right). \tag{6.59}$$

In our analysis we do not consider the cross-correlation between CMB lensing and polarization, while the EB and TB cross-spectra vanish for parity arguments.

6.3.3. Cosmological constraints

The specifications used in our analysis for the optical surveys are reported in Tab. 6.8, while for the CMB anisotopies noise we refer to Thornton et al. (2016) as it depends on many parameters. We used linear predictions for the matter power spectrum (obtained with the software `hi_class` code (Zumalacárregui et al., 2017)). Halo model corrections (Smith et al., 2003; Takahashi et al., 2012) were applied to this linear power spectrum and baryonic effects modelled via a fitting formula by Schneider & Teyssier (2015), which adds some nuisance parameters to our analysis such as the average halo mass M_c. Similarly to what done in the KiDSxGAMA case, screening mechanisms were modelled in a phenomenological way, by applying a scale-dependent filter to the effective Newtonian coupling, the gravitational slip and the linear growth factor:

$$\begin{aligned} \mu(k, a) &\to \mu_{\mathrm{GR}} + \mu_{\mathrm{MG}}(a, k) \exp(-(k/k_s)^2) \\ \gamma(k, a) &\to \gamma_{\mathrm{GR}} + \gamma_{\mathrm{MG}}(a, k) \exp(-(k/k_s)^2) \\ D_+(k, a) &\to D_{+\mathrm{GR}}(a) + D_{+\mathrm{MG}}(a, k) \exp(-(k/k_s)^2). \end{aligned} \tag{6.60}$$

This achieves the recovery of General Relativity predictions on scales $k > k_s$.

In producing our forecasts we choose the time parameterization for the α functions given by the proportionality to the dark energy density fraction

$$\alpha_i = \hat{\alpha}_i \Omega_{\mathrm{DE}} \tag{6.61}$$

Parameter	fiducial	σ	$\sigma_{\rm F}$	prior	comments
$\hat{\alpha}_{\rm K}$	0.05	-	-	$[0,\infty)$	fixed
$\hat{\alpha}_{\rm B}$	0.1	0.067	0.037	$[0,\infty)$	
$\hat{\alpha}_{\rm M}$	0.2	0.050	0.035	$[0,\infty)$	
$\hat{\alpha}_{\rm T}$	0.01	-	-	$[0,\infty)$	fixed
w_0	-0.8	0.0135	0.0062	$[-1,1]$	
$w_{\rm a}$	0.2	0.030	0.017	$[0,\infty)$	
$\sum m_\nu$[eV]	0.05	0.017	0.012	$[0,\infty)$	
$\Omega_{\rm m0}$	0.314	0.0011	0.0009	$[0,\infty)$	
σ_8	0.834	0.0021	0.0017	$[0,\infty)$	
h	0.674	0.0013	0.0009	$[0,\infty)$	
$\Omega_{\rm b}$	0.0486	0.00019	0.00013	$[0,\infty)$	
$n_{\rm s}$	0.962	0.00145	0.00139	$[0,\infty)$	
$z_{\rm re}$	11.357	0.28	0.026	$[0,\infty)$	
$k_{\rm s}$	0.1	0.011	0.013	$[0,\infty)$	
b	0.68	0.0016	0.0013	$[0,\infty)$	
$M_{\rm c}$	0.26	0.0024	0.0022	$[0,\infty)$	
η_b	0.5	0.0063	0.006	$[0,\infty)$	

Table 6.9.: Parameters used for the inference process and the 68% marginalised errors from the Fisher analysis, $\sigma_{\rm F}$ and the MCMC σ. The fiducial values are taken from Planck Collaboration et al. (2016a), Schneider & Teyssier (2015) and Ferraro et al. (2015).

and set constraints on the proportionality coefficients. We vary only α_B and α_M, following the recent constraints on α_T from gravitational waves detection and considering that the kineticity α_K is unconstrained by cosmological observations. Additionally, we vary over the standard cosmological parameters reported in Tab. 6.9. We consider a time-varying equation of state for the dark energy component, parameterised by w_0 and w_a and include other phenomenological and astrophysical parameters that form in total a 17-dimensional parameter space, such as the reionisation redshift z_{re}, the bias parameter b and the baryons-to-photons ratio η_b.

In Fig. 6.11 we show constraints on α_B and α_M obtained marginalising over all the other parameters. In Fig. 6.12 we show a comparison between the 68% probability contours obtained with both an MCMC sampling and a Fisher matrix analysis. Each contour is marginalised over all the other parameters. The plot clearly shows that the Fisher matrix performs usually well in predicting degeneracies directions, however this is not true for the combination given by the sum of neutrino masses and the braiding proportionality coefficient. Furthermore, the errors obtained with the Gaussian approximation are constantly underestimated, confirming the Cramér-Rao bound.

6.4. Discussion and conclusions

In this Chapter we have investigated constraints achievable from current and future cross-correlation analyses of cosmic shear with other probes. We have analysed real data from the KiDS and GAMA survey, in a joint framework for tomographic cosmic shear power spectra, galaxy-galaxy lensing power spectra and angular clustering power spectra. Our numerical implementation reproduces the results of van Uitert et al. (2018) and extends their analysis to a Horndeski scenario. As far as future surveys are concerned, we have produced forecasts, with both a full MCMC reconstruction of the posterior distribution and a Fisher matrix analysis, for a combination of cosmic shear, galaxy clustering and CMB primaries and lensing, as can be achieved by future Stage IV surveys such as Euclid and CMB-S4.

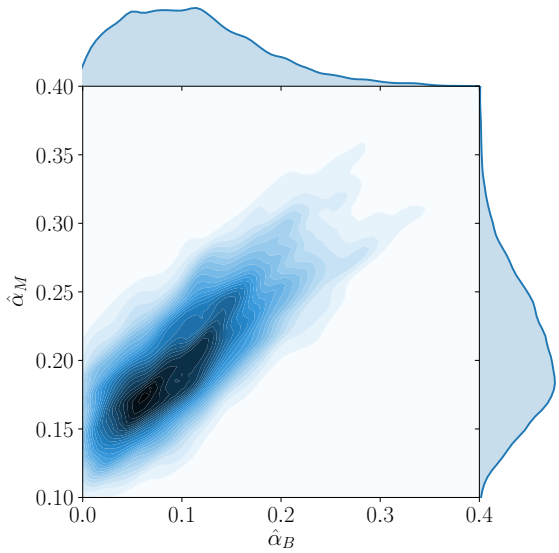

Figure 6.11.: Joint plot of the resulting posterior estimated from the MCMC for the remaining free modified gravity parameters. All other parameters have been marginalised over. Note that the contours shown do not correspond to certain σ-regions, but rather depict the full posterior distribution.

In the KiDSxGAMA analysis we have considered two different time parameterizations for the α functions: the one where these EFT functions are proportional to $\Omega_{\rm DE}$ has also been analysed in the work on future Stage IV surveys. We find that, as expected, constraints achievable from future surveys are tighter than those obtained with current data, especially considering that in the forecast analysis we had a larger parameter space including a parameterised dark energy component and a neutrino component. We estimate that future Stage IV surveys will allows us to constrain the Horndeski gravitational parameters with relative errors below 10%; this number is likely to be even reduced when cosmological surveys are considered in synergy with gravitational waves experiments, which may provide us with tight bounds on $\alpha_{\rm M}$ (Lombriser & Taylor, 2016).

A comparison between the two analyses must be carried out with care, however, due to the fact that the probes considered are not the same: in particular, the forecast analysis includes CMB primaries and lensing. On the other hand, in the KiDSxGAMA we have included a model for the intrinsic alignments contributions, while these have been neglected in our forecasts. Future work should go into the direction of extending these analyses with e.g. evaluating the KiDSxGAMA likelihood jointly with the CMB likelihood from Planck Collaboration et al. (2016a) and including an intrinsic alignment model in the forecast analysis.

Common to both analyses is the difficult interpretation of parameter space as far as the Horndeski functions are considered. Assuming proportionality to $\Omega_{\rm DE}$, in the forecasts the priors for the proportionality coefficients have been set to be strictly positive, while in the KiDSxGAMA analysis these parameters have been let free to vary over negative values as well. However, in the latter analysis we have found a very distinctive cutoff in the $\hat{\alpha}_{\rm B} - \hat{\alpha}_{\rm M}$ plane, which prevents the chains to end up in regions of parameter space where both proportionality coefficients are significantly negative. This is due to some stability checks (in particular those concerning the positivity of the speed of sound) which reject points ending up in the aforementioned regions of parameter space. These regions are, however, interesting because they may refer to models of relevance, such as $f(R)$ for which the relation $\alpha_B = -\alpha_M$ holds. While we have found that some of these stability conditions may in some cases be harmlessly bypassed, because they reject points in parameter space even when they produce a negative speed of sound *only* at very early times, the exploration of these stability conditions and their feedback

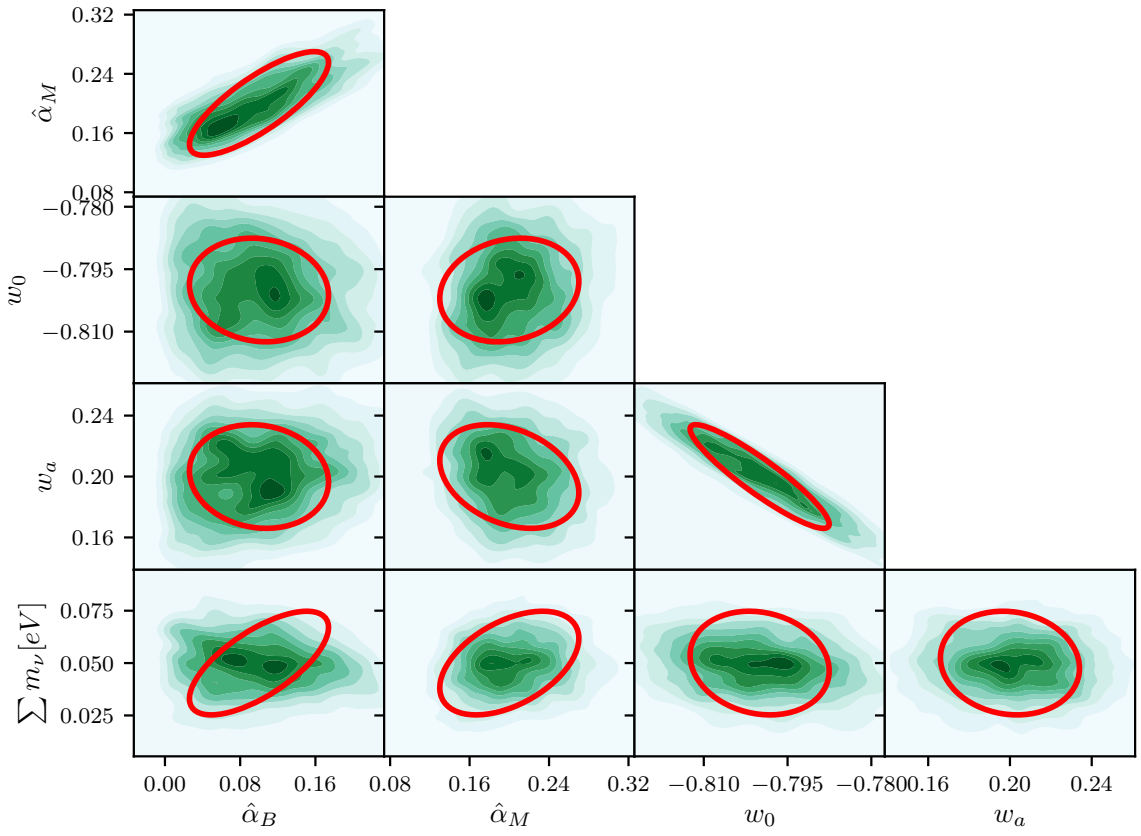

Figure 6.12.: Comparison of the MCMC results (green) with the Fisher-matrix forecast (red ellipse). The contour plot shows the probability only in the 68% region and can therefore be seen as a direct comparison with the Fisher-matrix, i.e. the outermost contour corresponds to the 1σ region. All contours are marginalised over all other parameters summarised in Tab. 6.9 where also the marginal errors are given. The experimental setup is the one described in Tab. 6.8.

on the class of valid models is an active area of research. The very recent analysis of Denissenya & Linder (2018), for example (but see also a discussion on this in Kreisch & Komatsu, 2017), has investigated the relation between parameterization and stability in Horndeski gravity, showing that the relation is not trivial. In particular, they have shown how stability evolves with redshift, picking out different regions of parameter space that can have complex structure: the final allowable stable part of parameter space is the intersection of stability for all redshifts. This can exhibit disconnected islands in parameter space and also shows significant sensitivity to the time dependent form assumed for the property functions, even for the case where only two property functions contribute, as considered in our analysis. Denissenya & Linder (2018) conclude that such sensitivity raises questions about the utility of EFT-like approaches to give robust, general conclusions about modified gravity, and that therefore simple, model independent parameterizations as binning in redshift (Casas et al., 2017) of the modified gravitational strengths (introduced in Sec. 2.6.2) can be a highly useful first step in uncovering signatures of modified gravity. The investigation of these stability conditions seems therefore a high priority issue to be addressed and had already been pointed out e.g. in the analysis of Planck Collaboration et al. (2016b).

To conclude, we stress that the contribution given by our KiDSxGAMA analysis is twofold. On the one hand we have a developed a first likelihood module for Horndeski gravity that can analyse cosmic shear data, as well as including correlations with galaxy-galaxy lensing and galaxy clustering. On the other hand, our likelihood module can be also run in a standard ΛCDM scenario and in this regard our results are in excellent agreement with those of van Uitert et al. (2018). In the future it will be interesting to carry out a similar analysis with larger data releases from the KiDS survey and with data from Stage IV surveys such as Euclid.

Chapter 7

3D cosmic shear: numerical challenges and 3D lensing random fields for cosmological inference

In the previous Chapters we highlighted the importance of cosmic shear as one of the most promising tools for current and future cosmological analyses. We also described how the spherical Fourier-Bessel decomposition of the cosmic shear field ("3D cosmic shear") is one way to maximise the amount of redshift information in a lensing analysis and therefore provides a powerful tool to investigate in particular the growth of cosmic structure, crucial for dark energy/modified gravity studies. However, we also already commented on how the computation of simulated 3D cosmic shear covariance matrices presents numerical difficulties, due to the required integrations over highly oscillatory functions.

In this Chapter we first present and compare two numerical techniques and relative implementations to perform these integrations; we then show how to generate 3D Gaussian random fields on the sky in spherical coordinates, starting from the 3D cosmic shear covariances. To validate our field-generation procedure, we calculate the Minkowski functionals associated to our random fields, compare them with the known expectation values for the Gaussian case and demonstrate the possibility of performing parameter inference through Minkowski functionals calculated from 3D cosmic shear data. This is a first step towards producing fully 3D Minkowski functionals for a lognormal field in 3D to extract Gaussian and non-Gaussian information from the cosmic shear field, as well as towards the use of Minkowski functionals as a probe of cosmology beyond the commonly used two-point statistics.

The 3D spherical Fourier-Bessel formalism presents challenging integrals to evaluate numerically: we highlight them while briefly reviewing the main equations of the 3D cosmic shear formalism, here specialised to the General Relativity case, in Sec. 7.1. Subsequently we describe two different numerical techniques used to evaluate those integrals, namely the numerical recipes underlying the results presented in Spurio Mancini et al. (2018a) and Taylor et al. (2018a,b): we present them in Secs. 7.2.1 and 7.2.2, respectively. In Sec. 7.3 we present a comparison of 3D cosmic shear covariance matrices obtained with the two numerical methods. Having two completely independent numerical techniques to tackle the 3D cosmic shear integrations, producing results in excellent agreement between them, is useful for a number of future applications that employ the simulated 3D covariance matrices. These include e.g. a cross-correlation analysis of 3D cosmic shear and galaxy clustering (see Lanusse et al., 2015, for a spherical-Bessel analysis of a spectroscopic galaxy clustering survey), or the development of a Bayesian Hierarchical Model for 3D cosmic shear power spectra estimation (see Alsing et al., 2016, 2017, for a Bayesian Hierarchical Model for tomography); at the end of the Chapter we discuss some of these future applications in Sec. 7.7. After having tested the solidity of our numerical implementation for the calculation of the 3D covariance matrices, in Sec. 7.4 we show how to make use of these matrices to generate Gaussian random fields on the sky. We explain in particular how to overcome the difficulties arising from the non-diagonality of the covariance matrices in the radial co-

ordinate, which originates from the inhomogeneity of the lensing field along the line of sight. We test the validity of our field-generation procedure in Sec. 7.5, where we briefly describe and then calculate the Minkowski Functionals of our generated Gaussian random fields, and compare them with their expectation values, known analytically in the Gaussian case. We also demonstrate how a likelihood analysis for cosmological inference can be carried out using the estimated Minkowski Functionals. Our conclusions are reported in Sec. 7.6.

Part of the content of this Chapter is based on the paper "3D cosmic shear: numerical challenges, 3D lensing random fields generation and Minkowski Functionals for cosmological inference", by A. Spurio Mancini, P. L. Taylor, R. Reischke, T. Kitching, V. Pettorino, B. M. Schäfer, B. Zieser, P. M. Merkel (Spurio Mancini et al., 2018b), submitted to PRD and available online at arXiv:1807.11461.

7.1. 3D cosmic shear

In the following Sections we will extensively make use of the 3D cosmic shear equations shown in the previous Chapter, as our first goal will be to describe two different numerical implementations that have been developed to perform the challenging integrations characterising the spherical Fourier-Bessel formalism applied to cosmic shear. Throughout the Chapter we will remain in the framework of General Relativity; our implementation is capable of producing forecasts for both a standard cosmological scenario and a modified gravity setting, but we choose the former because we would like to compare our predictions with those produced by another, completely independent implementation, whose computations however need to assume General Relativity. Therefore in the following we review the main expressions for the spherical Fourier-Bessel formalism describing cosmic shear, specialised to the General Relativity case. Since we already derived these equations in the previous Chapter in a more general non-standard gravitational scenario, here we limit ourselves to remind the reader of the final, main equations that we will need in the following: intermediate expressions, not reported here for brevity, can be derived from the more general ones for a non-standard gravitational scenario and reported in the previous Chapter, by setting $\mu = \eta = 1$.

We will be mainly concerned with the covariance of the coefficients $\bar{\gamma}_{\ell m}(k)$ in the spherical Fourier-Bessel expansion of the cosmic shear field, considering both the signal and noise contribution (Eqs. 5.11 and 5.17) which *for General Relativity* read

$$\boxed{\left\langle \bar{\gamma}_{lm}(k)\bar{\gamma}^{*}_{\ell' m'}(k')\right\rangle = \frac{9\Omega_m^2}{16\pi^4\chi_H^4}\frac{(\ell+2)!}{(\ell-2)!}\int\frac{\mathrm{d}\tilde{k}}{\tilde{k}^2}\,G_\ell(k,\tilde{k})\,G_\ell(k',\tilde{k})\,\delta^K_{\ell\ell'}\,\delta^K_{mm'}} \tag{7.1}$$

and

$$\boxed{\langle\bar{\gamma}_{\ell m}\,(k)\,\bar{\gamma}_{\ell' m'}(k')\rangle_{\mathrm{SN}} = \frac{\sigma_\epsilon^2}{2\pi^2}\int \mathrm{d}z\, n_z(z)\, j_\ell\,[k\chi_0(z)]\; j_{\ell'}\,[k'\chi_0(z)]\;\delta^K_{\ell\ell'}\delta^K_{mm'}} \tag{7.2}$$

respectively. In these formulas

$$G_\ell(k,k') = \int \mathrm{d}z\, n_z(z)\, F_\ell(z,k)\, U_\ell(z,k')\;, \tag{7.3}$$

$$F_\ell(z,k) = \int \mathrm{d}z_p\; p(z_p|z)\; j_\ell[k\chi^0(z_p)]\;, \tag{7.4}$$

$$U_\ell(z,k) = \int_0^{\chi(z)} \frac{\mathrm{d}\chi'}{a(\chi')}\frac{\chi-\chi'}{\chi\chi'}\, j_\ell(k\chi')\,\sqrt{P_\delta\,(k,z\,(\chi))} \tag{7.5}$$

and we have assumed that the intrinsic ellipticities of galaxies are uncorrelated.

We remind the reader that $\bar{\gamma}$ are estimates of the shear modes that, in addition to the pure lensing effect, keep into account the redshift distribution of the lensed galaxies $n_z(z)$ and the conditional probability $p(z_p|z)$ of estimating the redshift z_p given the true redshift z. The contribution to the total signal coming from sources situated at different distances is governed by the source density $n_z(z)$; through this term, the survey depth affects the strength of the overall signal. Angular variations are assumed to be negligible by considering a uniform source density: the number of sources per steradian and redshift interval is approximated by the mean $n_z(z)/(4\pi)$ across the sky. The influence of incomplete sky coverage is ignored in this formalism: for applications to a Fisher matrix analysis, for example, the effect of partial sky coverage can be well compensated by a multiplying factor $f_{\rm sky}$ denoting the fraction of sky spanned by the survey, prepended to the expression for the Fisher matrix (Heavens et al., 2006; Spurio Mancini et al., 2018a). A finite field of view can be incorporated in the analysis by considering a suitable window function $W(\hat{\mathbf{n}})$ that represents the angular distribution of the sources (e.g. a top hat filter corresponding to a rectangular field of view). For details on the extended formalism to include in the analysis this inhomogeneous sampling we refer the reader to e.g. Heavens (2003), while e.g. Leistedt et al. (2015) consider alternative methods, such as wavelets, to deal with survey geometry in 3D cosmic shear.

Statistical isotropy guarantees that the covariance in Eq. 7.1 does not depend on the multipole order m, while the assumed full sky coverage also prevents mixing of different ℓ modes. If the finite field of view is taken into account, statistical isotropy is broken (e.g. by the absence of data across parts of the sky) leading to a coupling of different ℓ-modes; furthermore, if the field of view is not square, even for a fixed ℓ there will be different results for different m-modes. The lensing weight function, the redshift errors and the redshift-dependence of the source distribution, instead, always introduce correlations between the amplitudes of the signal on different scales; the covariance matrix then acquires off-diagonal terms, the calculation of which is numerically involved (see Kitching et al., 2014, for how to take these into account using a pseudo-C_ℓ approach in 3D). The basis of spherical Bessel functions leads to integrals with rapidly oscillatory kernels, which in the inference process have to be solved for a large number of parameter combinations.

7.2. Numerical implementation

In this Section we will briefly describe the two methods used to calculate the correlations from Eqs. 7.1 and 7.2. While one code implements in C++ the Levin collocation method (Levin, 1996, 1997) that makes use of the periodic oscillations of the Bessel functions and has been used to produce the results of Zieser & Merkel (2016) and Spurio Mancini et al. (2018a), the other implements the integrations by matrix multiplications and appropriate use of the Limber approximation (Kaiser, 1992, 1998; Loverde & Afshordi, 2008) at high ℓ. The second method is a `Python` module, implemented in the code `GLaSS` and used in Taylor et al. (2018a,b).

7.2.1. Levin integration

The method presented in (Levin, 1996, 1997) can be used for efficient evaluation of rapidly oscillatory integrals, once certain conditions are satisfied. The main idea behind the method is to transform the quadrature problem into the solution of a system of linear ordinary differential equations. These are then tackled by collocation, i.e. choosing candidate solutions (polynomials) and a number of points in the domain (called collocation points), and selecting that solution which satisfies the given equations at the collocation points.

As seen in Sec. 7.1 the 3D cosmic shear signal and noise (cf. Eqs. 7.1 and 7.2) present several

integrals of the form

$$I_1[h] = \int_{z_1}^{z_2} \mathrm{d}z\, h(z)\, j_\ell(k\chi(z)) \tag{7.6}$$

or

$$I_2[h] = \int_{z_1}^{z_2} \mathrm{d}z\, h(z)\, j_\ell(k_1\chi(z))\, j_\ell(k_2\chi(z))\,. \tag{7.7}$$

The comoving distance between two events at redshift z_1 and z_2 is given by

$$\chi(z_1, z_2) = \int_{a(z_2)}^{a(z_1)} \frac{c\mathrm{d}a}{a^2 H(a)} = \chi_H \int_{z_1}^{z_2} \frac{\mathrm{d}z}{E(z)}\,, \tag{7.8}$$

where a is the scale factor and $\dot{a}/a = H(a) = H_0 E(a)$ is the Hubble function. Rather than redshift integrals, Eqs. 7.6 and 7.7 can be rewritten, using $\mathrm{d}z = \mathrm{d}\chi\, E[z(\chi)]$, with the comoving distance as the integration variable:

$$I_1[h] = \frac{1}{\chi_H} \int_{\chi(z_1)}^{\chi(z_2)} \mathrm{d}\chi\, E[z(\chi)]\, h[z(\chi)]\, j_\ell(k\chi)\,, \tag{7.9}$$

$$I_2[h] = \frac{1}{\chi_H} \int_{\chi(z_1)}^{\chi(z_2)} \mathrm{d}\chi\, E[z(\chi)]\, h[z(\chi)]\, j_\ell(k_1\chi) j_\ell(k_2\chi)\,. \tag{7.10}$$

Due to the highly oscillatory nature of the spherical Bessel functions, especially at high ℓ or k, the numerical solution of these integrals by standard quadrature routines is extremely inaccurate when a large number of zero-crossings occurs in the interval $[\chi(z_1), \chi(z_2)]$, unless an enormous number of points is used to sample the integrand: however, the procedure becomes then exceedingly time-consuming, especially if many combinations of ℓ and k need to be considered.

Here we describe an alternative method, presented by Levin (1996, 1997), which we use to evaluate our integrals. It is applicable to integrals of the form

$$I[F] = \int_a^b \mathrm{d}x\, F^T(x) w(x) = \int_a^b \mathrm{d}x\, \langle F, w\rangle(x)\,, \tag{7.11}$$

where $F(x) = [F_1(x), \ldots, F_d(x)]^T$ and $w(x) = [w_1(x), \ldots, w_d(x)]^T$ are vectors of functions, for which the second equality of Eq. 7.11 defines a scalar product $\langle , \rangle$ and the functions $w_i(x), i = 1, 2, ..., d$, but not $F_i(x)$, are rapidly oscillatory across the integration domain. A matrix of functions $A(x)$ is defined, such that the derivatives of $w(x)$, denoted by $w'(x)$, fulfill

$$w'(x) = A(x) w(x)\,. \tag{7.12}$$

The components $A_{iq}(x)$ should not be highly oscillatory. We show below an example of such a matrix for the particular cases given in Eqs. 7.9 and 7.10. In the Levin formalism a vector $p(x)$ is constructed to approximate the integrand in Eq. 7.11 by

$$\langle p, w\rangle' = \left\langle p' + A^T p, w \right\rangle \approx \langle F, w\rangle\,. \tag{7.13}$$

The first equality follows from applying the Leibniz rule for derivatives and Eq. 7.12, with $\langle p, Aw\rangle = \left\langle A^T p, w \right\rangle$. If such a vector is found, then the integral in Eq. 7.11 can be approximated by

$$I[F] \approx \int_a^b \mathrm{d}x\, \langle p, w\rangle'(x) = \langle p, w\rangle(b) - \langle p, w\rangle(a)\,. \tag{7.14}$$

This can be achieved by demanding that both terms should be equal, $\langle p, w\rangle' = \langle F, w\rangle$, at n collocation points x_j, $j = 1, 2, ..., n$. The requirement

$$\left\langle p' + A^T p - F, w \right\rangle (x_j) = 0, \quad j = 1, ..., n \tag{7.15}$$

generally means that the vector $\left\langle p' + A^T p - F \right\rangle$ must be orthogonal to w at the points x_j, for example by demanding that it should be the null vector:

$$p'(x_j) + A^T(x_j)p(x_j) = F(x_j). \tag{7.16}$$

Finding a vector p which has this property can be achieved by choosing a set of n linearly independent and differentiable basis functions $u_m(x)$ and writing each component $p_i(x)$ as a linear combination:

$$p_i(x) = c_i^{(m)} u_m(x), \quad i = 1, ..., d; \quad m = 1, ..., n\ . \tag{7.17}$$

Equation 7.16 then leads to the following linear system of equations for the $d \times n$ coefficients $c_i^{(m)}$:

$$c_i^{(m)} u'_m(x_j) + A_{qi} c_q^{(m)} u_m(x_j) = F_i(x_j), \quad i, q = 1, ..., d; j, m = 1, ..., n\ . \tag{7.18}$$

Levin (1996) showed how to concretely apply this algorithm to several cases of integrals with highly oscillatory kernels. The performance varies depending on the integrand, but accuracies below 10^{-6} can often be achieved with less than 10 collocation points. As suggested by Levin (1996), in our implementation we use equidistant collocation points

$$x_j = a + (j-1)\frac{b-a}{n-1}, \quad j = 1, ..., n \tag{7.19}$$

and choose the n lowest-order polynomials as basis functions:

$$u_m(x) = \left(\frac{x - \frac{a+b}{2}}{b-a}\right)^{m-1}, \quad m = 1, ..., n\ . \tag{7.20}$$

We note that the polynomials with $m > 1$ and the derivatives with $m > 2$ share the root $x = (a+b)/2$: to prevent the linear system of equations from becoming singular, that root should not be used as a collocation point. The factor $1/(b-a)$ is included for numerical reasons: if $b \gg 1$ or $b \ll 1$, the values of polynomials of different order may differ by several orders of magnitude; the normalising factor guarantees that $|u_m(x)| \leq 1$ across the integration domain, in order to regulate the range of the coefficients of the linear system of equations in Eq. 7.18 and thus the condition of the corresponding matrix. Suitable vectors w for the integrals in Eqs. 7.9 and 7.10 can be identified by considering the following recurrence relations for the spherical Bessel functions (Abramowitz et al., 1988):

$$\frac{\mathrm{d}}{\mathrm{d}x} j_\ell(x) = j_{\ell-1}(x) - \frac{\ell+1}{x} j_\ell(x)\ , \tag{7.21}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} j_{\ell-1}(x) = -j_\ell(x) + \frac{\ell-1}{x} j_{\ell-1}(x)\ . \tag{7.22}$$

Rewriting these relations in the form $w' = Aw$, one finds that

$$w(\chi) = \begin{pmatrix} j_\ell(k\chi) \\ j_{\ell-1}(k\chi) \end{pmatrix}, \quad A(\chi) = \begin{pmatrix} -\frac{\ell+1}{\chi} & k \\ -k & \frac{\ell-1}{\chi} \end{pmatrix} \tag{7.23}$$

is a suitable choice for the integral in Eq. 7.9, with $F(\chi) = \{E[z(\chi)]h[z(\chi)], 0\}^T$. It is easy to verify that neither the entries of the matrix A nor the integral kernels F are rapidly oscillatory. For integrals of the type in Eq. 7.10, four-dimensional vectors are needed:

$$w(\chi) = \begin{pmatrix} j_\ell(k_1\chi) j_\ell(k_2\chi) \\ j_{\ell-1}(k_1\chi) j_\ell(k_2\chi) \\ j_\ell(k_1\chi) j_{\ell-1}(k_2\chi) \\ j_{\ell-1}(k_1\chi) j_{\ell-1}(k_2\chi) \end{pmatrix}, \quad A(\chi) = \begin{pmatrix} -\frac{2(l+1)}{\chi} & k_1 & k_2 & 0 \\ -k_1 & -\frac{2}{\chi} & 0 & k_2 \\ -k_2 & 0 & -\frac{3}{\chi} & k_1 \\ 0 & -k_2 & -k_1 & \frac{2(\ell-1)}{\chi} \end{pmatrix}. \tag{7.24}$$

Similarly, $F(\chi) = \{E[z(\chi)]h[z(\chi)], 0, 0, 0\}^T$.

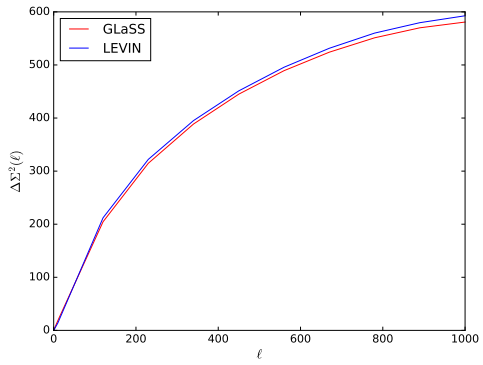

Figure 7.1.: Comparison of the differential signal-to-noise curve (Eq. 7.30) as a function of the angular multipole. The two curves have been obtained from the signal and noise parts of the covariance matrices produced with `GLaSS` (*red*) and the Levin method (*blue*).

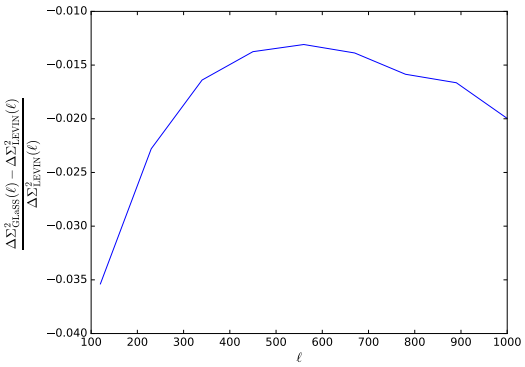

Figure 7.2.: Relative difference of the signal-to-noise curve calculated with the `GLaSS` and Levin method, as a function of the multipole ℓ.

7.2.2. *GLaSS*

The Generalised Lensing and Shear Spectra (`GLaSS`) code is written in `Python` and integrated into the modular cosmological package `Cosmosis` (Zuntz et al., 2015). Cosmological information can be read from an external source as in this work, or directly from the `Cosmosis` pipeline. More information can be found in Taylor et al. (2018b). `GLaSS` is written to compute the lensing spectra for an arbitrary weight function $W_\ell[k\chi^0(z_p)]$ which takes the place of the Bessel functions in Eq. 7.4; see Taylor et al. (2018a) for more details about this *generalized spherical-transform*. Nevertheless, 3D cosmic shear comes as an in-built run-mode option.

All nested integrals in Eqs. 7.1-7.5 are computed as matrix multiplications because this is one of the few operations that releases the Global Interpreter Lock in `Python` allowing parallelisation. For example,

$$U_\ell(z,k) \approx \sum_{\chi'} A(\chi(z),\chi')\, B(\chi',k)\,, \tag{7.25}$$

$A(\chi,\chi') \equiv \Delta\chi' \frac{F_K(\chi,\chi')}{a(\chi')}$, where $\Delta\chi'$ is the spacing between the sampled points in χ' and $B(\chi,\chi') \equiv j_\ell(k\chi')\,\sqrt{P(k;\chi')}$.

To further speed up computations all Bessel functions data is pre-computed in `GLaSS`. To save memory, values of the Bessel functions $j_\ell(x)$ are stored in a 2D look up table in ℓ and x and the $j_\ell(k\chi)$ are found as needed. This procedure was first described in Seljak & Zaldarriaga (1996); Kosowsky (1998).

Ω_m	Ω_b	Ω_r	Ω_k	w_0	w_a	σ_8	n_s	h
0.315	0.0486	9.187×10^{-5}	0.0	-1.0	0.0	0.834	0.962	0.674

Table 7.1.: Values of the cosmological parameters in the fiducial model assumed for the code comparison.

While computing the lensing spectra in terms of nested matrix multiplications allows for easy parallelisation, this procedure does not efficiently sample the z-k space as efficiently as the Levin integration. At high-ℓ where the Bessel functions oscillate quickly this means the lensing spectra must be evaluated at very high resolutions. To reduce the resolution at which the lensing spectra must be evaluated, `GLaSS` takes the extended Limber approximation (Loverde & Afshordi, 2008) above $\ell > 100$. This was shown to have negligible impact for stage IV surveys (Kitching et al., 2017). Taking the Limber approximation, equation 7.5 can be rewritten as:

$$U_\ell(\chi, k) = \frac{F_k(\chi, \nu(k))}{ka(\nu(k))} \sqrt{\frac{\pi}{2(\ell + 1/2)}} \sqrt{P(k, \nu(k))}\,, \tag{7.26}$$

where $\nu(k) \equiv \frac{\ell+1/2}{k}$. Meanwhile at low-$\ell$ the Bessel functions oscillate slowly and the nested integrals can be evaluated at lower resolution.

7.3. Code comparison

In the following we compare the predictions for the 3D cosmic shear covariance matrices produced with the Levin method and with the algorithm implemented in the `GLaSS` code. For the code comparison we fix the fiducial cosmological model to a flat cosmology with parameters given in Tab. 7.1. The source distribution and the redshift error probability need to be the same for the two codes. For the source distribution we follow Amendola et al. (2016) and choose

$$n_z(z) \propto n_0 \left(\frac{\sqrt{2}}{z_m}\right)^3 z^2 \exp\left[-\left(\frac{\sqrt{2}z}{z_m}\right)^{3/2}\right], \tag{7.27}$$

where z_m is the median redshift of the survey and n_0 is the observed redshift-integrated source density. We set $z_m = 0.9, n_0 = 30\,\mathrm{arcmin}^{-2}$. We take the redshift error distribution to be a Gaussian

$$p(z_p|z) = \frac{1}{\sqrt{2\pi}\sigma(z)} \exp\left[-\frac{(z_p - z)^2}{2\sigma^2(z)}\right], \tag{7.28}$$

with a redshift-dependent dispersion

$$\sigma(z) = \sigma_z(1 + z)\,. \tag{7.29}$$

We first compare the signal-to-noise curve. The cumulative signal-to-noise ratio, summed over the contributions at different multipoles up to a maximum multipole ℓ, is defined as

$$\Sigma^2(\leq \ell) = f_\mathrm{sky} \sum_{\ell'=\ell_\mathrm{min}}^{\ell} \frac{2\ell' + 1}{2} \mathrm{Tr}\left[\boldsymbol{C}_{\ell'}^{-1} \boldsymbol{S}_{\ell'} \boldsymbol{C}_{\ell'}^{-1} \boldsymbol{S}_{\ell'}\right] \equiv \sum_{\ell'=\ell_\mathrm{min}}^{\ell} \Delta\Sigma^2(\ell')\,, \tag{7.30}$$

where $\boldsymbol{S}$ is the signal covariance (Eq. 7.1) only, while $\boldsymbol{C}$ refers to the sum of signal and shot noise, i.e. Eqs. 7.1+7.2. The signal-to-noise curves produced by both codes are shown in Fig.7.1, depicting the differential contributions to the signal-to-noise coming from the different multipoles. The Levin and `GLaSS` predictions for the signal-to-noise curves show agreement with each other for the multipoles considered, reaching differences below 4%, as evidenced by Fig. 7.2 where the relative difference between the predictions of the two codes are shown. `GLaSS` has slightly lower signal-to-noise at high

ℓ. This is because GLaSS is not designed specifically for 3D cosmic shear and the signal-to-noise converges as the resolution of the computation grid is increased. This problem will be exacerbated as one includes higher and higher ℓ (e.g. $\ell = 3000$ where there is still useful signal to add) because the Bessel functions oscillate more quickly. For our comparison we used 2000 k-modes linearly spaced between $k = 0.005\,h/\mathrm{Mpc}$ and $k = 2.0\,h/\mathrm{Mpc}$, and restricted the comparison to multipoles $\ell \leq 1000$; see Taylor et al. (2018b) for details on how much information is captured by GLaSS at different k-resolutions and Taylor et al. (2018a) for a discussion of the run time at different resolutions.

As a second diagnostic for our comparison, we consider individually the signal and noise contributions to the covariance matrices (Eqs. 7.1 and 7.2, respectively) for two different multipoles, $\ell = 100$ and $\ell = 500$. For both signal and noise we compare the elements on the diagonal $C_\ell(k, k)$, and plot them respectively in Fig. 7.3 and Fig. 7.4. In the noise case we also multiply the curves by k^2, to check that they effectively become flat as expected. The predictions show good agreement, with differences of at most a few percent (in the lower k range for the signal, and over the entire k range for the noise), as visible also from Fig. 7.5, where we show the differences between the codes, normalised to the sum of their predictions. The disagreement in the signal plot towards the higher end of the k range is due to the numerical noise present in the GLaSS computations; however, this discrepancy can be disregarded because the contributions from those k-regimes ($k \gtrsim 0.2h/\mathrm{Mpc}$ for $\ell = 100$, $k \gtrsim 0.4h/\mathrm{Mpc}$ for $\ell = 500$) are many orders of magnitude smaller than the main contributions around the peak of the curves, and also much smaller than contributions from the noise (cf. Fig 7.4). For $\ell = 100$, the Levin and GLaSS predictions coincide until approximately $k \simeq 0.2h/\mathrm{Mpc}$: at this point the behaviour of the curve for GLaSS starts being dominated by numerical noise, while the Levin signal decreases in a smoother way. The same happens for $\ell = 500$, but the disagreement starts at approximately $k \simeq 0.4\,h/\mathrm{Mpc}$. In both cases however, the signal predictions in those k-regimes are at least 3-4 orders of magnitude smaller than the contributions around the peak of the curves, just before and after approximately $k \simeq 0.1h/\mathrm{Mpc}$, respectively. Importantly, the values of the signal curves for those k-regimes are even smaller than the contributions from the noise, which dominates in that regime by many orders of magnitude. This means that for practical purposes we can safely ignore the contributions from those k-regimes where the codes are apparently in disagreement in their signal predictions. In Figs. 7.3,7.4, and the left panels of Fig. 7.5 we demonstrate this point by shading the regions where the signal contribution represents a fraction $\leq 1/1000$ of the noise contribution at the same k. These regions turn out to be the same where the signal predictions of the two codes disagree, thus demonstrating that this discrepancy can be safely disregarded. In the bottom panel of Fig. 7.3 we plot the same comparison between the signal predictions produced by both two codes, with a linear scale on the y-axis instead of the logarithmic one used in the top panel; this is another way to appreciate how subdominant the contributions coming from the higher end of the k-range are with respect to the signal coming from the lower k-range.

It is interesting to note that the disagreement is practically only evident in the signal predictions, while the noise part is much less affected. This may be due to the increased number of matrix multiplications that need to be performed in the calculation of the signal with respect to the noise (cf. Eqs. 7.1 and 7.2). The fact that the number of integrations to carry out for the noise is higher means, in the GLaSS implementation, that more matrix multiplications are required and these are sensitive to the resolution in k. Additionally, since in the noise part of the covariance matrix there are no multiplications by Bessel functions, this may suggest that the spikes at high-k in the signal may also be due to the Bessel function resolution breaking down (as explained in Sec. 7.2.2, in GLaSS the Bessel functions $j_\ell(x)$ are precomputed in a look up table in ℓ and x). The code implementing the Levin method sources the matter power spectrum from the Einstein-Boltzmann solver Cosmic Linear Anisotropy Solving System (CLASS, Lesgourgues, 2011), while for this code comparison the matter power spectrum used by GLaSS has been sourced from the Code for Anisotropies in the Microwave Background (CAMB, Lewis et al., 2000). The CLASS and CAMB codes have been compared in their predictions (Lesgourgues, 2011). Therefore, in comparing the Levin and GLaSS methods, the matter power

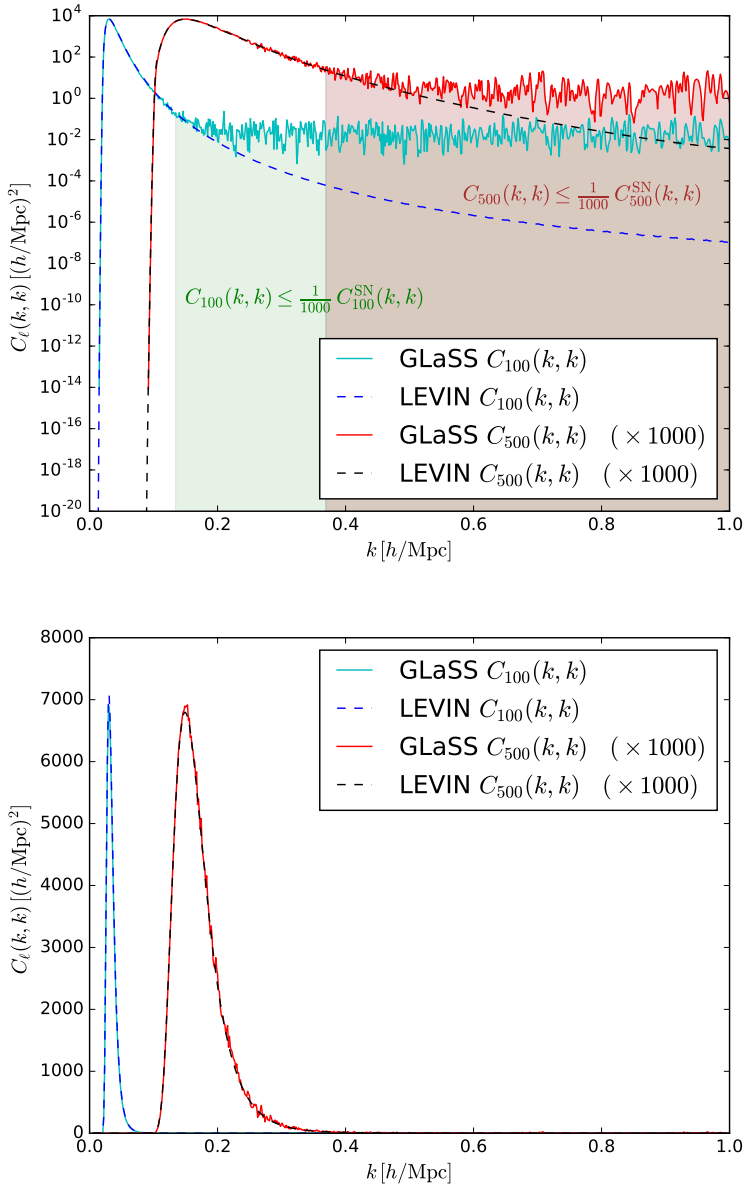

Figure 7.3.: Comparison of the diagonal elements of the signal part of the covariance matrices (Eq. 7.1) for two multipoles $\ell = 100$ and $\ell = 500$, produced with `GLaSS` (*solid* lines, *cyan* and *red* for $\ell = 100$ and $\ell = 500$, respectively) and the Levin method (*dashed* lines, *blue* and *black* for $\ell = 100$ and $\ell = 500$, respectively). All curves have been plotted without performing any interpolation. We show the same curves using a linear (*upper* panel) and a logarithmic (*bottom* panel) scale on the y-axis. The differences at higher k ($k \gtrsim 0.2\,h/\mathrm{Mpc}$ for $\ell = 100$, $k \gtrsim 0.4\,h/\mathrm{Mpc}$ for $\ell = 500$) arise from the higher numerical noise present in the `GLaSS` computations in that k regime. However, these contributions are many orders of magnitude smaller than the main contributions around the peaks of the curves, and much smaller than the contributions from the noise (cf. Fig. 7.4), therefore can be safely neglected. We demonstrate this point in the upper panel by indicating the shaded region for each multipole ℓ where the signal represents a fraction $\leq 1/1000$ of the noise: these regions correspond to the k-ranges where the `GLaSS` and Levin predictions for the signal are in apparent disagreement (cf. also Fig. 7.5). In both panels the curves for $\ell = 500$ have been multiplied by a factor 1000 for easier visualisation.

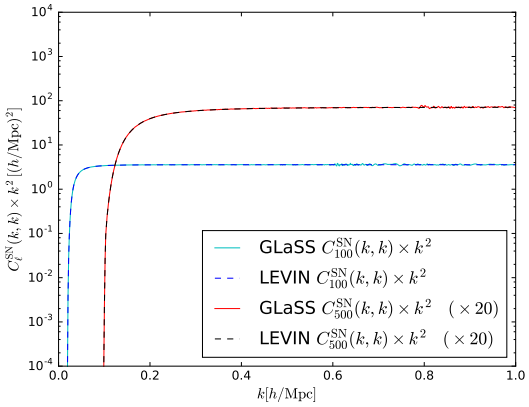

Figure 7.4.: Comparison of the diagonal elements of the noise part of the covariance matrices (Eq. 7.2) for the same two multipoles $\ell = 100$ and $\ell = 500$ of Fig. 7.3, produced with GLaSS (*solid*, *cyan* for $\ell = 100$ and *red* for $\ell = 500$) and the Levin method (*dashed*, *blue* for $\ell = 100$ and *black* for $\ell = 500$). All curves have been plotted without performing any interpolation and multiplied by a factor k^2. In the case $\ell = 500$ the curves produced by both methods have also been multiplied by a factor 20 for easier visualisation.

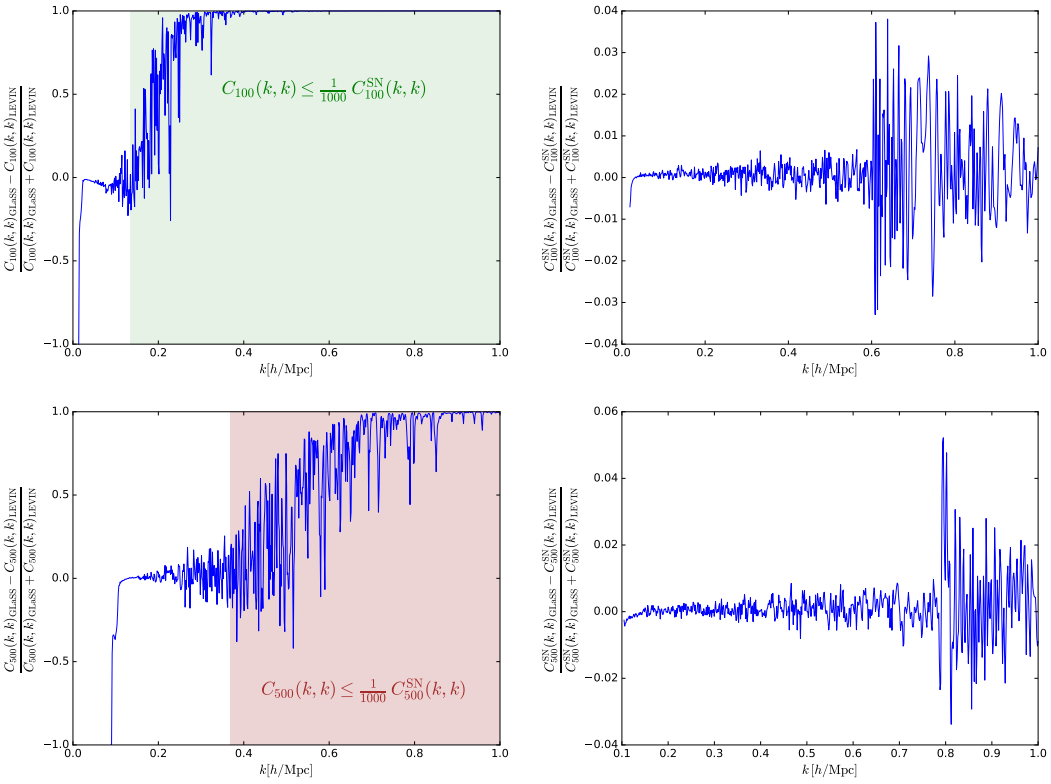

Figure 7.5.: Differences between the predictions for the signal (*left* panels) and noise (*right* panels) contributions to the covariance matrices for multipoles $\ell = 100$ (*top* panels) and $\ell = 500$ (*bottom* panels), normalised to their sum. We stress here again that the discrepancies at high k should not be a concern because the k-regimes where they originate produce contributions very much subdominant with respect to the peaks of the signal curves, and also with respect to the relevant contributions from the noise (cf. Figs. 7.3, 7.4). In the signal plots we shade the regions where the signal is a fraction $\le 1/1000$ of the noise (cf. Fig. 7.3): these regions correspond tho the k values where the differences between the two codes are bigger, however since the signal contributions from these regions are negligeable, this discrepancy can be safely ignored.

spectrum has been ruled out as a possible source of discrepancy.

We conclude this Section with a note on the performance of the two codes. The code implementing the Levin integration has been developed explicitly for the production of precise 3D cosmic shear cosmological forecasts and has been recently used to this purpose in Spurio Mancini et al. (2018a). As one can see from Figs. 7.3 and 7.4, the curves produced with the Levin method are very smooth, showing the high precision achieved by the method. This compensates for the relatively low speed of the code, necessary to achieve that precision. `GLaSS` on the other hand, has not been developed for 3D cosmic shear only; in Taylor et al. (2018a,b) it is introduced as a means to compute lensing spectra for arbitrary weighting functions and, importantly, for integration within the cosmological module `Cosmosis`. This means that speed has been a crucial goal in developing the code and the method used for the matrix multiplications indeed allows for greater speed than the one achieved with the Levin method. However, numerical noise remains higher: to overcome this issue, one would need to increase the resolution at which the matrix multiplications are performed, but this would inevitably imply a slower performance of the code. We conclude that the use of the Levin or the `GLaSS` method depends on the task to perform: if a high level of precision is required, the Levin method should be preferred, while if speed is a crucial requirement, `GLaSS` can be a better option. For our purposes in this paper, i.e. the demonstration of a method for generating 3D lensing random fields on the sky and the calculation of Minkowski Functionals associated to these fields, both methods are equally valid for the computation of the 3D cosmic shear covariance matrices, which represent the starting point of the algorithms described in the following sections.

7.4. Generation of spin-2 random fields on the sky

In this section we show how to generate random fields on the sky starting from the full 3D cosmic shear covariance matrix. As shown in Sec. 7.1, the full covariance matrix can be decomposed in $C_\ell(k,k')$ for each multipole ℓ, given that the assumed isotropy of the shear field implies multipole independence $\langle\gamma_{\ell m}(k)\gamma_{\ell' m'}(k')\rangle = C_\ell(k,k')\delta_{\ell\ell'}\delta_{mm'}$. We detail our procedure considering for simplicity the convergence κ, as it is a scalar field and therefore easier to analyse. The convergence shares essentially the same covariance matrix with the shear field, each ℓ-block only being rescaled by a prefactor $\frac{\ell(\ell+1)}{(\ell+2)(\ell-1)}$ (Castro et al., 2005) that plays a role only for the very largest angular scales (cf. Sec. 4.2.4). The generalisation to the spin-2 case for the shear field simply requires starting from the original shear covariance matrix and replacing the transforms from Fourier coefficients to configuration space with their spin-2 extensions. To demonstrate our field generation procedure we use the covariance matrices produced with the Levin integration.

Our aim is to generate modes of the convergence field in Fourier space $\kappa_{\ell m}$ and to transform them back into configuration space using the HEALPix (Górski et al., 2005) in-built function `alm2map`, in its scalar version for the convergence case (for the shear, one simply needs to activate the option `pol = True` that allows the user to deal with spin-2 fields). This way we can obtain samples of the convergence field in configuration space $\kappa(r,\theta,\phi)$, on spherical shells corresponding to different values of the radius; on each shell, the field can be discretised on a HEALPix map. An example of 3D reconstruction on 3 slices in redshift or, equivalently, comoving distance is shown in Fig. 7.6.

The procedure described in the following is similar to the one used in the code FLASK (Xavier et al., 2016) to generate samples of the density, convergence and shear fields on redshift slices, starting from tomographic weak lensing covariance matrices $C_{ij}(\ell)$, where the indices i and j run over the redshift slices and the type of field (density, convergence or shear). In FLASK, the problem of generating correlated random fields across different redshift slices is dealt with by means of a Cholesky decomposition of the correlated covariance matrices $C_{ij}(\ell)$. The Cholesky decomposition rewrites the covariance as the product of an upper and lower triangular matrix (see Appendix A.4 for mathematical details on using a Cholesky decomposition to generate correlated random fields). Here, the situation

is similar in that we also have correlated multipoles belonging to the different radial slices, however the correlation is in terms of the wavector k rather than the tomographic/field index i. This difference originates from the fact that we start from the 3D cosmic shear covariance matrices $C_\ell(k, k')$, as opposed to the tomographic $C_{ij}(\ell)$ matrices in FLASK. Additionally, in FLASK correlations between density, convergence and shear fields can be considered if the user desires, while here we concentrate singularly on the generation of convergence or shear fields and do not consider their cross-correlations. The fact that the random fields at different wavectors are correlated is ultimately due to the fact that the lensing field is not homogenenous along the line of sight, due to the mode-coupling effect of the lensing kernel, the source redshift distributions and the redshift error probability (cf. Eqs. 7.1-7.5).

The assumption of statistical isotropy implies that modes $\kappa_{\ell m}$ of the convergence field at different multipoles ℓ and m can be generated independently. The number of ℓ multipoles is in principle infinite, however practically there will be a $\ell_{\rm max}$ which sets the maximum resolution. We use $\ell_{\rm max} = 3N_{\rm side}$, where $N_{\rm side}$ is a HEALPix parameter describing the resolution of the HEALPix grid (Górski et al., 2005). The choice $\ell_{\rm max} = 3N_{\rm side}$ is the same made by Lim & Simon (2012) in their CMB analysis and guarantees that the grid size is comparable to the smallest angular scale considered, corresponding to $\ell_{\rm max}$. For each ℓ value, m ranges from $-\ell$ to $+\ell$, so that there are $2\ell + 1$ m values for each multipole ℓ. However, due to the hermiticity of the convergence field, we actually consider only $\ell + 1$ modes from 0 to ℓ. We employ a Cholesky decomposition of the covariance matrices to deal with the fact that modes corresponding to different k values are correlated:

$$C_\ell(k, k') = \sum_p T_\ell(k, p) T_\ell(p, k'), \tag{7.31}$$

where $\mathbf{T}(\ell)$ are (lower) triangular matrices, which we can later use to generate correlated random variables $\kappa_{\ell m}(k)$, e.g. Gaussian distributed,

$$\kappa_{\ell m}(k) = \sum_p T_\ell(k, p)\, n_{\ell m}(k), \tag{7.32}$$

where $n_{\ell m}(k)$ are independent, Gaussian distributed random variables with zero mean and unit variance (see Appendix A.4 for a motivation for this method as a means to generate correlated random fields). To obtain our convergence field samples in configuration space, we transform back from Fourier space, first by multiplying by a spherical Bessel function and k^2 and integrating over k, as indicated by Eq. 4.41 which defines the spherical Fourier-Bessel transform and its inverse, and then acting with the HEALPix routine `alm2map` to obtain the field samples on a discretised grid in the angular coordinates.

We summarise schematically our procedure in Algorithm 4. We implemented it in a Python routine, leveraging parallelisation on multiple cores with `joblib`. The problem is embarassingly parallel, since the correlation of the fields on different radii is preserved by the starting cosmic shear covariance matrix, while different realisations of the random fields are completely independent. The fact that the covariance does not depend on m, but only on the multipole ℓ, can be used to speed up calculations, as one needs to perform the Cholesky decomposition only once per each multipole ℓ, and can then use the decomposition for all m's pertaining to that ℓ mode. In Fig. 7.10 we report a study of the time scaling with $N_{\rm side}$ of the full procedure including the random fields generation process and the estimation of the Minkowski Functionals, described in next Section.

7.5. Minkowski Functionals of scalar fields on the sphere

In this Section we briefly introduce the notion of Minkowski Functionals (MFs) and apply them to the generated random fields introduced in Sec. 7.4. For Gaussian random fields the MFs can be calculated analytically. We will compare these theoretical predictions with the MFs calculated directly from the HEALPix maps as a proof of concept. In particular we will calculate the MFs on spheres of different

Algorithm 4 Algorithm for generation of lensing Gaussian random fields on spherical shells

input Covariance matrix (e.g. for the convergence) $\langle \kappa_{\ell m} \kappa_{\ell' m'} \rangle = C_\ell(k, k') \delta_{\ell \ell'} \delta_{m m'}$

output $\kappa(r, \theta, \phi)$. For each fixed radius r in $r_1, \ldots r_{N_\chi}$, create a HEALPix map on discretised θ and ϕ

method $\forall r \in [r_1, \ldots r_{N_\chi}]$:

$\forall \ell \in [0, \ell_{\rm max}]$:

Cholesky decompose $\boldsymbol{C}_\ell = \boldsymbol{T}_\ell \boldsymbol{T}_\ell^T$

$\forall m \in [0, \ell]$:

sample $z \sim \mathrm{N}(0, I)$

$\kappa_{\ell m}(k) = \boldsymbol{T}_\ell z$

$\kappa_{\ell m}(k) \rightarrow \int \mathrm{d}k\, k^2\, j_\ell(kr) \rightarrow \kappa_{\ell m}(r)$

$\kappa_{\ell m}(r) \rightarrow$ HEALPix `alm2map` $\rightarrow \kappa(r, \theta, \phi)$

radii (cf. Fig. 7.6) and estimate the covariance between the different MFs at those radii. Repeating this whole procedure for different starting lensing covariance matrices (e.g varying each time one cosmological parameter), we can then produce a likelihood function dependent on the underlying cosmology.

7.5.1. Mathematical background on Minkowski Functionals

Here we define the MFs, concentrating on the aspects that are more interesting for cosmological applications and referring the reader to e.g. Mecke et al. (1994) for further mathematical details. In our definitions we follow the notation of Schmalzing & Gorski (1998) and Lim & Simon (2012).

MFs are integral morphological measures that characterise the geometry and topology of spatial patterns; in particular, they characterize the morphological properties of convex, compact sets in a d dimensional space. A morphological property must be invariant under translation and rotations. Hadwiger's theorem (Hadwiger, 1957) states that on a d-dimensional convex ring embedded in a d dimensional space there exist $d + 1$ linearly independent morphological functionals. On the 2-sphere, $\mathbb{S}_2$, there are $2 + 1 = 3$ MFs which, up to normalization, have clear geometrical interpretations. Let us expand these ideas, starting from an introduction to the role of MFs in integral geometry.

Integral geometry and properties of Minkowski Functionals

Let us first introduce integral geometry in flat space, or, to be more precise, in a d–dimensional Euclidean space E^d. We wish to characterize the morphology of a suitable set $Q \subseteq E^d$. A possible starting point for integral geometry (Santaló, 1976) is some manifold with a group of transformations $\mathcal{G}$. Usually we are dealing with the d-dimensional Euclidean space, for which the natural choice is the group that contains as subgroups rotations and translations. One can then consider the set $\mathcal{K}$ of convex bodies embedded in this space and, as an extension, the so–called convex ring R of all finite unions of convex bodies. In order to characterize a body B from the convex ring one looks for scalar functionals M that satisfy the following requirements (Schmalzing et al., 1996):

- *translation invariance*: the functional should be independent of the body's position and orientation in space, $M(gB) = M(B)$ for any $g \in \mathcal{G}$, $B \in R$;
- *additivity*: When two bodies B_1 and B_2 are united, the functional of the union is $M(B_1 \cup B_2) = M(B_1) + M(B_2) - M(B_1 \cap B_2)$ for any $B_1, B_2 \in R$;
- *conditional continuity*: The functionals of convex approximations to a convex body converge to the functionals of the body, $M(K_i) \rightarrow M(K)$ as $K_i \rightarrow K$ for K, K_i in $\mathcal{K}$.

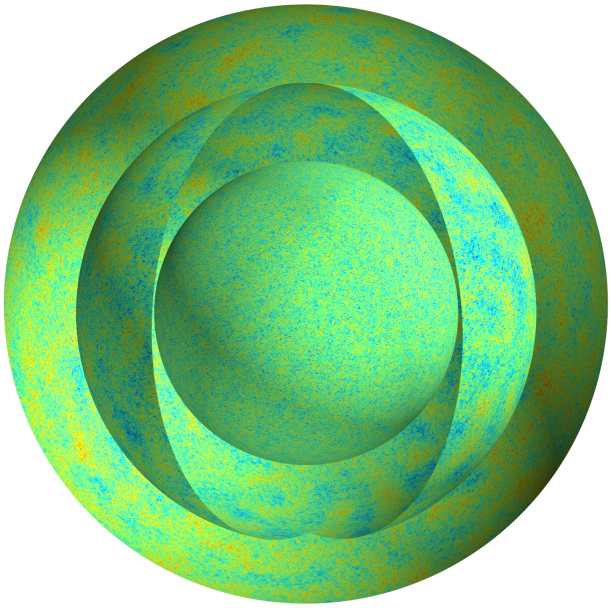

Figure 7.6.: Convergence field sampled at three different values of the radius χ, with the observer situated in the centre. A section of the outer and middle sphere has been removed to facilitate visualisation. The lensing covariance matrix which we used for sampling the random field is given by Eq.7.1. We consider only contributions from the signal part of the covariance matrix, and use 30 ℓ modes ranging between 10 and 1000. We use a linear matter power spectrum for the calculation of the covariance, since the condition $\ell \leq 1000$ ensures that we are less sensitive to the highly non-linear k-regime (cf. also Fig.7.3).

While we might be led to think that these fairly general requirements leave a vast choice of such functionals, Hadwiger's theorem states that under a few simple requirements there are only $d + 1$ independent such functionals if the space is d-dimensional: these are the so–called Minkowski Functionals V_j, with j ranging from 0 to d.

If the set Q has a smooth boundary ∂Q, its Minkowski Functionals – except for the d–dimensional volume V_0, which is calculated by volume integration – are given by simple surface integrals (Schneider, 1993). We use ω_j to denote the surface area of the j–dimensional unit sphere. Some special values are $\omega_0 = 2$, $\omega_1 = 2\pi$, $\omega_2 = 4\pi$, while in general

$$\omega_j = \frac{2\pi^{(j+1)/2}}{\Gamma((j+1)/2)}. \tag{7.33}$$

MFs are formally defined as:

$$\begin{aligned} V_0(Q) &= \int_Q \mathrm{d}v, \\ V_j(Q) &= \frac{1}{\omega_{j-1}\binom{d}{j}} \int_{\partial Q} \mathrm{d}s \mathfrak{S}_j(\kappa_1 \ldots \kappa_{d-1}). \end{aligned} \tag{7.34}$$

Here $\mathrm{d}v$ and $\mathrm{d}s$ denote the volume element in $\mathbb{E}^d$ and the surface element on Q, respectively, κ_1 to κ_{d-1} are the boundary's $d-1$ principal curvatures, and $\mathfrak{S}_j$ is the j-th elementary symmetric function defined by the polynomial expansion

$$\prod_{i=1}^{d-1}(x + \kappa_i) = \sum_{j=1}^{d} x^{d-j} \mathfrak{S}_j(\kappa_1 \ldots \kappa_{d-1}); \tag{7.35}$$

d	1	2	3
V_0	length	area	volume
V_1	χ	circumference	surface area
V_2	–	χ	total mean curvature
V_3	–	–	χ

Table 7.2.: Some of the $d+1$ Minkowski Functionals in d–dimensional Euclidean space may be interpreted as familiar geometric quantities (apart from numerical factors). This table summarizes the geometric interpretations of all Minkowski Functionals for one, two and three dimensions. The symbol χ denotes the connectivity, a purely topological quantity that describes a topological space's shape or structure regardless of the way it is bent.

hence $\mathfrak{S}_1 = 1$, $\mathfrak{S}_2 = \kappa_1 + \ldots + \kappa_{d-1}$, and so on up to $\mathfrak{S}_d = \kappa_1 \ldots \kappa_{d-1}$. Table 7.2 summarizes the geometric interpretations of the Minkowski functionals in one, two and three dimensions. In three-dimensional Euclidean space, in particular, the first functional equals the volume V of the body, the second one is the surface area A. The third functional corresponds to the integral mean curvature of the body's surface and provides information about the shape. Lastly, the fourth functional can be interpreted as the Euler characteristic χ which is a purely topological quantity that can be calculated using the simple formula

$$\chi = \text{number of components} - \text{number of tunnels} + \text{number of cavities}. \tag{7.36}$$

Spaces of constant curvature

Let us now consider the d–dimensional space of constant curvature kK. The sign k equals $+1$, 0 or -1, for the spherical space $\mathbb{S}^d$, the Euclidean space $\mathbb{E}^d$ and the hyperbolic space $\mathbb{H}^d$, respectively. K is a positive constant of dimension [Length]$^{-2}$, hence its inverse square root $K^{-1/2}$ can be interpreted as the radius of curvature. Santaló (1976) shows how to obtain an integral geometry on such spaces. Curvature integrals as in Eq. (7.34) can still be defined, provided that care is taken in using the geodesic curvatures κ_i. If we do so, we can consider these quantities the Minkowski Functionals in curved spaces.

Some of the geometric interpretations are altered with respect to the flat case. While in flat space the curvature integral $V_d(Q)$ is equal to the Euler characteristic $\chi(Q)$, curved spaces require a generalized Gauss–Bonnet Theorem proved for arbitrary Riemannian manifolds by Chern (1944). The theorem states that the Euler characteristic is a linear combination of all Minkowski functionals as defined by Eq. (7.34),

$$\chi(Q) = \sum_{j=0}^{d} c_j V_j(Q), \tag{7.37}$$

with the coefficients c_j given by

$$c_j = \begin{cases} \binom{d}{j} \dfrac{2(kK)^{(d-j)/2}}{\omega_{d-j}} & \text{if } d-j \text{ even}, \\ 0 & \text{if } d-j \text{ odd}. \end{cases} \tag{7.38}$$

Note that from the point of view of Hadwiger's theorem, which is also valid on curved spaces, all linear combinations of Minkowski functionals are equally suitable as morphological descriptors, so one may both use the integrated curvature V_d and the Euler characteristic χ as the last Minkowski functional. [14] In the following, we will consider V_d, having in mind the connection $V_d \leftrightarrow \chi$ just described.

[14] To develop an intuitive understanding of the Euler characteristic χ we could consider its 'classical' definition for the

Two–dimensional unit sphere

We now focus on the supporting space for our convergence and shear maps, the sphere $\mathbb{S}^2$ of radius R. The parameters introduced in the previous section now take the values $d = 2$ for the dimension, $k = +1$ for the curvature sign, and $K = R^{-2}$ for the absolute value of the curvature.

Let us now consider a smooth scalar field $u(\mathbf{x})$ on $\mathbb{S}^2$, for example the convergence field. We wish to calculate the Minkowski functionals of the excursion set Q_ν over a given threshold ν. An excursion set for the field u is a region where the field exceeds some threshold level: for a given threshold ν, the excursion set Q_ν and its boundary ∂Q_ν for a smooth scalar field u on the sphere are mathematically defined as

$$Q_\nu = \{x \in \mathbb{S}_2 \,|\, u(x) > \nu\}\,, \tag{7.39}$$

$$\partial Q_\nu = \{x \in \mathbb{S}_2 \,|\, u(x) = \nu\}\,. \tag{7.40}$$

On the 2-sphere, $\mathbb{S}_2$, there are $2 + 1 = 3$ MFs which, up to normalization, have clear geometrical interpretations and represent the area, circumference and integrated geodesic curvature of the excursion set. Rewriting the definition in Eq. (7.34), we obtain the Minkowski functionals for a set $Q \subseteq \mathbb{S}^2$ with smooth boundary ∂Q by

$$V_0(Q) = \int_Q \mathrm{d}a, \tag{7.41}$$

$$V_1(Q) = \frac{1}{4}\int_{\partial Q} \mathrm{d}\ell, \tag{7.42}$$

$$V_2(Q) = \frac{1}{2\pi}\int_{\partial Q} \mathrm{d}\ell\,\kappa, \tag{7.43}$$

where $\mathrm{d}a$ and $\mathrm{d}\ell$ denote the surface element of $\mathbb{S}^2$ and the line element along ∂Q, respectively. Being a linear object, the boundary ∂Q has only one geodesic curvature κ. Using the generalized Gauss–Bonnet Theorem in Eq. (7.37) with the coefficients for two dimensions substituted, we can calculate the Euler characteristic $\chi(Q)$ from the Minkowski functionals via

$$\chi(Q) = V_2(Q) + \frac{1}{2\pi R^2} V_0(Q). \tag{7.44}$$

The zeroth Minkowski functional V_0, i.e. the area, can be evaluated by integration of a Heaviside step function over the whole sphere

$$V_0(Q_\nu) = \int_{\mathbb{S}^2} \mathrm{d}a\,\Theta(u - \nu). \tag{7.45}$$

The other Minkowski functionals are actually defined by line integrals along the isodensity contour in Eq. (7.43), but they can be transformed into surface integrals by inserting a delta function, and the

surfaces of polyhedra, i.e. according to the formula $\chi = V - E + F$, where V, E, and F are respectively the numbers of vertices (corners), edges and faces in the given polyhedron. Any convex polyhedron's surface has Euler characteristic $V - E + F = 2$. This is known as the Euler's polyhedron formula and corresponds to the Euler characteristic of the sphere (i.e. $\chi = 2$), and applies identically to spherical polyhedra. A spherical polyhedron is a tiling of the sphere in which the surface is partitioned into bounded regions called spherical polygons. The most familiar spherical polyhedron is the football, which can be thought of as a spherical truncated icosahedron.

appropriate Jacobian

$$V_1(Q_\nu) = \frac{1}{4}\int_{\partial Q_\nu} \mathrm{d}\ell = \frac{1}{4}\int_{\mathbb{S}^2} \mathrm{d}a\,\delta(u-\nu)|\nabla u|,$$
$$V_2(Q_\nu) = \frac{1}{2\pi}\int_{\partial Q_\nu} \mathrm{d}\ell\,\kappa = \frac{1}{2\pi}\int_{\mathbb{S}^2} \mathrm{d}a\,\delta(u-\nu)|\nabla u|\,\kappa. \tag{7.46}$$

The geodesic curvature, κ, describes the deviation of the curve γ from being geodetic, i.e. a straight line. For a normalised tangent, i.e. $|\dot{\gamma}| = 1$, it is defined through

$$\kappa := |\nabla_{\dot{\gamma}}\dot{\gamma}|\,, \tag{7.47}$$

where $\nabla_{\dot{\gamma}}$ represents the covariant derivative along the tangent vector $\dot{\gamma}$ of the curve. Thus κ vanishes if and only if γ is a geodesic. However, for the numerical calculation of V_2, it is convenient to express κ in terms of u. To do so, one can use the fact that u does not change along γ and thus $\mathrm{d}u(\dot{\gamma}) = 0$ which implies that $\dot{\gamma}^\mu = \epsilon^{\mu\nu}\nabla_\nu u$, where $\epsilon^{\mu\nu}$ is the totally antisymmetric second-rank tensor. Upon normalization, this can be used in equation (7.47) to yield κ in terms of the metric and derivatives of u in a local coordinate system. Schmalzing & Gorski (1998) show how to calculate κ on a generic manifold, which in the case of $\mathbb{S}_2$ reads

$$\kappa = \frac{2u_{;\theta}u_{;\phi}u_{;\theta\phi} - u_{;\theta}^2 u_{;\phi\phi} - u_{;\phi}^2 u_{\theta\theta}}{u_{;\theta}^2 + u_{;\phi}^2}, \tag{7.48}$$

where the semicolon denotes a covariant derivative. Since the integrands can now be written as second–order invariants, we have succeeded in expressing all Minkowski functionals as surface integrals over the whole sphere $\mathbb{S}^2$,

$$V_j(Q_\nu) = \int_{\mathbb{S}^2} \mathrm{d}a\,\mathcal{I}_j, \tag{7.49}$$

with integrands $\mathcal{I}_j$ depending solely on the threshold ν, the field value u and its first– and second–order covariant derivatives. In summary,

$$\mathcal{I}_0 = \Theta(u-\nu), \tag{7.50}$$
$$\mathcal{I}_1 = \frac{1}{4}\delta(u-\nu)\sqrt{u_{;1}^2 + u_{;2}^2}, \tag{7.51}$$
$$\mathcal{I}_2 = \frac{1}{2\pi}\delta(u-\nu)\frac{2u_{;1}u_{;2}u_{;12} - u_{;1}^2 u_{;22} - u_{;2}^2 u_{;11}}{u_{;1}^2 + u_{;2}^2}. \tag{7.52}$$

In the following, when we refer to the MFs we will actually mean the surface densities of the Minkowski functionals, i.e. the same quantities normalised by a constant numerical factor given by the area of $\mathbb{S}^2$,

$$v_j(\nu) = \frac{1}{4\pi R^2}V_j(Q_\nu) = \frac{1}{4\pi R^2}\int_{\mathbb{S}^2} \mathrm{d}a\,\mathcal{I}_j. \tag{7.53}$$

Summarising:

- The first MF $V_0(\nu)$ is the area fraction of Q_ν, given by

$$V_0(\nu) := \frac{1}{4\pi}\int_{\mathbb{S}_2} \mathrm{d}\Omega\,\Theta(u-\nu)\,, \tag{7.54}$$

  where Θ is the Heaviside function.

- The second MF is proportional to the total boundary length of Q_ν and is thus

$$V_1(\nu) := \frac{1}{16\pi}\int_{\partial\mathbb{S}_2} \mathrm{d}l = \frac{1}{16\pi}\int_{\mathbb{S}_2} \mathrm{d}\Omega\,\delta(u-\nu)\,|\nabla u|\,. \tag{7.55}$$

  Here δ is the delta distribution and $|\nabla u|$ is the norm of the gradient of u.

- Finally, the third MF is the integral of the geodesic curvature κ along the boundary

$$V_2(\nu) := \frac{1}{8\pi^2}\int_{\partial\mathbb{S}_2} \mathrm{d}l\,\kappa = \frac{1}{8\pi^2}\int_{\mathbb{S}_2} \mathrm{d}\Omega\,\delta(u-\nu)\,|\nabla u|\,\kappa\,. \tag{7.56}$$

Expectation values for a Gaussian random field

Minkowski functionals and other geometric characteristics of Gaussian random fields are extensively studied by Adler (1981). Analytical expressions for the average Minkowski functionals of a Gaussian random field in arbitrary dimensions were derived by Tomita (1986); in the special case of two dimensions, the results for the isodensity contour at threshold ν are[15]

$$\begin{aligned} v_0(\nu) &= \frac{1}{2} - \frac{1}{2}\Phi\left(\frac{\nu-\mu}{\sqrt{2\sigma}}\right), \\ v_1(\nu) &= \frac{\tau^{1/2}}{8\sigma^{1/2}}\exp\left(-\frac{(\nu-\mu)^2}{2\sigma}\right), \\ v_2(\nu) &= \frac{\tau}{2\pi^{3/2}\sigma}\frac{\nu-\mu}{\sqrt{2\sigma}}\exp\left(-\frac{(\nu-\mu)^2}{2\sigma}\right). \end{aligned} \tag{7.57}$$

Note that these expressions contain only three parameters, namely μ, σ, and τ. All three are easily estimated from a given realization of the Gaussian random field, by taking averages of the field itself, its square, and the sum of its squared derivatives; then

$$\begin{aligned} \mu &= \langle u \rangle \\ \sigma &= \left\langle u^2 \right\rangle - \mu^2 \\ \tau &= \frac{1}{2}\left\langle u_{;i} u_{;i} \right\rangle. \end{aligned} \tag{7.58}$$

With these relations and the spherical harmonics expansion of u, the parameters σ and τ may also be calculated directly from the angular power spectrum C_ℓ, with the results

$$\begin{aligned} \sigma &= \sum_{\ell=1}^{\infty}(2\ell+1)C_\ell, \\ \tau &= \sum_{\ell=1}^{\infty}(2\ell+1)C_\ell\frac{\ell(\ell+1)}{2}. \end{aligned} \tag{7.59}$$

7.5.2. Numerical calculation of MFs

In order to numerically estimate the MFs calculated on our realisations of the lensing fields we make use of the HEALPix suite of tools (Górski et al., 2005). We first generate full sky maps of e.g. the convergence field, on concentric spherical shells at different radii, starting from the 3D covariance matrices; to this purpose we follow the procedure described in sec. 7.4. We then calculate numerically

[15] The function $\Phi(x)$ is the Gaussian error function given by $\Phi(x) = \frac{2}{\sqrt{\pi}}\int_0^x \mathrm{d}t\exp(-t^2)$.

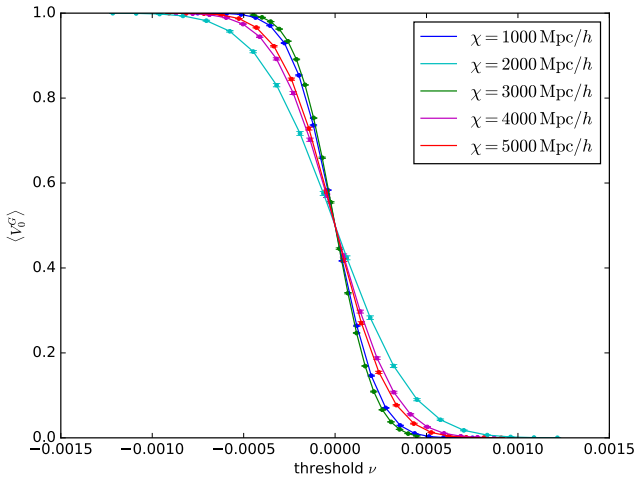

Figure 7.7.: Numerical estimations of the first MF V_0^G (*dots*), calculated on our generated Gaussian fields at different values of the radius (represented by different colours), compared with the theoretical predictions given by Eq. 7.67 (joined by *lines*), as a function of the threshold ν. The range of the thresholds always varies between $-4\sqrt{\sigma}$ and $+4\sqrt{\sigma}$, where σ is the (average) variance of the lensing field at a fixed radius.

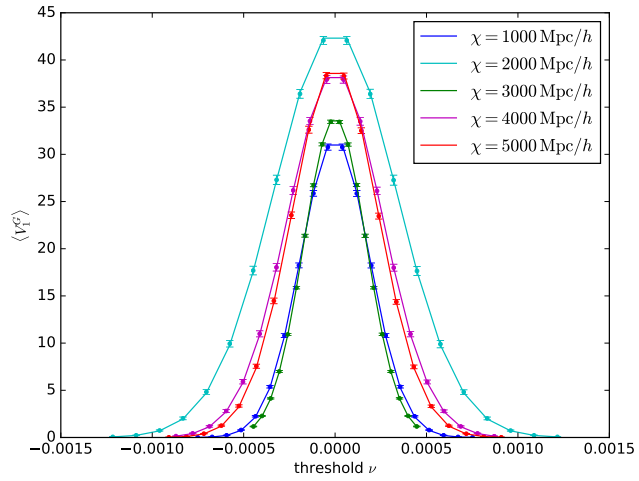

Figure 7.8.: Numerical estimations of the second MF V_1^G (*dots*), calculated on our generated Gaussian fields at different values of the radius (represented by different colours), compared with the theoretical predictions. The colour scheme is the same as in Fig. 7.7.

the MFs by directly implementing the integrals in Eqs. 7.54-7.56; our algorithm closely follows the one used in Schmalzing & Gorski (1998) and Lim & Simon (2012) and is detailed in the following.

Given a pixelated map with field values $u(x_i)$, HEALPix easily allows for the calculation of first and second partial derivatives at each pixel in (ℓ, m) spherical harmonic space. As seen in 7.5.1, the three numerical MFs for $\mathbb{S}_2$, which we label $V_i (i = 0, 1, 2)$ can be computed via a sum over all pixels

$$V_i(\nu) := \frac{1}{N_{\mathrm{pix}}} \sum_{j=1}^{N_{\mathrm{pix}}} \mathcal{I}_i(\nu, x_j) \tag{7.60}$$

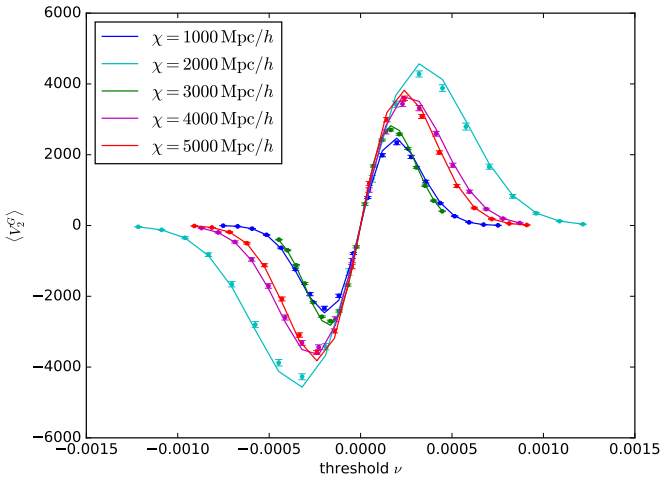

Figure 7.9.: Numerical estimations of the third MF V_2^G (*dots*), calculated on our generated Gaussian fields at different values of the radius (represented by different colours), compared with the theoretical predictions. The colour scheme is the same as in Fig. 7.7.

of the respective integrands

$$\mathcal{I}_0(\nu, x_j) := \Theta(u - \nu)\,, \tag{7.61}$$

$$\mathcal{I}_1(\nu, x_j) := \frac{1}{4}\delta(u - \nu)\sqrt{u_{;\theta}^2 + u_{;\phi}^2}\,, \tag{7.62}$$

$$\mathcal{I}_2(\nu, x_j) := \frac{1}{2\pi}\delta(u - \nu)\frac{2u_{;\theta}u_{;\phi}u_{;\theta\phi} - u_{;\theta}^2 u_{;\phi\phi} - u_{;\phi}^2 u_{\theta\theta}}{u_{;\theta}^2 + u_{;\phi}^2}\,, \tag{7.63}$$

where the semicolon indicates a covariant derivative.

Using the well–known parametrization of the unit sphere through azimuth angle ϑ and polar angle φ we can express the covariant derivatives at a point (ϑ, φ) in terms of the partial derivatives[16];

$$\begin{aligned}
u_{;\vartheta} &= u_{,\vartheta}, \\
u_{;\varphi} &= \frac{1}{\sin\vartheta}u_{,\varphi}, \\
u_{;\vartheta\vartheta} &= u_{,\vartheta\vartheta}, \\
u_{;\vartheta\varphi} &= \frac{1}{\sin\vartheta}u_{,\vartheta\varphi} - \frac{\cos\vartheta}{\sin^2\vartheta}u_{,\varphi}, \\
u_{;\varphi\varphi} &= \frac{1}{\sin^2\vartheta}u_{,\varphi\varphi} + \frac{\cos\vartheta}{\sin\vartheta}u_{,\vartheta}.
\end{aligned} \tag{7.64}$$

The partial derivatives in turn are best calculated from the spherical harmonics expansion

$$u(\vartheta, \varphi) = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} a_{\ell m}Y_{\ell m}(\vartheta, \varphi). \tag{7.65}$$

This is simply done by replacing the harmonic function $Y_{\ell m}$ with its appropriate partial derivative. Since the functions $Y_{\ell m}$ depend on φ via sine and cosine functions only, the derivatives with respect to φ can be obtained analytically. Partial derivatives with respect to ϑ are calculated via recursion

[16]Note that we use indices following a semicolon, such as $u_{;i}$ to denote covariant differentiation of u with respect to the coordinate i, as opposed to partial derivatives where we write indices following a comma, e.g. $u_{,i}$.

formulae constructed by differentiating the recursion for the associated Legendre functions P_ℓ^m, given for example by Abramowitz et al. (1988). In HEALPix, these partial derivatives are implemented in the function `alm2mapder`, which given a set of $a_{\ell m}$ returns a pixelised map in the angular coordinates of a field, as well as its angular partial derivatives.

We still have to account for the finite number of sample points. The integrands I_1 and I_2 involve the delta function, which is numerically approximated through a discretization of threshold space in bins of width $\Delta\nu$ by the Heaviside function

$$\delta_N(x) := (\Delta\nu)^{-1}[\Theta(x + \Delta\nu/2) - \Theta(x - \Delta\nu/2)] \,. \tag{7.66}$$

This approximation of the delta function produces some numerical noise, which Lim & Simon (2012) demonstrate to be due to the delta function discretization rather than some random noise which should disappear averaging over n_R realisations. For our purposes, we do not consider the corrections proposed by Lim & Simon (2012) to remove this discretisation effect and simply average over many realisations of the field. This is enough for our purposes, as our main goal is to test the field generation procedure rather than using the MFs to study e.g. non-Gaussianity as in Lim & Simon (2012) (in which case these corrections should be taken into account).

For Gaussian fields, as the ones we are considering here, the expectation values for the MFs are known analytically and equal to

$$\bar{V}_0^G(\nu) := \left\langle V_0^G(\nu) \right\rangle = \frac{1}{2}\left(1 - \mathrm{erf}\left(\frac{\nu - \mu}{\sqrt{2\sigma}}\right)\right) , \tag{7.67}$$

$$\bar{V}_1^G(\nu) := \left\langle V_1^G(\nu) \right\rangle = \frac{1}{8}\sqrt{\frac{\tau}{\sigma}}\exp\left(-\frac{(\nu - \mu)^2}{2\sigma}\right) , \tag{7.68}$$

$$\bar{V}_2^G(\nu) := \left\langle V_2^G(\nu) \right\rangle = \frac{1}{(2\pi)^{3/2}}\frac{\tau}{\sigma}\frac{\nu - \mu}{\sqrt{\sigma}}\exp\left(-\frac{(\nu - \mu)^2}{2\sigma}\right) . \tag{7.69}$$

Therefore we can compare our numerical estimates with the theoretical expectation values as a check for the validity of our field generation procedure. We perform this comparison in Figs. 7.7, 7.8, 7.9, where we overplot our numerical estimates and their expectation values. We consider all three MFs and show the comparison for five values of the radii, corresponding to five concentric shells over which we generate our lensing field. We calculate our MFs over a set of thresholds that always ranges between $-4\sqrt{\sigma}$ and $+4\sqrt{\sigma}$, where σ is the variance of the lensing field at a certain radius. The error bars associated to our numerical estimates of the MFs are taken as the square root of the diagonal elements of the covariance matrix of the MFs, computed as

$$\mathrm{Cov}_{ij} = \frac{1}{n_R - 1}\sum_{m=1}^{n_R}\Big(V_i^m - \langle V_i\rangle\Big)\Big(V_j^m - \langle V_j\rangle\Big), \quad i, j = 0, \cdots 3 \cdot n_\chi \cdot n_\nu \tag{7.70}$$

where the indices i, j run over the type of Minkowski Functional (the three MFs V_0, V_1, V_2), the number of radii n_χ and the number of thresholds n_ν. $\langle V_i\rangle$ denotes the mean of the MFs over all realisations n_R, $\langle V_i\rangle = \frac{1}{n_R}\sum_{m=1}^{n_R} V_i^m$.

An example of this matrix is presented in Fig. 7.11. We consider the covariance between all three MFs (V_0, V_1, V_2, each of them as a function of the threshold ν, ranging from ν_1 to $\nu_{\max}$), and we include the correlations between MFs belonging to each different radius (labelled by different χ value, from χ_1 to $\chi_{\max}$). The values of the radii are the same used for Figs. 7.7, 7.8, 7.9. We stress here that the error bars depicted in Figs. 7.7, 7.8, 7.9, associated to the MFs calculated for each value of the threshold are not independent. Also, since the realisations of the random fields on different radial shells are not statistically independent, the MFs on different radii are not independent either. To give

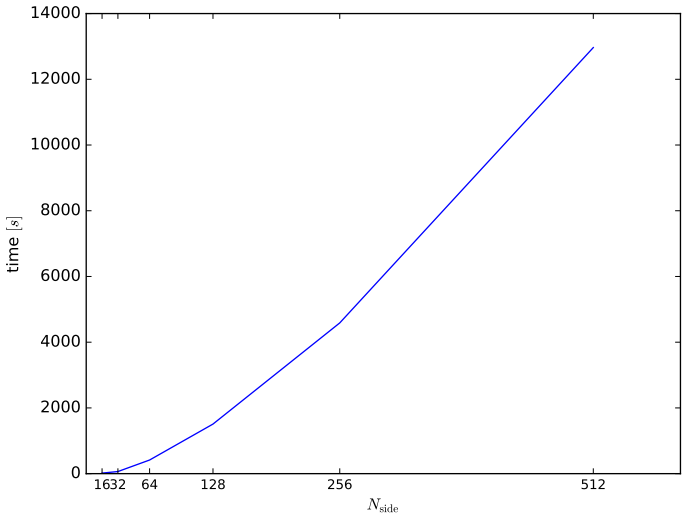

Figure 7.10.: Scaling of the run time in seconds for the code implementing the generation of a convergence random field and the estimation of the relative Minkowski Functionals, as a function of the HEALPix parameter $N_{\rm side}$ used to discretise the map. All values refer to a single realisation of the random field at $\chi = 1000{\rm Mpc}/h$. The scaling is approximately $O(N_{\rm side}^{1.5})$.

a flavour of the correlations between the MFs, in Fig. 7.12 we show the elements of the correlation matrix, i.e. the Pearson correlation coefficients r_{ij} calculated as

$$r_{ij} = \frac{{\rm Cov}_{ij}}{\sqrt{{\rm Cov}_{ii}}\sqrt{{\rm Cov}_{jj}}}, \quad i,j = 0, \cdots 3 \cdot n_\chi \cdot n_\nu \tag{7.71}$$

from the covariance matrix ${\rm Cov}_{ij}$ of our numerical estimates of the MFs, Eq. 7.70. It follows from the Cauchy-Schwarz inequality that $-1 \leq r_{ij} \leq 1$. A value of $r_{ij} = 1$ indicates a perfect linear correlation between the two variables i and j; a common interpretation is that in this case all data points in a sample lie on a straight line. This is also true if $r_{ij} = -1$, but the slope of the line is negative. A vanishing correlation coefficient implies that there is no linear correlation. If the correlation coefficient is positive, deviations of both variables from the mean tend to have the same sign, whereas opposite signs lead to a negative correlation coefficient.

As expected, we notice in particular a strong anti-correlation for V_0 centered around $\nu = 0$, as was expected by looking at Fig. 7.7. The same is true for V_2 (cf. Fig. 7.9), while V_1 is strongly positively correlated (cf. Fig 7.8). This high amount of (anti)correlation suggests that in the Gaussian case analysed here it is not necessary to consider a very high number of threshold values; however, this may not be true in the non-Gaussian case, where a higher resolution in the threshold values may be important to identify non-Gaussian features. Crucially important is, in all cases, a sufficient resolution in the HEALPix maps used at the beginning for the generation of the random fields, and later for the calculation of the MFs (in our estimates, we used the HEALPix parameters $N_{\rm side} = 256$ and $\ell_{\rm max} = 3N_{\rm side} = 768$). This affects considerably the speed of the numerical implementation of these computations, however as mentioned earlier in Sec. 7.4 the generation of random fields and, separately, the calculation of the MFs (both happening at each realisation and at each radius) are embarrassingly parallel problems; this can be leveraged in practical implementations by employing parallelisation across multiple cores and nodes, without the need to worry about inter-process communication. In Fig. 7.10 we report a study of the scaling of the code performance (considering one realisation of the convergence field and the estimation of the associated Minkowski Functionals) with the HEALPix parameter $N_{\rm side}$. Overall, we identify a scaling approximately $O(N_{\rm side}^{1.5})$.

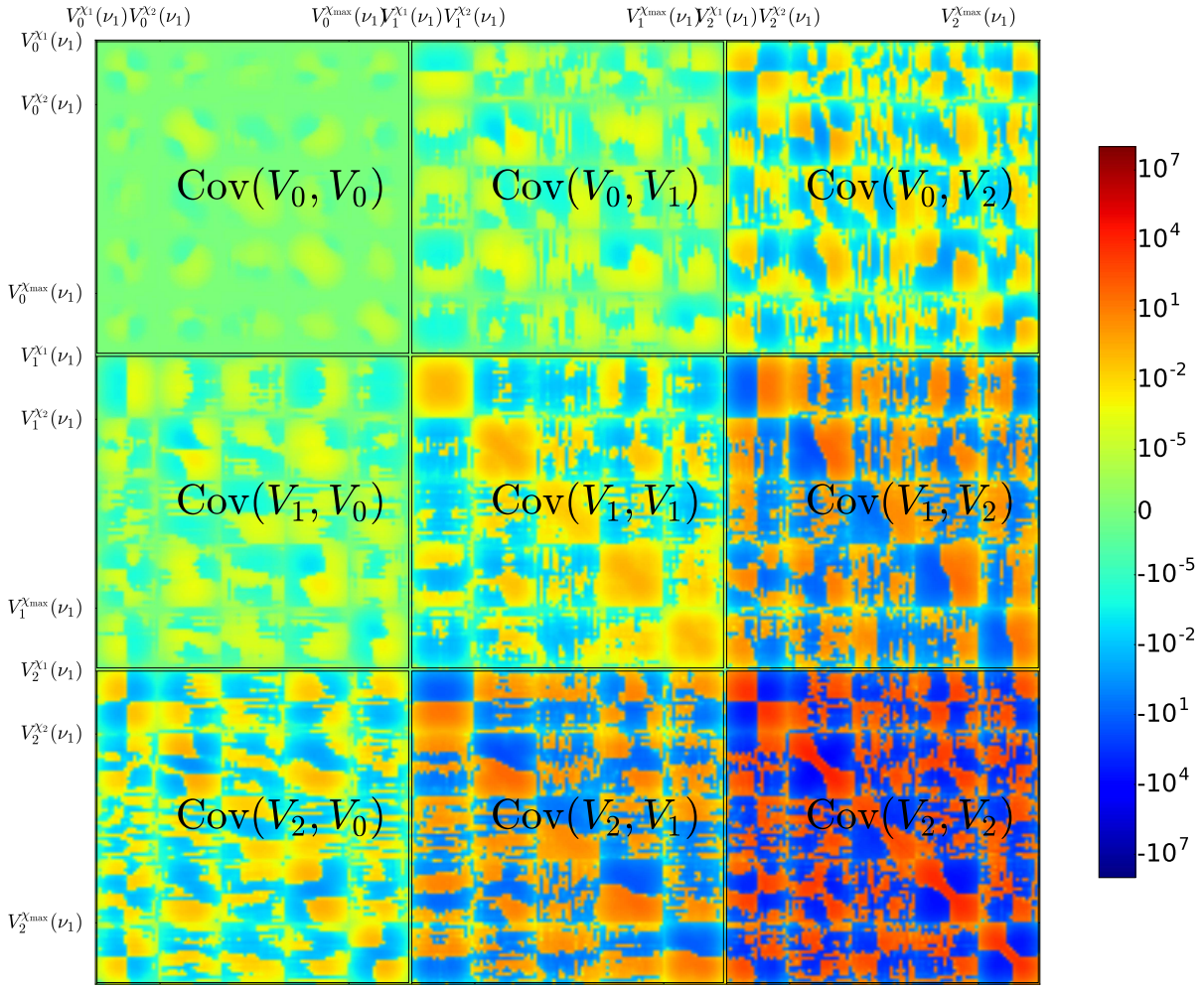

Figure 7.11.: Covariance matrix between different MFs at different radii. We consider correlations between all three MFs V_0, V_1, V_2, all functions of the threshold ν (ranging from ν_1 to $\nu_{\max}$), as calculated at different radii (labelled by different χ values, ranging from χ_1 to $\chi_{\max}$ and specifically equal to 1000, 2000, 3000, 4000 and 5000 Mpc/h, as in Figs. 7.7, 7.8, 7.9). In the matrix we indicate the block sub-matrices that represent the covariance between the three MFs. We used a logarithmic scale for both positive and negative values to highlight the many orders of magnitude spanned by the entries of the matrix and the different contibutions given by the three MFs.

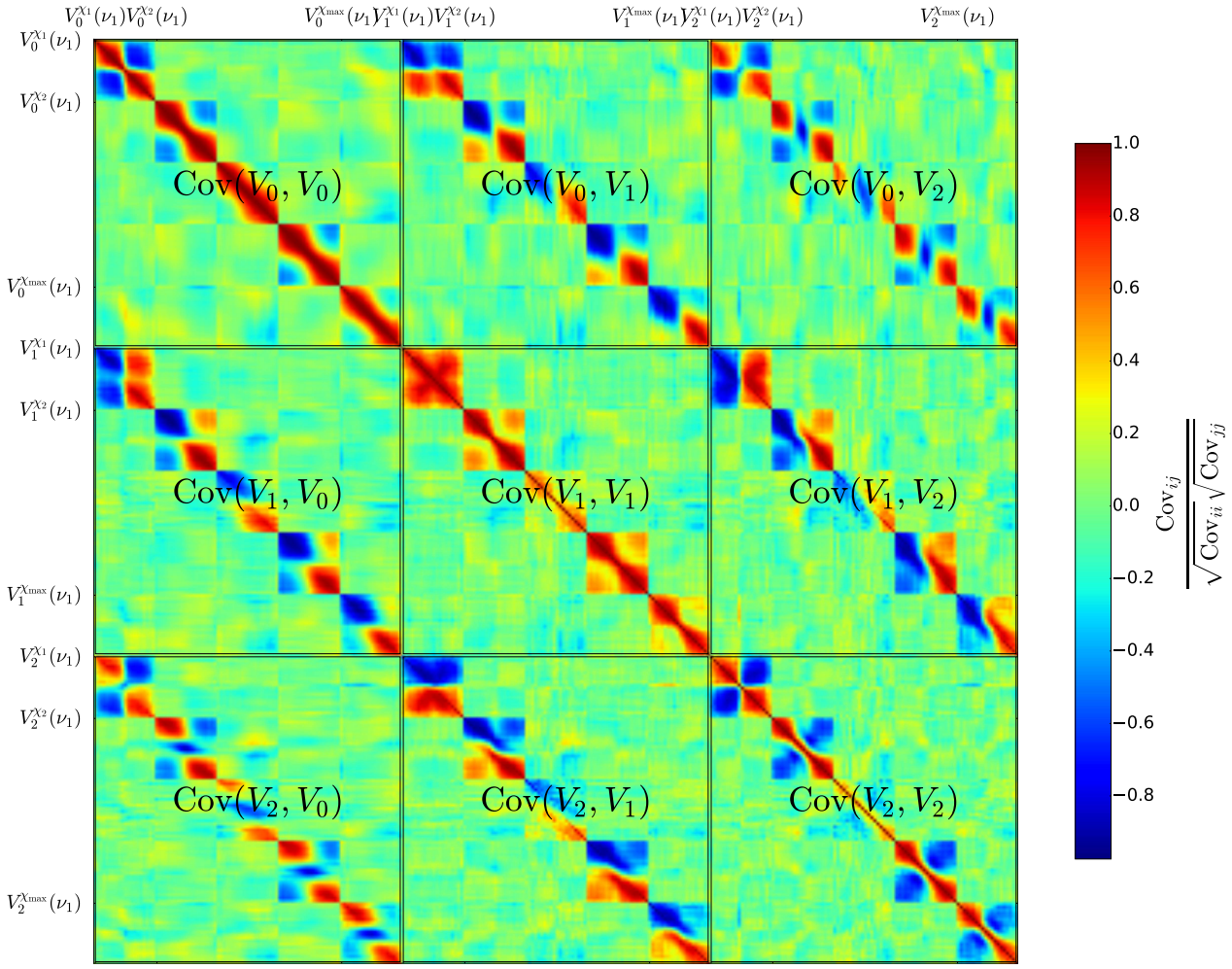

Figure 7.12.: Correlation matrix between different MFs at different radii; the matrix entries represent the Pearson correlation coefficient, obtained from the covariance matrix entries (the same plotted in in Fig. 7.11) following Eq. 7.71. We consider correlations between all three MFs V_0, V_1, V_2, all functions of the threshold ν (ranging from ν_1 to $\nu_{\max}$), as calculated at different radii (labelled by different χ values, ranging from χ_1 to $\chi_{\max}$ and specifically equal to 1000, 2000, 3000, 4000 and 5000 Mpc/h, as in Figs. 7.7, 7.8, 7.9). In the matrix we indicate the block sub-matrices that represent the correlation between the three MFs.

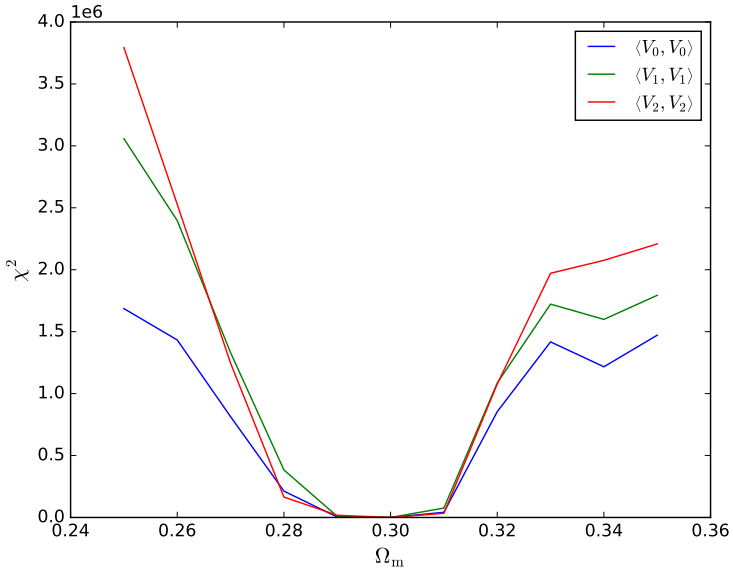

Figure 7.13.: χ^2 obtained considering the covariance of different combinations of MFs, i.e. considering the three MFs singularly ($\langle V_0, V_0 \rangle$ (*blue*), $\langle V_1, V_1 \rangle$ (*green*) and $\langle V_2, V_2 \rangle$ (*red*)). Our fiducial model is represented by the choice $\Omega_m = 0.3$.

7.5.3. Inference from Minkowski Functionals of Gaussian fields

Introduced in cosmology by Mecke et al. (1994), the main applications of MFs so far have been as probes of primordial non-Gaussianities (Schmalzing & Buchert, 1997; Winitzki & Kosowsky, 1998; Schmalzing & Gorski, 1998), widely used in two and three dimensions, for instance on WMAP CMB data (Hikage et al., 2008), Planck CMB data (Ducout et al., 2013; Novaes et al., 2016; Buchert et al., 2017) and on the SDSS galaxy catalogue (Park et al., 2005; Hikage et al., 2006). In the CMB case, MFs constitute an attractive alternative to an analysis with polyspectra for a number of reasons. Firstly, contrary to the bispectrum, they are defined in configuration rather than in Fourier space, so that a robust implementation for MFs becomes in practice easier to achieve. Secondly, MFs are sensitive to the full hierarchy of higher order correlations, instead of third order only, and can provide additional information on all the non-linear coupling parameters f_{NL}, g_{NL}, ... which appear in the perturbative development of the primordial curvature perturbation (Komatsu & Spergel, 2001; Okamoto & Hu, 2002). Additionally, MFs can be analytically determined for Gaussian random fields; lastly, they are additive which makes accounting for complicated survey geometries much easier compared to estimators of polyspectra.

In this work we propose (for the first time, as to our knowledge) MFs as an alternative probe of Gaussianity, in addition to non Gaussianity, in the sense specified in the following. We show how, assuming our MFs to be Gaussian distributed, we can use the MFs to probe the cosmology dependence of the fields realisations. This can be leveraged in future work to develop a full cosmological inference process based on the MFs calculated on lensing fields, of which we provide a first example here.

From a Bayesian perspective, assuming that our likelihood $L(V_i|\Omega)$ (the probability of having MFs V_i given the cosmological parameters Ω) is Gaussian is equivalent, considering a flat prior $p(\Omega)$ on the cosmological parameters Ω, to having a Gaussian posterior $p(\Omega|V_i)$, since by virtue of Bayes theorem

$$p(V_i|\Omega) \propto L(\Omega|V_i)p(\Omega). \tag{7.72}$$

It follows that we are allowed to consider the likelihood and the posterior equivalently. In the Gaussian case, defining $\mathcal{L} = -\ln L$ and ignoring additive constants, we have that $-2\mathcal{L} = \chi^2$, where the χ^2 can be

evaluated as

$$\chi^2(\theta) = \sum_{i,j=1}^{n_\nu \cdot n_\chi} \Big(\langle V_i(\theta) \rangle - \langle V_i(\theta_0) \rangle \Big) \, \mathrm{cov}_{ij}^{-1}(\theta_0) \Big(\langle V_j(\theta) \rangle - \langle V_j(\theta_0) \rangle \Big), \tag{7.73}$$

where the averages are performed over the number of realisations n_R, while the indices i, j run over the length of our data vector, i.e. we consider the MFs evaluated at all the n_ν thresholds and all the n_χ radii. The MFs depend on the cosmological parameters and so does the covariance matrix; for the calculation of the chi-square, we use the inverse evaluated at the fiducial model θ_0.

We calculate the χ^2 statistics with MFs obtained from the realisations of the lensing random fields (we consider the convergence in this example) at different values of one cosmological parameter, for simplicity. We consider 11 values of Ω_m, ranging from 0.25 to 0.35 in equidistant intervals of 0.01 centered on the fiducial value of 0.3. For each of the Ω_m values we produce our 3D cosmic shear covariance matrix following the equations in Sec. 7.1, with either the Levin or the `GLaSS` method. Once the full lensing covariance matrix is available, we use it to generate, according to the procedure described in Sec. 7.4, n_R realisations of the convergence field at r_{N_χ} values of the radius in configuration space. On each shell and for each realisation we also calculate the associated MFs, and store them in memory. Subsequently we use them to build the full covariance matrix, exactly as the one shown in the previous subsection, however this time we will have one covariance matrix of the MFs for each starting value of Ω_m. Inverting the covariance corresponding to our fiducial value $\Omega_\mathrm{m} = 0.3$, we can then use it to calculate the χ^2 following Eq. 7.73.

The calculation of this inverse covariance matrix poses a numerical problem, in that its entries are very small and standard methods such as Gaussian elimination fail in producing a sensible inverse. We use therefore a Moore-Penrose pseudo-inverse matrix (Dresden, 1920; Penrose, 1955), after checking that it effectively produces an inverse covariance matrix that, multiplied by the covariance, gives back the identity matrix to within numerical precision.

We calculate the χ^2 isolating the different MFs in our data vector. This implies isolating from the full covariance matrix the relevant sub-blocks for the auto-correlation of V_0, V_1 and V_2 (which we will in the following schematically indicate with $\langle V_0, V_0 \rangle$, $\langle V_1, V_1 \rangle$, $\langle V_2, V_2 \rangle$, or $\mathrm{Cov}(V_0, V_0)$, $\mathrm{Cov}(V_1, V_1)$, $\mathrm{Cov}(V_2, V_2)$). These sub-blocks can be visualised by looking at the corresponding sub-blocks in the covariance matrix plotted in Fig.7.11 (e.g. the correlation $\langle V_0, V_0 \rangle$ isolates the top left corner block); in Fig. 7.13 we plot the χ^2 curves obtained with the three MFs. We notice how the χ^2 increases going from V_0 to V_2.

7.6. Discussion and conclusions

3D cosmic shear constitutes an alternative to a traditional tomographic analysis of a cosmic shear survey. The spherical-Bessel expansion of the shear field at the core of its formalism maximises the amount of redshift information; however, the calculation of the covariance matrices presents numerical difficulties due to the numerous integrations over highly oscillatory functions.

In this Chapter we described and compared two methods for the calculation of simulated 3D cosmic shear covariance matrices. While the first method implements the Levin technique for integration of the periodic oscillations of the Bessel functions, the second method, implemented in the code `GLaSS`, tackles the integrations by matrix multiplications and appropriate use of the Limber approximation.

We first compared the predictions of the two codes in terms of covariance matrices and found excellent agreement. We compared the output of the codes both in terms of the total signal-to-noise ratio and the single contributions to the covariance matrices C_ℓ, for two different values of the multipole ℓ, for both the signal and noise parts.

Once tested the accuracy of the predictions for the covariance matrices, we used the simulated matrices to generate Gaussian lensing fields on the sky. The procedure we described, based on a Cholesky decomposition of the C_ℓ matrices, allowed us to generate correlated Gaussian fields at different slices in comoving distance. The generation of normal and lognormal fields (the latter being easily obtainable from the former, by exponentiation of the Gaussian maps) can be used in future work to compute a realistic covariance matrix for a full 3D cosmic shear likelihood analysis. This should improve upon e.g. the CFHTLenS analysis for 3D cosmic shear (Kitching et al., 2014), where a covariance implementation similar to `GLaSS` was used. Kitching et al. (2014) constructed a likelihood, in which the parameter dependency was in the covariance rather than the mean shear transform coefficients. This could be improved by having a likelihood in which the covariance is used as the mean and the 4-point covariance of the covariance used.

The generation of normal and lognormal random fields, starting from the 3D cosmic shear covariance matrices, also constitutes the first step for the development of a Bayesian Hierarchical Model for 3D cosmic shear power spectra estimation (following e.g. the work of Alsing et al., 2016, 2017, and extending it to a spherical-Bessel formalism), which can be investigated in future work.

Finally, we tested our random field generation procedure by calculating Minkowski Functionals associated to our Gaussian random fields and comparing them with their known expectation values. We found good agreement between our numerical estimates and their theoretical expectation values. We calculated our Minkowski Functionals separately on each spherical shell, however we stress here that the realisations of the random fields on different radial shells are not statistically independent, as one can appreciate from the correlation matrix presented in Fig. 7.12. Future work should concentrate on estimating the full correlation between the Minkowski Functionals at different values of the radii, implementing a fully three-dimensional approach for their calculation (see e.g. Hikage et al., 2003; Gleser et al., 2006; Yoshiura et al., 2017; Appleby et al., 2018, for examples of Minkowski Functionals in 3D). Producing fully 3D Minkowski Functionals for a lognormal field in 3D can be used in particular to extract non-Gaussian information from the shear field.

Finally, we showed how Minkowski Functionals can also be used to extract Gaussian information by means of a likelihood analysis. We show an example of this in Fig. 7.13, where we plot the χ^2 obtained from the covariance of the different Minkowski Functionals as a function of the varying cosmological parameter Ω_m. This is a first example of a full cosmological inference process, making use of the Minkowski Functionals, that we plan to develop in future work.

7.7. Future applications

We already mentioned earlier in this Chapter that a reliable algorithm for the realisation of 3D normal and lognormal lensing random fields generated from the 3D cosmic shear covariance matrices is a necessary ingredient for the development of a Bayesian Hierarchical Model for 3D cosmic shear power spectra estimation. In this concluding Section we will explain in greater detail what this means and how such a method could be concretely developed in future work. We will also describe prospects for a future joint 3D cosmic shear - galaxy clustering analysis, which could also be developed within the context of a Bayesian Hierarchical Model.

Computing the posterior

To introduce the statistical concept of a Bayesian Hierarchical Model let us step back to the ultimate goal of Bayesian inference analysis, already introduced in Sec. 3.1: the posterior probability. A very general description of an experiment is that there is some prior information I, some new data d, and one or more models M with parameters θ; we wish to infer the parameters of a given model in the

light of the new data collected by the experiment. From a Bayesian perspective, the ultimate goal of any statistical analysis is therefore the posterior probability,

$$p(\theta|d, I) \propto p(d|\theta, I)p(\theta|I) \tag{7.74}$$

where we have used Bayes' theorem to write it in terms of the likelihood and the prior, and the dependence on the model has been suppressed. Once the posterior is known, one can, if desired, form Bayesian credible intervals or estimates of the parameters, but all these are ultimately byproducts of the posterior itself.

Therefore, if possible, one should always try to reconstruct the posterior distribution from an experiment. Unfortunately, this turns out to be effectively impossible in a direct way for all but the simplest experiments, since the posterior normally does not have an analytic form; for this reason, in Bayesian analysis one typically draws samples from the posterior distribution, as already discussed in Sec. 3.4 where we reviewed different sampling algorithms to target the posterior distribution. Asymptotically, these methods will sample the posterior with a density of samples that is proportional to the target, so that with a sufficient number of samples, the chain can characterise the target *in principle* to any arbitrary accuracy required. The caveat is that for some situations this may be a very demanding computational task.

However, while the posterior itself may be complicated, it may be broken down into a hierarchy of elements, each of which we can sample from. This forms what is called a 'Bayesian Hierarchical Model': a statistical model, written in multiple levels (hierarchical form) that estimates the parameters of the posterior distribution using the Bayesian method (Allenby et al., 2005). The sub-models combine to form the hierarchical model, and Bayes' theorem is used to integrate them with the observed data and to account for all the uncertainty that is present. The result of this integration is the posterior distribution.

Cosmic shear as a Bayesian Hierarchical Model

Drawing inferences about cosmology from cosmic shear survey data is computationally very challenging, because it involves a number of complex modelling elements. Cosmological parameter inference from observed galaxy shapes and redshifts requires first of all a forward model relating the cosmology to the cosmic shear field and in turn its impact on observed galaxies. Measuring galaxy shapes and redshifts from pixelized images and photometric data requires detailed models for the telescope point-spread function (PSF), seeing effects, pixel noise and other instrumental effects, as well as models for the intrinsic distributions of galaxy properties which determine their physical and photometric appearance (see Schneider et al., 2015, for a discussion).

This complex forward model entails a huge number of model parameters and data, connected via conditional probability distributions. To visualise this complex network, let us for simplicity isolate the following sub-problem, relating the cosmic shear power spectrum $\mathbf{C}$ to the realisation of the field $\mathbf{s}$. We will assume the shear field is Gaussian and fully characterized by its two-point statistics, i.e. its power spectrum (covariance matrix) $\mathbf{C}$. The model is summarized in Fig. 7.14: a prior is specified for the power spectrum $P(\mathbf{C})$ which generates a power spectrum $\mathbf{C}$, which in turn generates a shear field $\mathbf{s}$ via the density $P(\mathbf{s}|\mathbf{C})$ (which is Gaussian under our assumptions); we then specify a noise covariance matrix $\mathbf{N}$ and add noise to the realised shear map to produce a realisation of the data, i.e. a noisy estimate of the shear field, via the conditional density $P(\mathbf{d}|\mathbf{s}, \mathbf{N})$. Note that so far $\mathbf{C}$ and $\mathbf{s}$ are not specified to be in either their tomographic or fully 3D versions: the forward model described here is valid for both formalisms. The graph in Fig 7.14 shows explicitly the conditional structure of the full posterior $P(\mathbf{C}, \mathbf{s}|\mathbf{d})$ since we can simply write the posterior as the product of the conditional densities (and priors) appearing on the edges of the graph:

$$P(\mathbf{C}, \mathbf{s}|\mathbf{d}, \mathbf{N}) = \frac{P(\mathbf{d}|\mathbf{s}, \mathbf{N})P(\mathbf{s}|\mathbf{C})P(\mathbf{C})}{P(\mathbf{d})}. \tag{7.75}$$

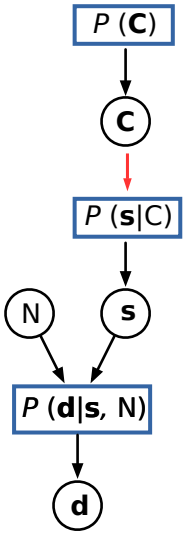

Figure 7.14.: Hierarchical forward model for noisy pixelized shear maps **d** from the shear power spectra **C**: the shear power spectrum **C** is drawn from some prior distribution, a realization of the shear field **s** is then generated given the power spectra, and finally noisy shear maps **d** are realised by adding noise with covariance **N**. Model parameters and data (nodes, represented by white circles) are connected via conditional probability distribution (represented by blue-framed boxes). The red arrow denotes the part of the model where the field-generation algorithm developed in this Chapter may represent a key tool allowing for an extension of the Bayesian Hierarchical Model developed for tomography by Alsing et al. (2016) to the 3D spherical Fourier-Bessel formalism.

In the more general case, considering the full complex forward hierarchical model for cosmic shear, the full statistical structure could be visualised with similar graphs [17] where model parameters and data (nodes) are connected via conditional probability distributions (represented by boxes). This representation clearly elicits the conditional structure of the inference problem; parameters and data which are directly connected (via a single conditional density) are dependent, whereas parameters and data which are not directly connected are conditionally independent. The posterior distribution for the full set of model parameters is straightforwardly obtained by taking the product of all distributions appearing in the graph. Furthermore, the distribution of a single parameter node conditional on all others is given by the product of the conditional densities on all incoming and outgoing edges for that parameter.

Each conditional probability distribution can be thought of as a separate modelling step; the hierarchical model thus breaks up the global problem into a number of sub-models, where one only needs to be able to write down conditional distributions for the various sub-sets of model parameters with all other parameters held fixed. Whilst the hierarchical model describes the global inference problem, it is still modular in the sense that the model neatly factorizes into a set of sub-problems that can be attacked in turn. The ultimate goal of the Bayesian Hierarchical Model is to solve the global inference problem, simultaneously inferring all of the model parameters given the data and marginalizing over all latent parameters that are not of direct interest. This approach correctly accounts for the complicated web of interdependencies; it is optimal in the sense that no information is lost, and rigorous in the sense that the uncertainties in all parameters are correctly and completely propagated throughout the analysis, from the raw data all the way through to the cosmological parameter inference. In contrast, the frequentist approach typically analyses each part of the model in a series of consecutive

[17] Sometimes in the literature these types of diagram are referred to as directed, acyclic bipartite graph, (see e.g. Constantinou & Fenton, 2018)

steps, where the results of each step are used as inputs for the next. This makes it challenging to both correctly account for the full statistical interdependency and to propagate uncertainties consistently, leading to potentially very dangerous biases that must be carefully corrected for. Whilst the typical frequentist approach estimates fixed values for parameters at each step and feeds them into the next, the hierarchical approach feeds the full probabilistic inference about each parameter throughout the analysis.

Reconstructing the cosmic shear power spectra posterior distribution

Alsing et al. (2016) developed a hierarchical Bayesian inference approach to inferring cosmological parameters from a weak lensing survey (a set of estimated shears and redshifts); rather than attempt to infer cosmological parameters directly, they achieve this via intermediate inference of the cosmic shear power spectrum.

The posterior distribution of the power spectrum is a desirable intermediate product, since cosmological models provide a deterministic relationship between cosmological parameters and the shear power spectrum. Therefore once the smooth posterior $P(\mathbf{C}|\mathbf{d})$ is obtained, cosmological parameter inference can be performed for a large number of cosmological (and systematics) models directly from $P(\mathbf{C}|\mathbf{d})$ without having to re-analyse the entire data-set.

In the following we propose a number of extensions to the work of Alsing et al. (2016), that leverage the algorithms developed in this Chapter.

3D cosmic shear power spectra estimation

From a practical perspective Alsing et al. (2016) constructed their actual implementation of the Bayesian Hierarchical Model using tomographic cosmic shear power spectra: the reason for this is mere computational simplicity. The statistical formalism developed in their work extends naturally to a fully 3D shear analysis (albeit at additional computational cost) and should be pursued.

Indeed, ideally one would like to analyse the full 3D shear field rather than the tomographic one in the context of Bayesian hierarchical inference. A forward model for the shapes of galaxies at points in 3D space necessarily requires some reference to the full 3D shear field; by restricting the analysis to a tomographic analysis without modelling the fluctuation of the field within the redshift bins, one cannot forward model a catalogue of individual galaxy shapes, angular positions and redshifts. It is possible, however, to write down a forward model for the average ellipticities of sources binned in redshift at angular positions on the sky with reference only to the tomographic shear fields. However, this way one is forced to process the catalogue of galaxy shapes and positions further into pixelized 2D maps of the average shapes in each pixel for sources in each redshift bin.

This demonstrates clearly that a tomographic analysis is sub-optimal on two counts: information is lost in compressing the data from a full 3D catalogue of galaxy shapes to a collection of averages (since the detailed redshift dependence of the power spectrum contains cosmological information), and secondly in a tomographic analysis it is not possible to include the full interdependence of the shear field and galaxy redshifts.

The algorithm we developed for the generation of lensing random fields starting from the 3D covariance matrices could be used in this Bayesian Hierarchical Model for 3D cosmic shear power spectra estimation, to link the spectra $\mathbf{C}$ to the field $\mathbf{s}$. In this case, $\mathbf{C}$ would specialise to the $C_\ell(k,k')$ and the field $\mathbf{s}$ would be given by the coefficients $\gamma_{\ell m}(k)$ in the spherical Fourier-Bessel formalism. The red arrow in Fig. 7.14 shows the passage that we made possible with our work. Normally, when given the task to generate Fourier modes from a given power spectrum, dependent only on the absolute value of the wavevector, one would generate e.g. Gaussian random fields in Fourier space from this power spectrum by exploiting the independence of the Fourier modes δ_k, which can be extracted from

a Gaussian distribution with variance equal to the power spectrum at the corresponding value of the wavevector:

$$\delta_k \sim \mathcal{N}(0, \frac{P(k)}{2}). \tag{7.76}$$

In the spherical Fourier-Bessel formalism, the cosmic shear power spectra $C_\ell(k, k')$ are not diagonal in k space. As explained in Sec.7.1, this reflects the inhomogeneity of the lensing fields, as a consequence of the lensing efficiency, the source redshift distribution and the redshift error distribution: all these effects contribute to the radial correlation of the lensing field. Employing a Cholesky decomposition of the covariance matrices, we demonstrated how to overcome this complication and generate 3D correlated lensing random fields.

Lognormal fields

The method developed by Alsing et al. (2016) is exact and optimal under the assumption of Gaussian shear fields.The general Bayesian hierarchical approach extends however naturally to include non-Gaussian models for the lensing fields, exploiting information beyond the power spectrum. This could be achieved by finding a transform of the shear field that Gaussianizes it. For example, there are good reasons to believe that the lensing fields are better described by lognormal rather than Gaussian statistics (Neyrinck et al., 2009). Under a lognormal model for the shear, the log-field is now fully characterised by its power spectrum and so the power spectrum inference (now of the log-field) remains an attractive intermediate product. In a 3D spherical Fourier-Bessel context, the generation of lognormal lensing fields can also be obtained thanks to our algorithms, by exponentation of the maps created in the Gaussian case. This would allow us to use the method for generating lognormal fields within a 3D Bayesian Hierarchical Model.

3D cross-correlation cosmic shear - galaxy clustering

A 3D Bayesian Hierarchical Model can (and should) be extended to perform joint inference of the lensing and clustering observables, including the cross-correlation between the two. A joint lensing-clustering analysis is very attractive for a number of reasons: lensing and clustering probe the same underlying cosmological field – the matter field – so that combining two independent measurements (including the cross-correlations) contains more statistical information than the two fields treated separately. At the same time the two observables are troubled by very different systematic effects; at the astrophysical level, lensing needs to account for intrinsic alignments whilst galaxy clustering has to model the galaxy bias. Jointly analysing lensing and clustering together allows for internal self-calibration of systematic effects, leading to a dramatic improvement in the cosmological constraints and more robust science.

Formalism for a fully 3D analysis of galaxy clustering has been already considered in the literature (see e.g. Heavens & Taylor, 1995; Rassat & Refregier, 2012; Nicola et al., 2014). A spherical Fourier-Bessel decomposition of the galaxy clustering observable has been recently considered in the literature by Lanusse et al. (2015). They found that the spherical Fourier-Bessel analysis is more robust than the tomographic one with respect to the inclusion of nuisance parameters. As a consequence of this, to the purpose of optimising a stage-IV type spectroscopic galaxy survey, the authors found that a given level of accuracy can be achieved for shallower surveys if a 3D spherical Fourier-Bessel analysis is performed, and that a given increase of the survey depth yields more information performing a 3D analysis than a tomographic one.

However, the cross-covariance between the lensing and clustering observables has never been studied in the spherical Fourier-Bessel formalism. A possible complication here arises from the treatment of the bias, relating the matter to the galaxy overdensity, as we shall see here.

The galaxy clustering observable does not have a broad kernel integration like in lensing, which in the spherical Fourier-Bessel cosmic shear covariance matrices appears in Eq. 7.5 and correlates different radial modes. However, relating the coefficients in the 3D expansion of the galaxy overdensity to those of the matter overdensity becomes complicated because of the bias function, therefore affecting also the relation between the cross-covariance of lensing-clustering coefficients and the matter power spectrum.

Let us first consider how the 3D expansion for the galaxy number density $n(\boldsymbol{\chi}) = (\chi, \theta, \varphi)$, the quantity observed in a galaxy clustering survey, looks like. The galaxy number density can be defined in terms of the galaxy overdensity δ_g through

$$n(\boldsymbol{\chi}) = \bar{n}(\chi)(1 + \delta_g(\boldsymbol{\chi}, z(\chi))) \,, \tag{7.77}$$

where $\bar{n}(\chi)$ is the mean number density of observed galaxies at comoving distance χ. In this expression, the time dependence of the observed overdensity as a function of comoving distance is made explicit through the $z(\chi)$ relation. In reality, we only partially observe this field due to the finite survey volume. The mean number density $\bar{n}(\chi)$ can be expressed in terms of the survey selection function $\phi(\chi)$ as

$$\bar{n}(\chi) = \phi(\chi)\bar{n} = \frac{N}{V}\phi(\chi) \,, \tag{7.78}$$

with $\bar{n}$ the mean number density of observed galaxies, N the total number of observed galaxies, and V the volume of the survey that fulfills $V = \int \phi(\chi) \mathrm{d}\boldsymbol{\chi}$. The observed field is thus no longer homogeneous because of the radial selection function. For the full galaxy number density field the expressions for the spherical Fourier-Bessel coefficients of a scalar field $n_{\ell m}(k)$, as introduced in Chapter 4 (cf. Eq. 4.29), apply

$$n_{\ell m}(k) = \sqrt{\frac{2}{\pi}} \int_0^\infty \mathrm{d}\chi\, \chi^2 \int \mathrm{d}\Omega\, n(\chi, \theta, \varphi)\, j_\ell(k\chi) Y^*_{\ell m}(\theta, \varphi). \tag{7.79}$$

The spherical Fourier-Bessel coefficients of the finite-depth field can be related to those of the field $n(\chi, \theta, \varphi)$ (Heavens & Taylor, 1995) by

$$n^{\mathrm{obs}}_{\ell m}(k) = \int_0^\infty dk'\, W_\ell(k, k') n_{\ell m}(k'), \tag{7.80}$$

with

$$W_\ell(k, k') = \frac{2}{\pi} \int \mathrm{d}\chi\, \chi^2 \phi(r) j_\ell(kr) j_\ell(k'r). \tag{7.81}$$

The observed 2-point function can be written as

$$\langle n^{\mathrm{obs}}_{\ell m}(k) n^{\mathrm{obs}*}_{\ell' m'}(k') \rangle = C_\ell(k, k') \delta_{\ell\ell'} \delta_{mm'} \tag{7.82}$$

then $C_\ell(k, k')$ can be expressed as

$$C_\ell(k, k') = \left(\frac{2}{\pi}\right)^2 \mathrm{d}k''\, k''^2 P(k'') W_\ell(k, k'') W_\ell(k', k''), \tag{7.83}$$

where $P(k)$ is the matter power spectrum.

In expression (7.77), the time (or redshift) dependence of the galaxy overdensity is due to the growth of structure and the evolution of galaxy bias with respect to the matter density field with time. Following the approach of Rassat & Refregier (2012) and Lanusse et al. (2015), in the linear regime this dependence on redshift can be separated in the form of growth and bias prefactors,

$$\delta_g(\boldsymbol{\chi}, z(\chi)) = b(\chi, k) D(\chi) \delta(\boldsymbol{\chi}) \,, \tag{7.84}$$

where $b(r,k)$ is a bias with a possible scale dependence, $D(r)$ is the growth factor, $\delta(\mathbf{r}) = \delta(\mathbf{r}, z = 0)$ is the matter overdensity field at present day, and we ignored for simplicity the contribution from a Poisson noise term arising from the discrete nature of the observed galaxy number density (which can be included in the noise part of the covariance matrix). We can define a modified selection function that includes the effects of bias and growth in the linear regime,

$$\phi^{\rm evol} = b(\chi,k)D(\chi)\phi(\chi)\ . \tag{7.85}$$

Using this modified selection function, the observed galaxy density can now be expressed directly as a function of the true matter overdensity at present time:

$$\frac{n(\mathbf{r})}{\bar{n}} = \phi(r) + \phi^{\rm evol}(r,k)\delta(\mathbf{r})\ . \tag{7.86}$$

Remembering that

$$\bar{\gamma}_{\ell m}(k) = -\frac{3\Omega_{\rm m}}{4\pi^2\chi_H^2}\sqrt{\frac{(\ell+2)!}{(\ell-2)!}}\int {\rm d}z n_z(z)\int {\rm d}z_p p(z_p|z) j_\ell[k\chi(z_p)]\int_0^{\chi(z)}{\rm d}\chi'\frac{\chi-\chi'}{\chi\chi'}\int {\rm d}k'\delta_{\ell m}(k')j_\ell(k'\chi') \tag{7.87}$$

The calculation of the cross-covariance between the coefficients for lensing and clustering leads then to

$$\left\langle n^{\rm obs}_{\ell m}(k)\bar{\gamma}_{\ell' m'}(k')\right\rangle = \left\langle\left(\int {\rm d}k\frac{2}{\pi}\int {\rm d}\chi\chi^2\phi(\chi)j_\ell(k\chi)j_\ell(k'\chi')\sqrt{\frac{2}{\pi}}\int {\rm d}\chi\chi^2\int {\rm d}\Omega\bar{n}\left(\phi(\chi)\left[1+b(\chi,k)D(\chi)\delta(\boldsymbol{\chi})\right]\right)\right.$$
$$\left.\left(-\frac{3\Omega_{\rm m}}{4\pi^2\chi_H^2}\sqrt{\frac{(\ell+2)!}{(\ell-2)!}}\int {\rm d}z n_z(z)\int {\rm d}z_p p(z_p|z) j_\ell[k\chi(z_p)]\int_0^{\chi(z)}{\rm d}\chi'\frac{\chi-\chi'}{\chi\chi'}\int {\rm d}k'\delta_{\ell m}(k')j_\ell(k'\chi')\right)\right\rangle \tag{7.88}$$

We see how the functional form of the bias is needed, in order to relate the covariance of the lensing-clustering observable to the matter power spectrum.

8 Chapter 8
Conclusions

In this thesis we have investigated the constraints achievable on dark energy/modified gravity theories by means of cosmic shear, in different ways.

After reviewing the cosmological concordance model and the dark energy/modified gravity issue in Chapter 2, we have also had a brief overview in Chapter 3 on the statistical techniques underlying the results presented in this thesis. Chapter 4 has served as an introduction to the gravitational lensing effect, the main cosmological probe considered in this thesis. In particular, we have shown explicitly how to derive expressions for the power spectra of the lensing signal in the two formalisms for a cosmic shear survey analysed in the following of the thesis: tomography and 3D cosmic shear.

In Chapter 5 we have demonstrated that a 3D cosmic shear analysis has the power to constrain parameters describing Horndeski gravity (as well as standard cosmological parameters) better than a tomographic one. We have performed this forecast for Stage IV surveys such as Euclid. The increase in sensitivity for the Horndeski parameters amounts to approximately 20%, and for all cosmological parameters considered in the analysis we have identified a neat increase in the sensitivity using a fully 3D spherical Fourier-Bessel decomposition over a tomographic analysis. This suggests that 3D cosmic shear should be considered and actively further developed as a way of analysing a cosmic shear survey. Even though the computational complexity of the calculation of the signal increases, the gain in sensitivity compensates for this.

Our analysis has been restricted to multipoles up to $\ell = 1000$ and radial wavenumbers below $k = 1\,h/\mathrm{Mpc}$, avoiding uncertainties associated with smaller scales. Modes above these cuts require an adequate and fully reliable description of the non-linear matter power spectrum, currently unavailable in ΛCDM and even more in a general modified gravity context as the Horndeski class of models. To give a flavour of the potential information content of these smaller scales, to which cosmic shear is highly sensitive, we have employed a non-linear prescription of the matter power spectrum that leverages the work of Mead et al. (2015) to account for baryonic physics. We stress again here that we do not claim this prescription to represent a rigorous approach to non-linearities in modified gravity; we have rather used the Mead et al. (2015) prescription, developed for a ΛCDM context, to provide a rough estimate of the increasing constraining power that may be achievable by including non-linear corrections. Furthermore, our analysis has not included intrinsic alignments – finite auto-correlations of the intrinsic shapes of galaxies and cross-correlations with the cosmic shear signal.

Extensions of this work should then go into the direction of improving the treatment of non-linear scales. If more reliable prescriptions for non-linearities become available, they should be implemented in our framework and the analysis should be extended to smaller scales. This also implies the inclusion of intrinsic alignments models in the forecasts and the investigation of alternative forms for the screening mechanism, which we have implemented with a phenomenological cutoff at a varying scale, prior to marginalising over it.

Some of these conclusions are common to the other analyses carried out in this thesis. For example, in Chapter 6 we have considered constraints that are achievable with current surveys (specifically, the KiDS and GAMA surveys) on Horndeski parameters, and we have found that as expected the constraints are not tight. The same analysis, performed with larger data releases from the KiDS survey, could in principle lead to tighter constraints. The importance of our work lies in the fact that the methodology and numerical implementation has now already been developed and tested, and can easily be applied in the future to larger datasets. Our implementation has been compared in the ΛCDM scenario against the previous fiducial analysis of the KiDS collaboration, and we have achieved excellent agreement with their results. This represents an important benchmark and strengthens the reliability of our implementation, which is completely independent from the previous analysis. Additionally, our likelihood module is able to carry out a modified gravity analysis. On this side, the same problem with non-linear corrections found in the 3D cosmic shear forecasts has been a common point to this analysis: for the non-linear power spectrum we have followed the official KiDS prescription, using the Mead et al. (2015) correction; however, this should again be interpreted as a means to give a flavour of the constraining power of the data, but more realistic prescriptions are needed in a modified gravity context. The same applies to the screening mechanism, which also in this analysis has been implemented in a phenomenological way.

All these points need to be addressed in future work towards a realistic setup for the analysis of data coming from Stage IV surveys. In this regard, we stress that our KiDSxGAMA likelihood can well be regarded as a first implementation that could set the path for the development of a likelihood module for Stage IV surveys such as Euclid. Regarding Stage IV surveys, in Chapter 6 we have produced forecasts with both a full likelihood analysis and a Fisher matrix approach on Horndeski parameters and found that the scope for constraining these parameters is great. Our implementation can and should be extended to the inclusion of intrinsic alignments, and similar conclusions reported earlier in these Conclusions as regards non-linear prescriptions and screening mechanism implementations apply here.

As far as cross-correlations are concerned, we believe that our work in this thesis represents an important step towards the development of a multi-probe analysis in the spherical Fourier-Bessel formalism, which this thesis has studied in detail for the cosmic shear case. Our code comparison work in Chapter 7 shows the numerical stability of our implementation for the calculation of the 3D cosmic shear integrals. With such a reliable technique at hand, many possibilities are open for an extension of the covariance matrix to other observables, first and foremost galaxy clustering as we outline in Sec. 7.7.

The random fields generation procedure described in Chapter 7 is a necessary ingredient for the development of a Bayesian Hierarchical Model for 3D cosmic shear power spectra estimation, which we aim at developing in future work, possibly extending the analysis to galaxy clustering for the reasons mentioned in Sec. 7.7.

Finally, our work in Chapter 7 on the Minkowski Functionals associated to our random fields has successfully demonstrated the feasibility of a cosmological inference pipeline making use of the estimated Minkowski Functionals. We plan to develop this approach further in future work, first by implementing a Fisher matrix analysis and then moving away from the Gaussian likelihood approximation.

In conclusion, we believe that our work has succeeded in providing an overview on current and future possibilities achievable in terms of understanding the laws of gravity by means of the study of the weak gravitational lensing effect.

Appendix A

A Appendices

A.1. Shot noise in 3D cosmic shear

Here we derive explicitly the expression for the shot noise contribution to the 3D cosmic shear covariance matrix, as given by Eq. 5.17. The noise contribution is present in all the literature on 3D cosmic shear (see e.g. the seminal papers Heavens, 2003; Heavens et al., 2006), however in the context of our code comparison for 3D cosmic shear it may arise difficulties due to the different conventions used for the spherical-Bessel formalism. We refer the reader also to Appendix C in Lanusse et al. (2015), where a derivation of the shot noise term for 3D galaxy clustering is presented.

Shot noise arises by discretising the survey in cells that either contain one or zero galaxies (Peebles, 1980). We will keep the discussion more general here, for a random field $f(\vec{x})$ that is discretised on our series of cells labelled by index i. We will later specialise to our intrinsic ellipticity field. n_i represents the occupation number of the cell and f_i the value of the field in cell i:

$$f(\vec{x}) = \sum_i \delta(\vec{x} - \vec{x}_i)\, n_i\, f_i. \tag{A.1}$$

We calculate the correlation (where V is the "volume" factor for our field):

$$\langle f(\vec{x}) f(\vec{x}) \rangle = \sum_{i,j} \left\langle \delta^D(\vec{x} - \vec{x}_i) \delta^D(\vec{x} - \vec{x}_i)\, n_i\, n_j\, f_i\, f_j \right\rangle \frac{1}{V^2} \tag{A.2}$$

$$= \sum_{i,j} \delta^D(\vec{x}_i - \vec{x}_j) \left\langle n_i^2 \right\rangle \left\langle f_i^2 \right\rangle \frac{1}{V^2} \tag{A.3}$$

$$= \sum_i \delta^D(0) \langle n_i \rangle \left\langle f_i^2 \right\rangle \frac{1}{V^2} \tag{A.4}$$

$$= \sum_i \langle n_i \rangle \frac{\langle f_i^2 \rangle}{V} \tag{A.5}$$

where we used the fact that n_i and f_i are uncorrelated, $\langle n_i^2 \rangle = \langle n_i \rangle$ due to Poisson sampling and we assumed that only equal cells are correlated. In the last step we used that $\delta^D(0) = V = 4\pi$. In our case, the random field we consider is the intrinsic ellipticity of the galaxies ϵ_S. This is because, as alredy mentioned in Sec. 5.1.2, we assume the observed ellipticity ϵ to be the sum of the shear γ and the intrinsic ellipticity ϵ_S, and neglect correlations between γ and ϵ_S as given by intrinsic alignments. We denote the intrinsic ellipticity dispersion as σ_ϵ (with a typical value $\sigma_\epsilon \simeq 0.3$). Expressing the field f in a spherical basis as

$$f_{\ell m}(k) = \sqrt{\frac{2}{\pi}} \sum_i n_i\, f_i\, j_\ell(k\chi_i)\, Y_{\ell m}(\hat{\mathbf{n}}_\mathbf{i}), \tag{A.6}$$

and taking into account the redshift distribution of galaxies, one arrives at Eq. 5.17.

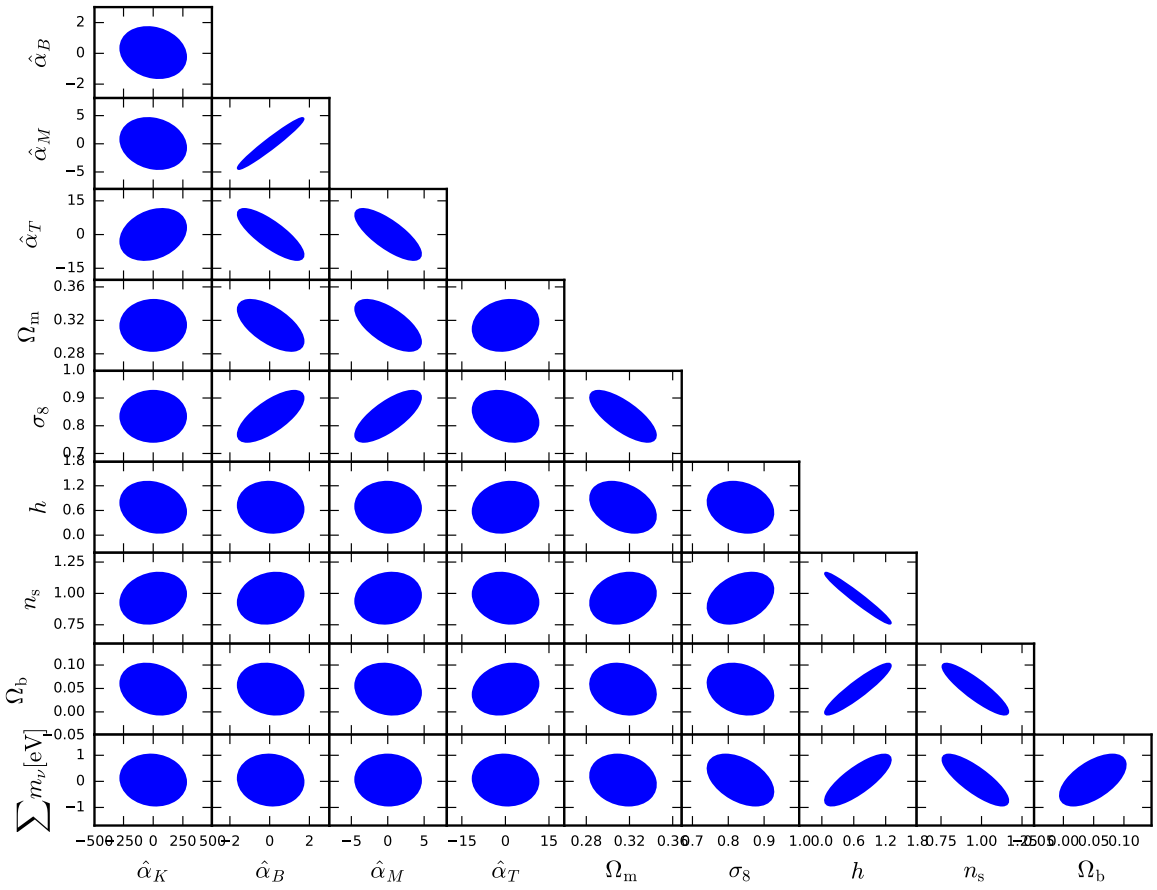

Figure A.1.: Fiducial values from Tab. 5.2 and survey specifications are given in Tab. 5.1. We point out that $\hat{\alpha}_K$ is unconstrained.

A.2. Varying $\hat{\alpha}_K$ and $\hat{\alpha}_T$

We show in Fig. A.1 the contour plots that we obtain if we also vary $\hat{\alpha}_K$ and $\hat{\alpha}_T$. We notice in particular that, as expected, $\hat{\alpha}_K$ is unconstrained, therefore we decide to fix it at its fiducial value.

A.3. Influence of k_V

In Fig. A.2 we show the influence of the screening length on the constraining power. Clearly, if k_V becomes smaller, GR is retained at larger scales already, thus decreasing the sensitivity on the modified gravity parameters. For a more complete discussion we refer the reader to Alonso et al. (2016).

A.4. Cholesky Decomposition

We review the Cholesky Decomposition algorithm to generate Gaussian random fields with a known covariance in Algorithm 5. This is widely used in statistics (see e.g. Fox & Parker, 2015), but is relatively new in cosmology.

The covariance matrix of any random vector $\boldsymbol{Y}$ is given as $\boldsymbol{E}(\boldsymbol{Y}\boldsymbol{Y}^T)$. Consider now a random vector $\boldsymbol{XY}$ consisting of uncorrelated random variables with each random variable, $\boldsymbol{X}_i$, having zero mean and

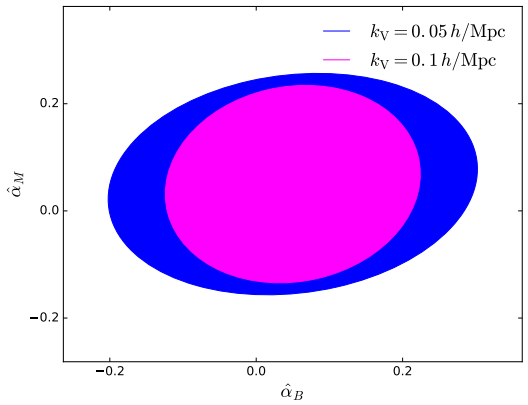

Figure A.2.: Constraints on $\hat{\alpha}_B$ and $\hat{\alpha}_M$ for two different choices of the screening length k_V.

unit variance. It follows that

$$\boldsymbol{E}(\boldsymbol{X}\boldsymbol{X}^T) = \boldsymbol{I}, \tag{A.7}$$

with $\boldsymbol{I}$ the identity matrix. To generate a random vector with given covariance matrix $\boldsymbol{Q}$, consider the Cholesky decomposition of $\boldsymbol{Q}$ such that $\boldsymbol{Q} = \boldsymbol{L}\boldsymbol{L}^T$, where $\boldsymbol{L}$ is lower triangular and L^T upper triangular. Note that it is possible to obtain a Cholesky decomposition of $\boldsymbol{Q}$ since by definition the covariance matrix $\boldsymbol{Q}$ is symmetric and positive definite.

Consider now the random vector $\boldsymbol{Z} = \boldsymbol{L}\boldsymbol{X}$. We have

$$\boldsymbol{E}(\boldsymbol{Z}\boldsymbol{Z}^T) = \boldsymbol{E}((\boldsymbol{L}\boldsymbol{X})(\boldsymbol{L}\boldsymbol{X})^T) = \boldsymbol{E}(\boldsymbol{L}\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{L}^T) \tag{A.8}$$

$$= \boldsymbol{L}\boldsymbol{E}(\boldsymbol{X}\boldsymbol{X}^T)\boldsymbol{L}^T = \boldsymbol{L}\boldsymbol{I}\boldsymbol{L}^T = \boldsymbol{Q}. \tag{A.9}$$

Hence, the random vector $\boldsymbol{Z}$ has the desired covariance matrix $\boldsymbol{Q}$.

Algorithm 5 Cholesky sampling to generate Gaussian Random Fields

Input Covariance matrix C_ℓ
Output $\gamma_\ell \sim \mathrm{N}(0, C_\ell)$
Cholesky factor $C_\ell = T_\ell T_\ell^T$
sample $z_\ell \sim \mathrm{N}(0, I)$
$\gamma_\ell = T_\ell z_\ell$

Cholesky factor $\boldsymbol{\Sigma} = \mathbf{C}\mathbf{C}^T$

sample $\mathbf{z} \sim N(0, \mathbf{I})$

$\mathbf{y} = \mathbf{C}\mathbf{z} \sim N(0, \boldsymbol{\Sigma})$

Bibliography

Abazajian, K. N., Adshead, P., Ahmed, Z., et al. 2016, ArXiv e-prints 1610.02743, 1610.02743

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, Physical Review Letters, 119, 161101

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, ApJ, 848, L12

Abbott, T., Abdalla, F. B., Allam, S., et al. 2016, Physical Review D, 94, 022001

Abramowitz, M., Stegun, I. A., & Romer, R. H. 1988, American Journal of Physics, 56, 958

Adamek, J., Durrer, R., & Tansella, V. 2016, J. Cosmol. Astropart. Phys., 1, 024

Adler, R. J. 1981, The Geometry of Random Fields

Akeret, J., Seehars, S., Amara, A., Refregier, A., & Csillaghy, A. 2012, ArXiv e-prints, arXiv:1212.1721

Alam, S., Ata, M., Bailey, S., et al. 2017, MNRAS, 470, 2617

Allenby, G., Rossi, P., & McCulloch, R. 2005

Alonso, D., Bellini, E., Ferreira, P. G., & Zumalacarregui, M. 2016, ArXiv e-prints, arXiv:1610.09290

Alsing, J., Heavens, A., & Jaffe, A. H. 2017, MNRAS, 466, 3272

Alsing, J., Heavens, A., Jaffe, A. H., et al. 2016, MNRAS, 455, 4452

Amendola, L. 2000, Phys. Rev. D, 62, 043511

Amendola, L., Kunz, M., Motta, M., Saltas, I. D., & Sawicki, I. 2013, Phys. Rev. D, 87, 023501

Amendola, L., Kunz, M., & Sapone, D. 2008, J. Cosmol. Astropart. Phys., 4, 013

Amendola, L., Appleby, S., Avgoustidis, A., et al. 2016, ArXiv e-prints, arXiv:1606.00180

Appleby, S., Chingangbam, P., Park, C., Yogendran, K. P., & Joby, P. K. 2018, ArXiv e-prints, arXiv:1805.08752

Armendáriz-Picón, C., Damour, T., & Mukhanov, V. 1999, Physics Letters B, 458, 209

Armendariz-Picon, C., Mukhanov, V., & Steinhardt, P. J. 2001, Phys. Rev. D, 63, 103510

Asgari, M., Taylor, A., Joachimi, B., & Kitching, T. D. 2018, Monthly Notices of the Royal Astronomical Society, doi:10.1093/mnras/sty1412

Audren, B., Lesgourgues, J., Benabed, K., & Prunet, S. 2013, Journal of Cosmology and Astroparticle Physics, 2013, 001

Ayaita, Y., Schäfer, B. M., & Weber, M. 2012, MNRAS, 422, 3056

Babichev, E., & Deffayet, C. 2013, Classical and Quantum Gravity, 30, 184001

Bacon, D. J., Refregier, A. R., & Ellis, R. S. 2000, MNRAS, 318, 625

Bailoni, A., Spurio Mancini, A., & Amendola, L. 2017, MNRAS, 470, 688

Baker, T., Bellini, E., Ferreira, P. G., et al. 2017, ArXiv e-prints, arXiv:1710.06394

Baker, T., & Bull, P. 2015, ApJ, 811, 116

Baldi, M., & Villaescusa-Navarro, F. 2016, ArXiv e-prints, arXiv:1608.08057

Baldi, M., Villaescusa-Navarro, F., Viel, M., et al. 2014, MNRAS, 440, 75

Barreira, A., Li, B., Hellwing, W. A., Baugh, C. M., & Pascoli, S. 2013, J. Cosmol. Astropart. Phys., 10, 027

Bartelmann, M., Fabis, F., Kozlikin, E., et al. 2017, New Journal of Physics, 19, 083001

Bartelmann, M., & Schneider, P. 2001, Phys. Rep., 340, 291

Battye, R. A., & Moss, A. 2014, Physical Review Letters, 112, doi:10.1103/PhysRevLett.112.051303

Baugh, C. M. 1996, Monthly Notices of the Royal Astronomical Society, 280, 267

Bekenstein, J. D. 1993, Physical Review D, 48, 3641

Bellini, E., Cuesta, A. J., Jimenez, R., & Verde, L. 2016, J. Cosmol. Astropart. Phys., 2, 053

Bellini, E., & Sawicki, I. 2014, Journal of Cosmology and Astroparticle Physics, 2014, 050

Bernstein, G., & Jain, B. 2004, ApJ, 600, 17

Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22

Bettoni, D., Ezquiaga, J. M., Hinterbichler, K., & Zumalacárregui, M. 2017, Phys. Rev. D, 95, 084029

Bettoni, D., & Liberati, S. 2013, Phys. Rev. D, 88, 084020

Bird, S., Viel, M., & Haehnelt, M. G. 2012, MNRAS, 420, 2551

Blas, D., & Lim, E. 2014, International Journal of Modern Physics D, 23, 1443009

Bloomfield, J. K., Flanagan, E., Park, M., & Watson, S. 2013, Journal of Cosmology and Astroparticle Physics, 2013, 010

Brans, C., & Dicke, R. H. 1961, Phys. Rev., 124, 925

Bridle, S., & King, L. 2007, New Journal of Physics, 9, 444

Brown, M. L., Taylor, A. N., Bacon, D. J., et al. 2003, MNRAS, 341, 100

Buchert, T., France, M. J., & Steiner, F. 2017, Classical and Quantum Gravity, 34, 094002

Cacciato, M., Bosch, F. C. v. d., More, S., Mo, H., & Yang, X. 2013, Monthly Notices of the Royal Astronomical Society, 430, 767

Cacciato, M., Lahav, O., Bosch, F. C. v. d., Hoekstra, H., & Dekel, A. 2012, Monthly Notices of the Royal Astronomical Society, 426, 566

Camera, S., Kitching, T. D., Heavens, A. F., Bertacca, D., & Diaferio, A. 2011, Monthly Notices of the Royal Astronomical Society, 415, 399

Carroll, S. M., de Felice, A., Duvvuri, V., et al. 2005, Phys. Rev. D, 71, 063513

Carroll, S. M., Duvvuri, V., Trodden, M., & Turner, M. S. 2004, Phys. Rev. D, 70, 043528

Casas, S., Amendola, L., Baldi, M., Pettorino, V., & Vollmer, A. 2016, J. Cosmol. Astropart. Phys., 1, 045

Casas, S., Kunz, M., Martinelli, M., & Pettorino, V. 2017, Physics of the Dark Universe, 18, 73

Castro, P. G., Heavens, A. F., & Kitching, T. D. 2005, Phys. Rev. D, 72, 023516

Catelan, P., Kamionkowski, M., & Blandford, R. D. 2001, Monthly Notices of the Royal Astronomical Society, 320, L7

Cembranos, J. A. R., Hallabrin, C., Maroto, A. L., & Jareo, S. J. N. 2012, Physical Review D, 86

Cembranos, J. A. R., Maroto, A. L., & Jareo, S. J. N. 2017, Journal of High Energy Physics, 2017

Charmousis, C., Copeland, E. J., Padilla, A., & Saffin, P. M. 2012, Physical Review Letters, 108, 051101

Chern, S.-S. 1944, Annals of Mathematics, 45, 747

Chevallier, M., & Polarski, D. 2001, International Journal of Modern Physics D, 10, 213

Clerkin, L., Kirk, D., Manera, M., et al. 2017, MNRAS, 466, 1444

Clifton, T., Ferreira, P. G., Padilla, A., & Skordis, C. 2012, Phys. Rep., 513, 1

Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, Monthly Notices of the Royal Astronomical Society, 362, 505

Constantinou, A. C., & Fenton, N. 2018, Significance, 15, 19

Creminelli, P., & Vernizzi, F. 2017, ArXiv e-prints, arXiv:1710.05877

Crisostomi, M., Koyama, K., & Tasinato, G. 2016, Journal of Cosmology and Astroparticle Physics, 2016, 044

de Jong, J. T. A., Verdoes Kleijn, G. A., Kuijken, K. H., & Valentijn, E. A. 2013, Experimental Astronomy, 35, 25

de Jong, J. T. A., Verdois Kleijn, G. A., Erben, T., et al. 2017, A&A, 604, A134

de Rham, C. 2014, Living Reviews in Relativity, 17

de Rham, C., Deskins, J. T., Tolley, A. J., & Zhou, S.-Y. 2017, Reviews of Modern Physics, 89, 025004

de Rham, C., & Gabadadze, G. 2010, Phys. Rev. D, 82, 044020

de Rham, C., Gabadadze, G., & Tolley, A. J. 2011, Physical Review Letters, 106, 231101

de Rham, C., & Tolley, A. J. 2010, J. Cosmol. Astropart. Phys., 5, 015

Deffayet, C., Esposito-Farèse, G., & Vikman, A. 2009, Phys. Rev. D, 79, 084003

Deffayet, C., Gao, X., Steer, D. A., & Zahariade, G. 2011, Phys. Rev. D, 84, 064039

Deffayet, C., Pujolàs, O., Sawicki, I., & Vikman, A. 2010, J. Cosmol. Astropart. Phys., 10, 026

Dekel, A., & Lahav, O. 1999, The Astrophysical Journal, 520, 24

Denissenya, M., & Linder, E. V. 2018, ArXiv e-prints, arXiv:1808.00013

Deser, S., & Woodard, R. P. 2007, Physical Review Letters, 99, 111301

Dresden, A. 1920, Bull. Amer. Math. Soc., 26, 385

Driver, S. P., Norberg, P., Baldry, I. K., et al. 2009, Astronomy & Geophysics, 50, 5.12

Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, Monthly Notices of the Royal Astronomical Society, 413, 971

Ducout, A., Bouchet, F. R., Colombi, S., Pogosyan, D., & Prunet, S. 2013, MNRAS, 429, 2104

Durrer, R., & Tansella, V. 2016, J. Cosmol. Astropart. Phys., 7, 037

Edge, A., Sutherland, W., Kuijken, K., et al. 2013, The Messenger, 154, 32

Efron, B. 1979, Ann. Statist., 7, 1

Eifler, T., Krause, E., Schneider, P., & Honscheid, K. 2014, Monthly Notices of the Royal Astronomical Society, 440, 1379

Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605

Eisenstein, D. J., & Zaldarriaga, M. 2001, The Astrophysical Journal, 546, 2

Ezquiaga, J., García-Bellido, J., & Zumalacárregui, M. 2016, ArXiv e-prints, arXiv:1603.01269

Ezquiaga, J., & Zumalacárregui, M. 2017, ArXiv e-prints, arXiv:1710.05901

Ezquiaga, J. M., & Zumalacárregui, M. 2018, ArXiv e-prints, arXiv:1807.09241

Fasiello, M., & Vlah, Z. 2017, Phys. Lett., B773, 236

Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601

Ferraro, S., Sherwin, B. D., & Spergel, D. N. 2015, Phys. Rev. D, 91, 083533

Fierz, M., & Pauli, W. 1939, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 173, 211

Fluri, J., Kacprzak, T., Sgier, R., Réfrégier, A., & Amara, A. 2018, ArXiv e-prints, arXiv:1803.08461

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Fox, C., & Parker, A. 2015, ArXiv e-prints, arXiv:1505.03512

Gerke, B. F., & Efstathiou, G. 2002, MNRAS, 335, 33

Gleser, L., Nusser, A., Ciardi, B., & Desjacques, V. 2006, MNRAS, 370, 1329

Gleyzes, J. 2017, Phys. Rev. D, 96, 063516

Gleyzes, J., Langlois, D., Piazza, F., & Vernizzi, F. 2013, J. Cosmol. Astropart. Phys., 8, 025

Gleyzes, J., Langlois, D., Piazza, F., & Vernizzi, F. 2015, Physical Review Letters, 114, 211101

Goldberg, J. N., Macfarlane, A. J., Newman, E. T., Rohrlich, F., & Sudarshan, E. C. G. 1967, Journal of Mathematical Physics, 8, 2155

Golovnev, A., Mukhanov, V., & Vanchurin, V. 2008

Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, Vol. 5, No. 1, p. 65-80, 2010, 5, 65

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759

Grassi, A., & Schäfer, B. M. 2014, MNRAS, 437, 2632

Gubitosi, G., Piazza, F., & Vernizzi, F. 2013, J. Cosmol. Astropart. Phys., 2, 032

Guth, A. H. 1981, Phys. Rev. D, 23, 347

Hadwiger, H. 1957, Vorlesungen über Inhalt, Oberfläche und Isoperimetrie (Springer)

Hamilton, A. J. S., & Tegmark, M. 2002, Monthly Notices of the Royal Astronomical Society, 330, 506

Hannestad, S., Tu, H., & Wong, Y. Y. 2006, J. Cosmol. Astropart. Phys., 6, 025

Harnois-Déraps, J., & van Waerbeke, L. 2015, MNRAS, 450, 2857

Hassan, S. F., & Rosen, R. A. 2012, Journal of High Energy Physics, 2, 126

Heavens, A. 2003, MNRAS, 343, 1327

Heavens, A. F., Kitching, T. D., & Taylor, A. N. 2006, MNRAS, 373, 105

Heavens, A. F., & Taylor, A. N. 1995, MNRAS, 275, 483

Heisenberg, L. 2018, ArXiv e-prints, arXiv:1807.01725

Heymans, C., Brown, M. L., Barden, M., et al. 2005, Monthly Notices of the Royal Astronomical Society, 361, 160

Heymans, C., Grocutt, E., Heavens, A., et al. 2013, MNRAS, 432, 2433

Hikage, C., Komatsu, E., & Matsubara, T. 2006, ApJ, 653, 11

Hikage, C., Matsubara, T., Coles, P., et al. 2008, Monthly Notices of the Royal Astronomical Society, 389, 1439

Hikage, C., Takada, M., Hamana, T., & Spergel, D. 2011, Monthly Notices of the Royal Astronomical Society, 412, 65

Hikage, C., Schmalzing, J., Buchert, T., et al. 2003, PASJ, 55, 911

Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, MNRAS, 465, 1454

Hinterbichler, K. 2012, Reviews of Modern Physics, 84, 671

Hinterbichler, K., & Khoury, J. 2010, Physical Review Letters, 104, 231301

Hinterbichler, K., Khoury, J., Levy, A., & Matas, A. 2011, Phys. Rev. D, 84, 103521

Hinterbichler, K., & Rosen, R. A. 2012, Journal of High Energy Physics, 2012

Hirata, C. M., & Seljak, U. 2004, Phys. Rev. D, 70, 063526

Hirata, C. M., & Seljak, U. 2010, Physical Review D, 82, doi:10.1103/PhysRevD.82.049901

Hoekstra, H., & Jain, B. 2008, Annual Review of Nuclear and Particle Science, 58, 99

Hoekstra, H., van Waerbeke, L., Gladders, M. D., Mellier, Y., & Yee, H. K. C. 2002, The Astrophysical Journal, 577, 604

Hollenstein, L., Sapone, D., Crittenden, R., & Schäfer, B. M. 2009, J. Cosmol. Astropart. Phys., 4, 012

Horava, P. 2009, Phys. Rev. D, 79, 084008

Horndeski, G. W. 1974, International Journal of Theoretical Physics, 10, 363

Hu, W. 1999, ApJ, 522, L21

Hu, W., & Jain, B. 2004, Phys. Rev. D, 70, 043009

Hu, W., & Okamoto, T. 2002, ApJ, 574, 566

Hu, W., & White, M. 2001, The Astrophysical Journal, 554, 67

Huterer, D. 2010, Gen. Rel. Grav., 42, 2177

Huterer, D., Knox, L., & Nichol, R. C. 2001, The Astrophysical Journal, 555, 547

Huterer, D., & Starkman, G. 2003, Physical Review Letters, 90, 031301

Jaccard, M., Maggiore, M., & Mitsou, E. 2013, Phys. Rev. D, 88, 044033

Jacobson, T., & Mattingly, D. 2001, Phys. Rev. D, 64, 024028

Jain, B., & Seljak, U. 1997, ApJ, 484, 560

Jain, B., & Taylor, A. 2003, Physical Review Letters, 91, 141302

Jee, M. J., Tyson, J. A., Hilbert, S., et al. 2016, The Astrophysical Journal, 824, 77

Jee, M. J., Tyson, J. A., Schneider, M. D., et al. 2013, The Astrophysical Journal, 765, 74

Joachimi, B., & Bridle, S. L. 2010, A&A, 523, A1

Joachimi, B., Schneider, P., & Eifler, T. 2008, Astronomy & Astrophysics, 477, 43

Joachimi, B., Taylor, A. N., & Kiessling, A. 2011, Monthly Notices of the Royal Astronomical Society, 418, 145

Joachimi, B., Cacciato, M., Kitching, T. D., et al. 2015, Space Sci. Rev., 193, 1

Joudaki, S., Blake, C., Heymans, C., et al. 2017, MNRAS, 465, 2033

Joyce, A., Lombriser, L., & Schmidt, F. 2016, Ann. Rev. Nucl. Part. Sci., 66, 95

Jullo, E., Rhodes, J., Kiessling, A., et al. 2012, The Astrophysical Journal, 750, 37

Kaiser, N. 1992, ApJ, 388, 272

Kaiser, N. 1998, ApJ, 498, 26

Kenney, J., & Keeping, E. 1951, Mathematics of Statistics No. pt. 2 (Van Nostrand)

Khoury, J., & Weltman, A. 2004, Phys. Rev. D, 69, 044026

Khoury, J., & Weltman, A. 2004, Phys. Rev. Lett., 93, 171104

Kilbinger, M. 2015, Reports on Progress in Physics, 78, 086901

Kilbinger, M., Benabed, K., Guy, J., et al. 2009, A&A, 497, 677

Kilbinger, M., Fu, L., Heymans, C., et al. 2013, MNRAS, 430, 2200

Kilbinger, M., Heymans, C., Asgari, M., et al. 2017, MNRAS, 472, 2126

Kitching, T. D., Alsing, J., Heavens, A. F., et al. 2017, MNRAS, 469, 2737

Kitching, T. D., & Heavens, A. F. 2017, Phys. Rev. D, 95, 063522

Kitching, T. D., Heavens, A. F., & Miller, L. 2011, MNRAS, 413, 2923

Kitching, T. D., Heavens, A. F., Taylor, A. N., et al. 2007, MNRAS, 376, 771

Kitching, T. D., Verde, L., Heavens, A. F., & Jimenez, R. 2016, MNRAS, 459, 971

Kitching, T. D., Heavens, A. F., Alsing, J., et al. 2014, MNRAS, 442, 1326

Knox, L. 1995, Phys. Rev. D, 52, 4307

Kobayashi, T., Yamaguchi, M., & Yokoyama, J. 2010, Physical Review Letters, 105, 231302

Köhlinger, F., Viola, M., Valkenburg, W., et al. 2016, Monthly Notices of the Royal Astronomical Society, 456, 1508

Köhlinger, F., Viola, M., Joachimi, B., et al. 2017, MNRAS, 471, 4412

Komatsu, E., & Spergel, D. N. 2001, Phys. Rev. D, 63, 063002

Kosowsky, A. 1998, arXiv preprint astro-ph/9805173

Krause, E., & Eifler, T. 2017, Monthly Notices of the Royal Astronomical Society, 470, 2100

Krause, E., Eifler, T. F., Zuntz, J., et al. 2017, ArXiv e-prints, arXiv:1706.09359

Kreisch, C. D., & Komatsu, E. 2017, ArXiv e-prints, arXiv:1712.02710

Kunz, M. 2012, Comptes Rendus Physique, 13, 539

Kwan, J., Sanchez, C., Clampitt, J., et al. 2017, Monthly Notices of the Royal Astronomical Society, 464, 4045

Langlois, D., & Noui, K. 2016, Journal of Cosmology and Astroparticle Physics, 2016, 034

Lanusse, F., Rassat, A., & Starck, J.-L. 2015, A&A, 578, A10

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints, arXiv:1110.3193

Laurent, P., Le Goff, J.-M., Burtin, E., et al. 2016, J. Cosmol. Astropart. Phys., 11, 060

Leistedt, B., McEwen, J. D., Kitching, T. D., & Peiris, H. V. 2015, Phys. Rev. D, 92, 123010

Lemos, P., Challinor, A., & Efstathiou, G. 2017, J. Cosmol. Astropart. Phys., 5, 014

Lesgourgues, J. 2011, ArXiv e-prints, 1104.2932

Lesgourgues, J. 2011, ArXiv e-prints, arXiv:1104.2934

Levin, D. 1996, Journal of Computational and Applied Mathematics, 67, 95

Levin, D. 1997, Journal of Computational and Applied Mathematics, 78, 131

Lewis, A., & Bridle, S. 2002, Phys. Rev. D, 66, 103511

Lewis, A., Challinor, A., & Lasenby, A. 2000, Astrophys. J., 538, 473

Lim, E. A., & Simon, D. 2012, J. Cosmol. Astropart. Phys., 1, 048

Lin, C.-A. 2016, ArXiv e-prints, arXiv:1612.04041

Lin, H., Dodelson, S., Seo, H.-J., et al. 2012, The Astrophysical Journal, 761, 15

Linder, E. V., & Jenkins, A. 2003, MNRAS, 346, 573

Liske, J., Baldry, I. K., Driver, S. P., et al. 2015, Monthly Notices of the Royal Astronomical Society, 452, 2087

Lombriser, L. 2016, J. Cosmol. Astropart. Phys., 11, 039

Lombriser, L., & Lima, N. A. 2017, Physics Letters B, 765, 382

Lombriser, L., & Taylor, A. 2016, J. Cosmol. Astropart. Phys., 3, 031

Lovelock, D. 1971, Journal of Mathematical Physics, 12, 498

Loverde, M., & Afshordi, N. 2008, Phys. Rev. D, 78, 123506

LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, ArXiv e-prints, arXiv:0912.0201

Maartens, R., Abdalla, F. B., Jarvis, M., Santos, M. G., & SKA Cosmology SWG, f. t. 2015, ArXiv e-prints, arXiv:1501.04076

MacCrann, N., Zuntz, J., Bridle, S., Jain, B., & Becker, M. R. 2014, arXiv:1408.4742 [astro-ph]

Mandelbaum, R. 2017, ArXiv e-prints, arXiv:1710.03235

Mandelbaum, R., Slosar, A., Baldauf, T., et al. 2013, Monthly Notices of the Royal Astronomical Society, 432, 1544

Martin, J. 2012, Comptes Rendus Physique, 13, 566

Martín-Moruno, P., Nunes, N. J., & Lobo, F. S. N. 2015, Phys. Rev. D, 91, 084029

Mead, A. J., Peacock, J. A., Heymans, C., Joudaki, S., & Heavens, A. F. 2015, MNRAS, 454, 1958

Mecke, K. R., Buchert, T., & Wagner, H. 1994, A&A, 288, 697

Merkel, P. M., & Schäfer, B. M. 2013, MNRAS, 434, 1808

Merkel, P. M., & Schäfer, B. M. 2017, MNRAS, 469, 2760

Munshi, D., & Kilbinger, M. 2006, A&A, 452, 63

Neal, R. M. 2012, ArXiv e-prints, arXiv:1206.1901

Nersisyan, H., Akrami, Y., Amendola, L., et al. 2017, Phys. Rev. D, 95, 043539

Newman, E., & Penrose, R. 1962, Journal of Mathematical Physics, 3, 566

Neyrinck, M. C., Szapudi, I., & Szalay, A. S. 2009, ApJ, 698, L90

Nicola, A., Refregier, A., & Amara, A. 2017, Physical Review D, 95, doi:10.1103/PhysRevD.95.083523

Nicola, A., Refregier, A., Amara, A., & Paranjape, A. 2014, Phys. Rev. D, 90, 063515

Nicolis, A., Rattazzi, R., & Trincherini, E. 2009, Phys. Rev. D, 79, 064036

Novaes, C. P., Bernui, A., Marques, G. A., & Ferreira, I. S. 2016, MNRAS, 461, 1363

Okamoto, T., & Hu, W. 2002, Phys. Rev. D, 66, 063008

Okamoto, T., & Hu, W. 2003, Phys. Rev. D, 67, 083002

Overduin, J. M., & Cooperstock, F. I. 1998, Phys. Rev. D, 58, 043506

Park, C., Choi, Y.-Y., Vogeley, M. S., et al. 2005, The Astrophysical Journal, 633, 11

Peebles, P. J. E. 1980, The large-scale structure of the universe

Peel, A., Lin, C.-A., Lanusse, F., et al. 2017, A&A, 599, A79

Peel, A., Pettorino, V., Giocoli, C., Starck, J.-L., & Baldi, M. 2018, ArXiv e-prints, arXiv:1805.05146

Peirone, S., Koyama, K., Pogosian, L., Raveri, M., & Silvestri, A. 2018, Phys. Rev. D, 97, 043519

Pen, U.-L. 1998, The Astrophysical Journal, 504, 601

Penrose, R. 1955, Mathematical Proceedings of the Cambridge Philosophical Society, 51, 406413

Perenon, L., Piazza, F., Marinoni, C., & Hui, L. 2015, arXiv:1506.03047

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565

Pettorino, V., Amendola, L., & Wetterich, C. 2013, Phys. Rev. D, 87, 083009

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016a, A&A, 594, A13

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016b, A&A, 594, A14

Pogosian, L., & Silvestri, A. 2016, Physical Review D, 94

Pratten, G., Munshi, D., Valageas, P., & Brax, P. 2016, Phys. Rev. D, 93, 103524

Pujolàs, O., Sawicki, I., & Vikman, A. 2011, Journal of High Energy Physics, 11, 156

Rassat, A., & Refregier, A. 2012, A&A, 540, A115

Ratra, B., & Peebles, P. J. E. 1988, Phys. Rev. D, 37, 3406

Reischke, R., Spurio Mancini, A., Schäfer, B. M., & Merkel, P. M. 2018, ArXiv e-prints, arXiv:1804.02441

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, The Astronomical Journal, 116, 1009

Sachs, R. K., & Wolfe, A. M. 1967, ApJ, 147, 73

Sakstein, J., & Jain, B. 2017, ArXiv e-prints, arXiv:1710.05893

Santaló, L. A. 1976, Integral Geometry and Geometric Probability

Sato, M., Hamana, T., Takahashi, R., et al. 2009, ApJ, 701, 945

Saunders, W., Sutherland, W. J., Maddox, S. J., et al. 2000, Monthly Notices of the Royal Astronomical Society, 317, 55

Sawicki, I., & Bellini, E. 2015, Phys. Rev. D, 92, 084061

Schäfer, B. M. 2018, ArXiv e-prints, arXiv:1807.06269

Schäfer, B. M., & Heisenberg, L. 2012, MNRAS, 423, 3445

Schmalzing, J., & Buchert, T. 1997, ApJ, 482, L1

Schmalzing, J., & Gorski, K. M. 1998, MNRAS, 297, 355

Schmalzing, J., Kerscher, M., & Buchert, T. 1996, in Dark Matter in the Universe, ed. S. Bonometto, J. R. Primack, & A. Provenzale, 281

Schneider, A., & Teyssier, R. 2015, J. Cosmol. Astropart. Phys., 12, 049

Schneider, M. D., Hogg, D. W., Marshall, P. J., et al. 2015, ApJ, 807, 87

Schneider, P., van Waerbeke, L., Kilbinger, M., & Mellier, Y. 2002, Astronomy & Astrophysics, 396, 1

Schneider, R. 1993, Convex bodies: the Brunn-Minkowski theory

Seljak, U., & Zaldarriaga, M. 1996, ApJ, 469, 437

Semboloni, E., Hoekstra, H., & Schaye, J. 2013, MNRAS, 434, 148

Semboloni, E., Hoekstra, H., Schaye, J., van Daalen, M. P., & McCarthy, I. G. 2011, MNRAS, 417, 2020

Shadmi, Y. 2017, ArXiv e-prints, arXiv:1708.00772

Simon, P., Hetterscheidt, M., Schirmer, M., et al. 2007, Astronomy & Astrophysics, 461, 861

Simon, P., King, L. J., & Schneider, P. 2004, A&A, 417, 873

Singh, S., Mandelbaum, R., & Brownstein, J. R. 2017, Monthly Notices of the Royal Astronomical Society, 464, 2120

Skilling, J. 2006, Bayesian Anal., 1, 833

Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311

Spergel, D., Gehrels, N., Breckinridge, J., et al. 2013, ArXiv e-prints, arXiv:1305.5422

Spurio Mancini, A., Reischke, R., Pettorino, V., Schäefer, B. M., & Zumalacárregui, M. 2018a, ArXiv e-prints, arXiv:1801.04251

Spurio Mancini, A., Taylor, P. L., Reischke, R., et al. 2018b, ArXiv e-prints, arXiv:1807.11461

Takada, M., & Hu, W. 2013, Phys. Rev. D, 87, 123504

Takada, M., & Jain, B. 2003a, MNRAS, 340, 580

Takada, M., & Jain, B. 2003b, MNRAS, 344, 857

Takada, M., & Jain, B. 2004, Monthly Notices of the Royal Astronomical Society, 348, 897

Takada, M., & Jain, B. 2009, Monthly Notices of the Royal Astronomical Society, 395, 2065

Takada, M., & White, M. 2004, The Astrophysical Journal Letters, 601, L1

Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, ApJ, 761, 152

Taruya, A., Takada, M., Hamana, T., Kayo, I., & Futamase, T. 2002, ApJ, 571, 638

Taylor, P. L., Kitching, T. D., & McEwen, J. D. 2018a, ArXiv e-prints, arXiv:1804.03667

Taylor, P. L., Kitching, T. D., McEwen, J. D., & Tram, T. 2018b, ArXiv e-prints, arXiv:1804.03668

Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, ApJ, 480, 22

Tegmark, M., Eisenstein, D., Strauss, M., et al. 2006, Physical Review D, 74, doi:10.1103/PhysRevD.74.123507

Thornton, R. J., Ade, P. A. R., Aiola, S., et al. 2016, ApJS, 227, 21

Tomita, H. 1986, Progress of Theoretical Physics, 76, 952

Troxel, M. A., & Ishak, M. 2015, Phys. Rep., 558, 1

van Daalen, M. P., Schaye, J., Booth, C. M., & Dalla Vecchia, C. 2011, MNRAS, 415, 3649

van Uitert, E., Joachimi, B., Joudaki, S., et al. 2018, MNRAS, 476, 4662

Van Waerbeke, L., Mellier, Y., Erben, T., et al. 2000, A&A, 358, 30

Vogelsberger, M., Genel, S., Springel, V., et al. 2014, MNRAS, 444, 1518

Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2013, Phys. Rep., 530, 87

Wetterich, C. 1988, Nuclear Physics B, 302, 668

Wetterich, C. 2004, Physics Letters B, 594, 17

Will, C. M. 2014, Living Reviews in Relativity, 17, 4

Winitzki, S., & Kosowsky, A. 1998, New A, 3, 75

Woodard, R. 2007, in Lecture optnotes in Physics, Berlin Springer Verlag, Vol. 720, The Invisible Universe: Dark Matter and Dark Energy, ed. L. Papantonopoulos, 403

Woodard, R. P. 2015, ArXiv e-prints, arXiv:1506.02210

Xavier, H. S., Abdalla, F. B., & Joachimi, B. 2016, MNRAS, 459, 3693

Yoshiura, S., Shimabukuro, H., Takahashi, K., & Matsubara, T. 2017, MNRAS, 465, 394

Zeldovich, Y. B. 1971, Soviet Journal of Experimental and Theoretical Physics Letters, 14, 180

Zhang, T.-J., Yuan, Q., & Lan, T. 2009, New A, 14, 507

Zieser, B., & Merkel, P. M. 2016, MNRAS, 459, 1586

Zumalacárregui, M., Bellini, E., Sawicki, I., Lesgourgues, J., & Ferreira, P. G. 2017, J. Cosmol. Astropart. Phys., 8, 019

Zumalacárregui, M., & García-Bellido, J. 2014, Phys. Rev. D, 89, 064046

Zumalacárregui, M., Koivisto, T. S., & Mota, D. F. 2013, Phys. Rev. D, 87, 083010

Zuntz, J., Paterno, M., Jennings, E., et al. 2015, Astronomy and Computing, 12, 45

Acknowledgments

First of all, I wish to thank Dr. Valeria Pettorino and Prof. Björn Malte Schäfer for their supervision in my PhD studies and for reviewing this thesis. The motivations for my gratitude are multiple, but in particular I would like to thank them here for the many conversations in which I learnt how a scientist thinks, how a scientist works, and how one gets along in science. I also would like to thank Prof. Hans-Christian Schultz-Coulon and Prof. Susanne Westhoff for kindly agreeing to be in the board of examiners.

I am grateful for funding received from the graduate college *Astrophysics of cosmological probes of gravity* by Landesgraduiertenakademie Baden-Württemberg. I wish to thank the German DFG Excellence Initiative for funding a mobility program, through the grant *Gravity on the largest scales and the cosmic large-scale structure*, within which some of the works presented in this thesis originated. I am most grateful to Prof. Alan Heavens, Dr. Benjamin Joachimi and Dr. Thomas Kitching for hosting me at Imperial College London, University College London and the Mullard Space Science Laboratory, respectively, and for making my stay enjoyable and fruitful. I also wish to thank Peter Taylor at MSSL for collaborating on the code comparison project. Many thanks to Prof. Hiranya Peiris at UCL for allowing me to use the Hypatia cluster and to Dr. Edd Edmondson for the technical support with it.

I wish to thank the Institute for Theoretical Physics for hosting me for three years, in particular Prof. Luca Amendola for interesting conversations on history and music. Thanks in particular to the secretaries at the ITP for their help with the 'paperwork' required throughout these three years. Thanks also to Dr. Heinzelmann and Prof. Klevansky at the Heidelberg Graduate School of Fundamental Physics (HGSFP) for their help and the funding provided by the HGSFP to allow me to attend many interesting conferences and schools. I wish to express my gratitude to people at the Institute for Theoretical Astrophysics, in particular to Prof. Matthias Bartelmann and Dr. Matteo Maturi, for kindly hosting me in their offices for a few days in the summer, allowing me to escape from my 50°C-office.

I would like to thank Robert Reischke and Fabian Köhlinger for sharing with me the harshness of programming and I wish them all the very best in particular for their future with their partners. I also would like to thank Miguel Zumalacárregui who has been an invaluable help with everything concerning `hi_class` and MontePython. Thanks to Britta Zieser and Philipp Merkel for having initiated the work on the Levin integration, which this dissertation has built upon. I would also like to express my gratitude to Nelson Lima and Javier Rubio for the nice words received in moments of need.

This thesis is dedicated to the memory of some people who are not with us anymore; Tham is one of them and I wish to remind him, wherever he is now, that I will never forget him.

As far as friends outside the academic environment are concerned, the list of people to thank would be too long. I will only mention here Johannes and Matteo from Heidelberg, and the ever present Kristian and Enzo at home.

Infine alla mia famiglia, il grazie più grande: a partire dai miei genitori e includendo nonne, zii e cugini. Grazie per il vostro sostegno e incommensurabile amore.