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

Ana Marta Pinho

born in: Paços de Ferreira, Portugal
Oral examination: May 13th, 2020

Assessing deviations to the ΛCDM model: the importance of model-independent approaches

Referees: Prof. Dr. Luca Amendola
Prof. Dr. Matthias Bartelmann

Beurteilung von Abweichungen zum ΛCDM Modell: Die Wichtigkeit von modellunabhängigen Ansätzen

Die bemerkenswerte Verbesserung der Genauigkeit von kosmologischen Daten in den letzten Jahren hat zu strikten Einschränkungen an die Parameter des ΛCDM-Modells geführt. Zum Beispiel haben die Gravitationswellenereignisse bestätigt, dass die Gravitationsgeschwindigkeit sehr nahe an der Lichtgeschwindigkeit ist. Dieses Ergebnis hat einige modifizierte Gravitationsmodelle ausgeschlossen. Die verbleibenden, zugelässigen Modelle sind im Datenvergleich fast ununterscheidbar von dem standardmäßigen ΛCDM-Modell. Eine Herangehensweise, um die Modelle zu unterscheiden, ist die Benutzung von dafür gebauten Schätzfunktionen, wie zum Beispiel eine modellunabhängige Bestimmung des anisotropischen Spannungsparameters. Von dieser Schätzfunktion kann man bestimmen, ob die Näherung der idealen Flüssigkeit in der allgemeinen Relativitätstheorie gültig ist, was jede Theorie testet, die diese Näherung beinhaltet. In dieser Dissertation benutzen wir die neuesten verfügbaren Daten von einigen kosmologischen Beobachtungen und drei verschiedene Methoden, um den anisotropischen Spannungsparameter zu rekonstruieren. Unsere Schlussfolgerungen hängen leicht von der Datenrekonstruktionsmethode ab, stimmen aber auf dem 2σ Level überein. Der resultierende, anisotropische Spannungsparameter könnte Standardgravitation auf 1-2σ ausschließen, abhängig von der Methode und der Rotverschiebung.

Eine wichtige Frage ist, wie die Informationsmenge in Daten bestimmt werden kann. Idealerweise möchten wir den Überzeugtheitsgrad in ΛCDM quantifizieren. In dieser Dissertation versuchen wir diese Fragen mit Informationstheorie zu beantworten. Wir berechnen die Entropie von Modellparametern für spezifische, kosmologische Beobachtungen. Wir vergleichen diesen Zugang mit der weitverbreiteten Fisher-Matrix, die typischerweise für Prognosen von zukünftigen, großräumigen Strukturbeobachtungen berechnet wird. Die Unsicherheiten von jedem Parameter werden erhalten und dadurch ist die Qualität der Daten üblicherweise mit den Eigenschaften der Fisher-Matrix verbunden. Informationsentropie kann auch messen, wie unterschiedliche Kombinationen von kosmologischen Beobachtungen die Modellparameter einschränken. Die gleiche Prozedur wird auf die jüngsten Datenungereimtheiten angewendet und es kann im Falle von Modellvergleichen benutzt werden. Wegen seiner analytischen Ausdrücke kann Informationsentropie äußerst nützlich sein, falls eine Gauß-Verteilung angenommen wird. Eine Verallgemeinerung zu allgemeinen Distributionen ist möglich.

Die Hauptbotschaft dieser Dissertation ist, dass neue Wege zum Testen von Gravitation nötig sind, insbesondere mit der abnehmenden Unsicherheit in kosmologischen Datensätzen und dem Auftauchen von Ungereimtheiten zwischen jenen Datensätzen. Wir müssen besser zwischen konkurrierenden Theorien unterscheiden können. Dies kann mit Schätzfunktionen erreicht werden, die sich nicht auf ein spezifisches Szenario verlassen. Eine andere Möglichkeit ist eine andere Perspektive auf den statischen Rückschluss zu finden, was insbesondere nützlich ist, um die Annahmen in der Datenreduktion neu zu bewerten.

Assessing deviations to the ΛCDM model: the importance of model-independent approaches

The remarkable improvement of the accuracy of cosmological data in the last years has provided tight constraints on the parameters of the ΛCDM model. For example, gravitational-wave events have confirmed that the speed of gravity is very close to the speed of light. This result has ruled out several modified gravity models. The remaining allowed models are nearly indistinguishable from the standard ΛCDM in data comparison. One approach to discriminate models is to use estimators built for that purpose, as, for example, a model-independent determination of the anisotropic-stress parameter. From this estimator, one can infer if the perfect fluid approximation done in General Relativity is valid, testing any theory that includes this approximation. In this dissertation, we use the latest available data from several cosmological probes and three different methods to reconstruct the anisotropic stress parameter in a model-independent way. Our conclusions depend mildly on the data reconstruction method but agree at the 2σ level. The resulting anisotropic stress may rule out standard gravity within a 1-2σ level depending on the method or redshift.

An important question is how the amount of information in the data can be measured. Ideally, we would like to quantify the degree of belief in ΛCDM. In this dissertation, we tackle these questions by using information theory. We compute the entropy of model parameters for specific cosmological probes. We compare this approach with the widely-used Fisher matrix, typically computed when forecasting future large-scale structure surveys. The uncertainties on each parameter are obtained and thus the quality of the data is usually associated with certain properties of the Fisher matrix. Information entropies can also measure how different combinations of cosmological probes constrain the parameters of a model. The same procedure is applied to the recently found data tensions, and it can be used in case of model comparison. Information entropies can be extremely useful due to its analytical expressions if a Gaussian distribution is assumed but a generalization to any distribution is possible.

The main message of this dissertation is that new ways of testing gravity are needed, specifically with the decreasing uncertainty in cosmological datasets and the appearance of discrepancies between datasets. We need to better discriminate competing theories. This can be done through estimators that should not rely on a specific scenario. Another possibility is to find a different perspective on statistical inference, which is particularly useful in order to re-evaluate the assumptions done in data reduction.

The eye sees only what the mind is prepared to comprehend.
(Robertson Davies)

Contents

List of Figures

List of Tables

List of abbreviations and acronyms

ΛCDM	Λ Cold Dark Matter
BBN	Big Bang Nucleosynthesis
BOSS	Baryon Oscillations Spectroscopic Survey
CMB	Cosmic Microwave Background
FLRW	Friedmann-Lemaître-Robertson-Walker
GC	Galaxy Clustering
HST	Hubble Space Telescope
LSS	Large-scale Structures
RSD	Redshift-Space Distortions
SDSS	Sloan Digital Sky Survey
SNIa	Supernova type Ia
TRGB	Tip of the Red Giant Branch
WL	Weak gravitational Lensing

1. Introduction

Throughout the past century, our understanding of the Universe has evolved significantly. Einstein formulated the theory of General Relativity in 1915. It brought a revolutionary new perspective of gravity with space and time as one single concept. As stated by J. A. Wheeler, spacetime tells matter how to move and matter tells spacetime how to curve. Shortly after, Slipher, Lemaître, and Hubble found evidence for an expanding universe. This was the starting point of Cosmology as we know today: a modern physics field with advanced astronomical observations. Later on, strong evidence for a "missing mass" problem suggested the need for a new species of matter, dubbed dark matter, to explain the rotation curves of galaxies. The discovery of the accelerated expansion of the Universe carried a new big challenge. It suggests the existence of dark energy component, a fluid with negative pressure that drives this accelerated phase of the Universe. Currently, observations point out that the Universe is quite well described by the ΛCDM model or the concordance model of Cosmology. Based on Einstein's General Relativity, this model is composed of three main contributions: a cosmological constant, Λ, associated with dark energy, a cold dark matter component, and ordinary matter, in a homogenous, isotropic, and flat spacetime.

The ΛCDM model has successfully explained several cosmological observations. The anisotropic temperature fluctuations from the Cosmic Microwave Background (CMB) radiation are exceptionally well fitted by this six parameter model. The accelerated expansion is usually obtained through the light curves of Supernova type Ia. These "standard candles" are well fitted by a ΛCDM cosmology. The existence of a "standard ruler" such as the Baryonic Acoustic Oscillations is a successful prediction from ΛCDM. Alongside with these observations, there are further checks done by the observed abundances of the first elements, the distribution of galaxy clusters (GC), and weak lensing (WL) that contribute to the robustness of this model.

Despite all these observational sucess of ΛCDM, there are also quite a few unsolved problems, one of them being the famous cosmological constant problem. One possible explanation for the cosmological constant Λ is to associate it with the vacuum energy. This intuitive solution implies that the fraction of dark energy observed at large scales matches the zero-point energy at the quantum level. Within reasonable assumptions, although a quantum gravity theory does not exist, these energy fractions differ by far too many orders of magnitude. This demands a careful choice of the cosmological constant, or in other words, to fine-tune it. Besides that, the nature of the cosmological constant is likewise lacking justification.

These issues have opened a whole new field of possible theories of gravity. In a simple way, one can formulate another theory by using a scalar field as dark energy instead of a cosmological constant. On the other hand, one can question the geometry of gravity and modify it. Although these alternative theories are not as sucessful as ΛCDM when compared with data, most of the current datasets do not fully rule out these other theories. This leads to a plethora of possibilities to explain the history and expansion of the Universe within the current precision of data. We are indeed in the precision era of cosmology but

not yet precise enough to confirm or rule out ΛCDM.

A few years ago, a statistical difference of 2σ was found between the cosmological parameters obtained by probing the CMB and the ones found by WL. More recently, the measured value for the Hubble expansion today given by the Supernova type Ia and Cepheid variable stars is about 4σ away from the value obtained through the last CMB survey. This particular tension carries not just the possibility of an unaccounted hidden systematical uncertainties, but also the idea of new physics as this discrepancy happens at the two extremes of the cosmic history. Although it is certainly important to explore new physics solutions for this data tension, there is another aspect that needs to be considered. Unlike to the standard candles and rules, most of these observables rely on a ΛCDM type of expansion and, therefore, the resulting data are model-dependent. Since ΛCDM is rather accepted to explain the current data, this may not sound like an issue or a wild guess. However, this could be a key point towards deepening our understanding of gravity. A completely unbiased and non-parametric approach to raw cosmological data is nearly impossible. Further data is inevitably needed but also new perspectives on this puzzle.

Here is where it becomes important to find ways of testing gravity, especially, in a model-independent way. Although it is a challenging task, to measure the properties of spacetime using the least amount of assumptions possible is a crucial step. This is the main motivation for the work presented throughout this dissertation. Instead of tackling the unanswered theoretical questions, the approach is to use the current knowledge and look for answers within both data and statistical frame. The two main subjects of this dissertation are estimators built to find deviations of the standard gravity scenario and an information-driven approach to current and future (forecasted) cosmological data.

The structure of this dissertation is described in more detail next. Chapter 2 describes specifically the state-of-the-art of cosmology. We give an overview of the standard model of ΛCDM which entails Einstein's general relativity theory, a cosmological constant Λ and a cold dark matter component. Although this theory has been generally quite successful when compared to observational data, its fine-tuning problems led to the formulation of alternative theories. We present a common generalization, the Horndeski theory, that encompasses several other models including quintessence and the Brans-Dicke theory. At the moment, several cosmological observables probe particular time frames and aspects of the universe. We characterize the technicalities of the Hubble parameter $H(z)$ and Hubble constant H_0, CMB, WL, and GC. Surprisingly, as already mentioned, the Hubble constant H_0, which expresses the expansion of the Universe, has currently different probes disagreeing on its value with high statistical significance. Additionally, there is a tension between CMB and weak lensing, visible on the (Ω_m, σ_8) plane. In order to solve these discrepancies, it is important to know the assumptions done during the data analysis.

In Chapter 3, we estimate the anisotropic stress parameter η in a model-independent way. To measure this parameter with the latest available data constitutes a test to the perfect fluid assumption done in general relativity. Consequently, it serves as a test to assess deviations to the ΛCDM scenario as it includes GR as a key component. We use data from the aforementioned probes and three data reconstruction methods with different assumptions for this work. In general, we find agreement with the standard model at a 2σ level. However, the methods can disagree at the 1σ level and even be in tension with the expected value for the ΛCDM model. Since the methods do not fully agree and tensions are known for these datasets, the presented work provides, for now, a promising framework to be applied to future better data.

While working on the aforementioned estimation, we noticed the theoretical and observational definitions of the E_G statistics. Chapter 4 outlines these definitions and the validity of its correspondence. The literature supports that these definitions are equivalent to the case of a flat ΛCDM Universe. This means that the E_G quantity is not a test of general relativity *per se* as this equivalence may not hold. We have derived the full expression and compared to the literature, although it is still work in progress.

In Chapter 5, we introduce the tools of Information Theory which can be useful in data analysis. The Fisher matrix formalism is very common to use when forecasting for future experiments. We make an analogy between these two approaches for the case of a Gaussian likelihood. Taking all the cross-correlations between the main cosmological probes - CMB and CMB lensing, WL and GC - we compute the common measures of statistical uncertainties as well as information entropy. Probe combination reduces the statistical errors likewise reduces the information entropy. We analyze the several data tensions and the role of Bayesian evidence in this comparison.

Finally, in Chapter 6 we make a general summary and outlook.

Publications

This dissertation was solely executed by the author. The results presented here were achieved through collaborations. Parts of the results are published. These publications are:

[4] **Model-independent reconstruction of the linear anisotropic stress η**
Ana Marta Pinho, Santiago Casas and Luca Amendola
Published in JCAP 1811, 027 (2018)
E-Print: arXiv:1805.00027 [astro-ph]
Comment: Chapter 3 is based on this publication.

[5] **Model-independent measures of gravity at large scales**
Luca Amendola, Ana Marta Pinho and Santiago Casas
Published in International Journal of Modern Physics A 33, 1844022 (2018)

[6] **Measuring gravity at cosmological scales**
Luca Amendola, Dario Bettoni, Ana Marta Pinho and Santiago Casas
Published in Universe 6, 2 (2020)
E-Print: arXiv:1902.06978 [astro-ph]
Comment: Part of Chapter 2 is based on this publication.

[7] **Information entropy in cosmological inference problems**
Ana Marta Pinho, Robert Reiscke and Björn Malte Schäfer
In preparation
Comment: Chapter 5 is based on this publication in preparation.

[8] **Revisiting the E_G statistics**
Luca Amendola and Ana Marta Pinho
In preparation
Comment: Chapter 4 is based on this publication in preparation.

2. Understanding the Universe: Theory and Observations

The theoretical foundation of our understanding of the Universe is General Relativity. More than a hundred years ago, Albert Einstein formulated General Relativity (GR) as an answer to the behaviour of gravity in a relativistic framework. This theory has been constantly tested ever since. After explaning the perihelion shift of Mercury, GR has passed solar system tests [9], pulsar timing tests [10], among others. The recent detection of gravitational waves from the LIGO/VIRGO collaboration [11] is another successful prediction. GR has proven to be a stable and sucessful theory but on its own it does not explain the accelerated expansion of the Universe. The evidence for a late-time acceleration [12, 13] has suggested a missing ingridient in the cosmic recipe. A cosmological constant, labeled Λ, added to the field equations of Einstein is one possible solution. Besides a component that drives the cosmic expansion, evidence was found for a missing mass in galaxies to explain their Keplerian rotation curves [14–16]. Currently, the concordance model of Cosmology has in its composition: gravity explained by GR, a cosmological constant and a dark matter component. This famous ΛCDM model explains most of the available cosmological data but it has a few theoretical unsolved issues. For that reason, several alternative theories have been formulated. Part of those alternative theories target the nature of dark energy and dark matter, others modify how gravity works.

2.1. Theoretical description of the Universe

In this chapter, we describe the principles of GR and the details of ΛCDM, the standard model of cosmology. We briefly discuss a few of the extensions of ΛCDM. The precision era of cosmology has begun, and it is possible to constrain cosmological parameters with unprecedented accuracy. We characterize a few cosmological probes such as the cosmic microwave background, weak lensing, galaxy clustering and the Hubble expansion. These probes attempt to estimate different properties of gravity at various time stamps in the cosmic history. We use the convention that Greek letters as indexes running from 0 to 3 and Latin letters as indexes running from 1 to 3. Here the speed of light c is going to be set to 1.

2.1.1. General Relativity

Albert Einstein first proposed General Relativity in 1915 [17]. However, its application to cosmology as the description of the Universe came two years later [18], also with the contribution of Willem de Sitter [19]. The field equations of Einstein connect the geometry of space and time with the matter present in it. To obtain the equations of motion, one must vary the Einstein-Hilbert action defined as

$$S = \frac{1}{16\pi G}\int \mathrm{d}^4x\sqrt{-g}(R - 2\Lambda) + S_m\,, \tag{2.1}$$

with respect to the metric field $g_{\mu\nu}$. In this expression, G is the gravitational constant and Λ is the cosmological constant. The volume element is $\mathrm{d}^4x\sqrt{-g}$ where g corresponds to the determinant of the metric. R is the Ricci scalar obtained from the contraction of the Riemann tensor $R^{\alpha}_{\beta\mu\nu}$ which are both functions of the metric. The matter action is represented by S_m. The field equations are given by

$$G_{\mu\nu} + g_{\mu\nu}\Lambda = 8\pi G T_{\mu\nu}\,, \tag{2.2}$$

where $G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ is the Einstein tensor, and the energy-momentum tensor $T_{\mu\nu}$ corresponds to

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S_m}{\delta g_{\mu\nu}}\,. \tag{2.3}$$

These field equations state that matter affects the structure of the spacetime and spacetime defines how matter moves. They are coupled non-linear differential equations of the metric. One uses symmetry assumptions on the spacetime to solve Eq. (2.2). Focusing on the cosmological solution to these equations, the starting point of modern cosmology is the cosmological principle. It states that at sufficiently large scales, the Universe is homogenous and isotropic. These are important assumptions that have been partially tested but they are very difficult to verify or falsify. The solution that fulfills these assumptions while solving Einstein's field equations in Eq. (2.2), is the Friedmann-Lemaître-Robertson-Walker (FLRW) metric,

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu} = -\mathrm{d}t^2 + a^2(t)\left[\frac{\mathrm{d}r^2}{1-Kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\mathrm{d}\phi^2)\right]\,, \tag{2.4}$$

where a is the scale factor and t the cosmic time coordinate. The spatial component has a radial coordinate, r, with the polar angle, θ, and the azimuthal angle, ϕ. K denotes the curvature, where the Universe is flat if $K = 0$, closed if $K = 1$, and open if $K = -1$. The metric $g_{\mu\nu}$ is an important object as it defines the size of spacetime intervals $\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu}$. It is also invariant under diffeomorphisms. Since any particle and field is coupled universally to the metric g, the equations of motion do not depend on the choice of coordinates.

2.1.2. Dynamics of the Universe

The dynamics of the Universe is described as a perfect fluid. Resembeling the cosmological principle idea, the perfect fluid assumption is justified by the fact that no viscous or dissipative forces should exist in the background at very large scales. This assumption allows to rewrite the energy-momentum tensor as

$$T^{\mu\nu} = (\rho + p)u^{\mu}u^{\nu} + pg^{\mu\nu}\,, \tag{2.5}$$

with the pressure $p = p(t)$ and the density $\rho = \rho(t)$, which due to homogeneity, are both only functions of the cosmic time t. Also, these functions are the sum of all matter and radiation components of the Universe. The background dynamics obeys the Friedmann equations. These are obtained by taking the (00) and (ii) components of the field equations in Eq. (2.2), with an FLRW metric, meaning,

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{K}{a^2}\,, \tag{2.6}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + p) + \frac{\Lambda}{3}\,, \tag{2.7}$$

where the dot denotes the derivative with respect to the cosmic time t. The Friedmann equations Eq. (2.7) can be reduced to one if an equation of state

$$P = w\rho\,, \tag{2.8}$$

is provided. The proportionality constant depends on the species. It corresponds to $w_m = 0$ for non-relativistic matter, $w_r = \frac{1}{3}$ for relativistic matter, and $w_0 = -1$ for the cosmological constant. In order to express each species density in a dimensionless quantity, we introduce the critical density $\rho_{\rm crit} = 3H_0^2/(8\pi G)$ for a flat Universe. The density parameters are thus defined as

$$\Omega_i(t) \equiv \frac{\rho_i(t)}{\rho_{\rm crit}}\,, \tag{2.9}$$

where $\rho_i(t)$ is the density function for each species that evolve differently with the scale factor. These density parameters obey the closure relation $\sum_i \Omega_i = 1$. The different dependences with the scale factor hint at the different epochs in cosmic history where a specific species dominated the expansion and evolution of the Universe. With these definitions, the expansion of the Universe is described by the Hubble function,

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)}\,. \tag{2.10}$$

The Hubble function is an important concept that plays a role in the definition of cosmological distances. The comoving distance is defined as

$$\chi(z) = -\int_0^z \frac{{\rm d}z'}{H(z')}\,, \tag{2.11}$$

This relates the redshift measured today to the distance of a source with redshift measured on a spatial hypersurface. Assuming a flat Universe, one can write,

$$D_A(z) = (1+z)\chi(z)\,, \qquad D_L(z) = \frac{D_A(z)}{(1+z)^2}\,, \tag{2.12}$$

where $D_A(z)$ is angular diameter distance and $D_L(z)$ is the luminosity distance. Since these distances are directly related to the cosmological parameters, it is possible to probe the content of the Universe. One example is the use of Supernova type Ia to measure such distances, as used when finding the accelerated expansion of the Universe [12, 13]. The Supernova type Ia are standardizable candles, which we describe in more detail in Sec. 2.2.4 and other ways of measuring the expansion of the Universe in Sec. 2.2.4.

2.1.3. The standard ΛCDM model

In the last decades, the outstanding improvement of precision in several cosmological probes led to the formulation of the concordance model. The ΛCDM model relies on the fact that the laws of physics are the same at any point in the Universe. GR is the description of gravity and the cosmological principle is assumed. It further assumes that the expansion of the Universe occurs from an early hot dense state and five key components.

These components are a dark energy that drives the current accelerated expansion of the Universe. This is further described by a cosmological constant that is associated with a vaccum energy. Besides dark energy, there is a dark matter component that is pressureless and interacts only gravitationally with normal matter. Normal atomic matter, photons and nearly massless neutrinos are also part of the composition. The curvature of the Universe is taken to be flat or very small. Finally, density fluctuations are neraly scale-invariant, adiabatic and Gaussian. Given this composition, the background evolution described by the Hubble function $H(z)$ can be written explicitly as

$$H^2(z) = H_0 \left[\Omega_c (1+z)^3 + \Omega_b (1+z)^3 + \Omega_r (1+z)^4 + \Omega_{\rm DE} \exp\left(\int_0^z \mathrm{d}z' \frac{3(1+w_{DE}(z'))}{1+z'} \right) \right], \tag{2.13}$$

with the redshift defined as $z \equiv= \frac{1}{a} - 1$. The equation of state $w_{\rm DE}$ denotes a general dark energy componentto allow for other definitions.

This discription has been successful in explaining accurately the cosmic microwave background radiation which we discuss in Sec. 2.2.1. The many cosmological probes such as weak lensing or galaxy clustering are also well suited with a ΛCDM-like Universe. However, there are a few problems yet to solve. The use of a cosmological constant as the driving mechanism of the late-time expansion is not fully justified, despite the fact that it explains well most observations. The main motivation for this choice is the idea of associating Λ with the vaccum energy. Given the composition of the Universe, we can estimate the density of dark energy as roughtly,

$$\rho_\Lambda = \frac{\Lambda}{8\pi G} \approx 10^{-47} \mathrm{GeV}^4 \,. \tag{2.14}$$

Let us now assume that the vaccum energy stems from a zero point energy of fields of mass m, and momentum k. Its energy reads $E = \frac{1}{2}\sqrt{k^2+m^2}$ and the energy density is the sum of all momenta up to some cut-off scale. If this cut-off scale is roughly the Planck mass $m_{\rm Pl}$, then we arrive to

$$\langle \rho_{\rm vac} \rangle \simeq 10^{74} \mathrm{GeV}^4 \,. \tag{2.15}$$

These estimates differ by 121 orders of magnitude. This energy comparison is the famous cosmological constant problem and the reason why one needs to carefully choose the cosmological constant, or in order words, to fine-tune it. Such a calculation is based on a very rough estimation from quantum field theory. For example, one can use a different regularization scheme resulting in $\langle \rho_{\rm vac} \rangle \approx 10^{10} \mathrm{GeV}^4$, which lowers the difference significantly [20]. However, to make a correct estimation of this vaccum energy, one needs to have a quantum gravity theory, which unfortunately does not yet exist [21, 22].

There is another interesting question left to answer. The density parameter today for a cosmological constant Ω_Λ and the one of matter $\Omega_{m,0}$ are surprisingly of the same order of magnitude. The proximity of these values raises the suspicion that we live in a special time of cosmic history, which seems unlikely. In other words, the redshift where the two densities coincide can be computed by

$$z_c = \left(\frac{\Omega_\Lambda}{1-\Omega_\Lambda} \right)^{\frac{1}{3}} - 1 \overset{\text{for } \Omega_\Lambda = 0.7}{\approx} 0.3 \,. \tag{2.16}$$

This means that very recently dark energy started dominating over matter and if that would have happened much longer ago, we would not see the current accelerated expansion.

Such problems have led to the formulation of alternative theories. To solve the cosmological constant problem, usually the question changes towards the nature of dark energy. Either by introducing a scalar field, whose dynamics drives the accelerated expansion, or modifying gravity, the problem remains unsolved but a plenthora of alternative theories have been formulated. We name a few of those next. The coincidence and cosmological problems are somewhat solved independently. In other words, the coincidence problem is still present in many theories that address the cosmological constant problem as dark energy explanation. Regarding the coincidence problem, one example solution are the tracker models, where the behaviour of the evolution of each density is similar but with an offset independent of initial conditions. This does not fully solve the problem but it mitigates the problem while building another framework [23]. Another example is the use of the anthropic principle. This principle stems from the idea that the existence of life on Earth is a needed consideration in physical theories. In more physical terms, it can be rephrased as the physical constants must have the values compatible with our existence as observers [24]. This explanation can be applied to both problems but it is not generally accepted.

2.1.4. Beyond the standard ΛCDM model

So far we have described the GR framework and the ΛCDM model while describing a few problems that need to be addressed. We now deal with some of the alternative theories to the ΛCDM model. These are summarized to two main categories: dark energy and modified gravity. A good criteria to make this distinction is proposed in [25] (see also see [26] for a more extensive review).

The dark energy solutions often involve the addition of a scalar field that adds an extra dynamical degree of freedom. This scalar field can be universally coupled to all matter species or to only one specifically, giving rise to the subcategories of universally-coupled theories and non-universally coupled theories. Examples of universally coupled theories are Quintessence[27–31], Horndeski theories [32–34], and effective field theories [35, 36]. The non-universal coupled theories usually let baryons remain uncoupled due to local tests. From these theories, some examples are the coupled dark energy model [37, 38] and growing neutrino quintessence [39, 40].

Modified gravity theories are associated with general modifications of GR, although any new form of matter such as the dark energy models aforementioned influences the geometry. These modifications can take the form of extra dimensions [41] or violate the Lorentz invariance [42]. Also, one can include two metrics instead of one [43], make the graviton massive [44], or both [45]. In this dissertation we would like to discuss briefly the Horndeski theory as one example of an alternative theory to the ΛCDM model.

Horndeski theory

The Horndeski theory [32] is the most general theory of a scalar field coupled to a metric with second-order equations of motion without ghost instabilities. First formulated in 1974, it only became more popular about 30 years later [34]. What started as a mathematical exercise turned out to be extremly useful to set a general structure that encapsulates many other theories, following an effective field theory approach. There is an unique action that only adds one extra dynamical scalar field while respecting homogeneity, isotropy and

the Weak Equivalence Principle. We define this action in the Jordan frame as

$$S = \int \mathrm{d}x^4\sqrt{-g}\Big[\frac{M_{\rm Pl}^2}{2}[1+\Omega(\tau)]R + \Lambda(\tau) - a^2c(\tau)\delta g^{00} + \frac{M_2^4(\tau)}{2}(a^2\delta g^{00})^2 \\ - \bar{M}1^3(\tau)2a^2\delta K\mu^\mu\delta g^{00} - \frac{\bar{M}2^2(\tau)}{2}(\delta K\mu^\mu)^2 - \frac{\bar{M}3^2(\tau)}{2}\delta K\nu^\mu\delta K^\nu_\mu + a^2\frac{\hat{M}^2(\tau)}{2}\delta g^{00}\delta R^{(3)} \\ + m_2^2(\tau)(g^{\mu\nu}+n^\mu n^\nu)\partial_\mu(a^2g^{00})\partial_\nu(a^2g^{00}) + \mathcal{L}m(g\mu\nu,\Psi_m)\Big]\,, \tag{2.17}$$

written in terms of the conformal time τ. Spacetime has been foliated for this action, where $n_\mu \equiv \frac{\partial_\mu\phi}{\sqrt{-(\partial\phi)^2}}$ is the preferred direction of time. Without going into deep details, we would like to point out that this action is a function of nine free functions of time. Once defininining these free functions of time, we can recover the Horndeski theory or theories that go beyond that [46].

For a ΛCDM background and fixed $\Omega(\tau)$, the functions $\Lambda(\tau)$ and $c(\tau)$ are also fixed [47]. Then, to remove theories with third-order spatial derivatives, one can also set $\bar{M}_2^2(\tau) = -\bar{M}_3^2(\tau)$ and $m_2^2(\tau) = 0$. This leaves five free functions and $H(\tau)$ related to the background cosmology. A commonly-used simplification of five free functions is the mapping done by the $\alpha_i(\tau)$ functions defined in [48]. These functions are

$$\begin{aligned} M_*^2(\tau) &= M_{\rm Pl}\Omega(\tau) + \bar{M}_2^2(\tau)\,,\\ M_*^2(\tau)H(\tau)\alpha_M(\tau) &= M_{\rm Pl}\dot{\Omega}(\tau) + \dot{\bar{M}}_2^2(\tau)\,,\\ M_*^2(\tau)H^2(\tau)\alpha_K(\tau) &= 2c(\tau) + M_2^4(\tau)\,,\\ M_*^2(\tau)H(\tau)\alpha_B(\tau) &= M_{\rm Pl}\dot{\Omega}(\tau) - \bar{M}_1^3(\tau)\,,\\ M_*^2(\tau)\alpha_T(\tau) &= -\bar{M}_2^2(\tau)\,,\\ M_*^2(\tau)\alpha_H(\tau) &= 2\hat{M}^2(\tau) - \bar{M}_2^2(\tau)\,, \end{aligned} \tag{2.18}$$

with $M_{\rm Pl}$ as the reduced Planck mass and M_*^2 as the effective Planck mass. These $\alpha_i(\tau)$ functions further provide an insight into the physics. α_M stems from mass run rate as the effective Planck mass changes and there is anisotropic stress. Also, α_K is labeled due to kineticity, present in quintessence or k-essence theories. α_B is associated to braiding causing clustering dark energy. The function α_T translates the tensor speed excess which produces deviations of the speed of gravitational waves with respect to the speed of light. Finally, the α_H accounts for the beyond Horndeski theories. These equations have been implemented in the `hi-CLASS` code [49] which is the Horndeski extension of the popular Boltzmann code called `CLASS` [50]. The `hi-CLASS` code is the foundation of the Fisher matrix code we use later in Chapter 5.

Recently, gravitational waves from a neutron star merger (GW170817) and its electromagnetic counterpart (GRB170817) were detected by the LIGO/VIRGO collaboration and many other telescopes [51–55]. One of the main results is that the difference between the speed of gravity c_T and the speed of light c is of about 10^{-15} [51]. Among other consequences of this measurement [56, 57], it strongly constraints a few of the above $\alpha_i(\tau)$ functions [58–61], in particular the α_T function, and thus rules out several models. However, to rule out these parts of the Horndeski theory, one needs to assume no extreme fine-tuning, the absence of attractors as expressed by [62], and an universal coupling between matter and the scalar field.

2.1.5. Modified gravity parametrizations

We do not address perturbation theory in great details but we briefly discuss another take on the metric that has that motivation. Typically to go beyond the background evolution, one starts by perturbing the metric and split it between a background contribution $\bar{g}_{\mu\nu}$ and a perturbation contribution $h_{\mu\nu}$. From the perturbed metric, one can identify scalar, vector and tensor terms from which the scalar are the relevant ones for structure formation. According to linear perturbation theory, if we chose the conformal Newtonian gauge, the line-element can be writen as

$$\mathrm{d}s^2 = -(1+2\Psi)\mathrm{d}t^2 + a^2(1-2\Phi)\mathrm{d}x^2\,, \tag{2.19}$$

where Ψ and Φ are the Bardeen potentials that are gauge-invariant (see [63] for a pedagogical introduction).

In many dark energy or modified gravity theories, the standard linear perturbation equations are not valid, and thus the Bardeen potentials provide different values than the ones from GR [24, 64]. For this reason it is common to parametrize the modifications of gravity. We focus on the particularly useful definitions that set the ground for the work presented in Chapter 3. The gravitational slip $\eta(z,k)$ is defined as

$$\eta(z,k) \equiv -\frac{\Phi(z,k)}{\Psi(z,k)}\,, \tag{2.20}$$

which is effective observable [65] that we later explore in Chapter 3. This definition is associated with modifications of GR but it can be applied to many different models [25, 29, 66].

We introduce also introduce a new function, $\mu(z,k)$, as a modification of the Poison equation in Fourier space,

$$-k^2\Psi \equiv 4\pi G H(z)^{-2}\mu(k,z)\rho_m(z)\delta_m(z,k)\,, \tag{2.21}$$

where k is the scale in terms of the cosmological horizon, $k = k_\chi(1+z)/H(z)$. ρ_m corresponds to the background average matter density and δ_m is the matter density contrast. Relativistic particles and radiation are neglected here since they are negligible at late times. The gradient of Ψ accelerates non-relativistic particles resulting in deviations of gravitational clustering. These deviations are encoded in the $\mu(z,k)$ function.

Regarding weak lensing observables, one also introduces the function $\Sigma(z,k)$ to account for deviations in the lensing potential $\Phi+\Psi$. This turns the respective Poisson equation into

$$-k^2(\Psi-\Phi) \equiv 4\pi G H(z)^{-2}\Sigma(k,z)\rho_m(z)\delta_m(z,k)\,. \tag{2.22}$$

If now we take the ratio between the Eq. (2.21) and Eq. (2.22) and a little bit of algebra, we can find a relation between all these definitions by

$$\Sigma(z,k) = \mu(z,k)(1+\eta(z,k)) \tag{2.23}$$

These definitions have a correspondence in GR where $\eta = 1$, $\mu = 1$, and $\Sigma = 2$.

In Chapter 3, we describe how to estimate Eq. (2.20) in a model-independent way [65], without assuming initial conditions of the Universe, neither on the primordial power spectrum, nor on the nature of dark matter or the details of galaxy bias. Unlike this estimation, one cannot estimate the matter background density or the matter overdensity

in a model independent way. For this reason, the same approach does not apply to a quantity like $\mu(k, z)$. In general, $\eta \neq 1$ hints at a deviation from standard gravity or a form of dark energy that cannot be approximated by a perfect fluid. However, in the limit of large (but still sub-horizon) scales and provided that the theory does contain at least one mass scale besides the Planck mass [67], one obtains $\eta = 1$.

2.1.6. Structure formation

The existence of structures in the Universe such as galaxy clusters, galaxies or the solar system would not be possible if the Universe was perfectly homogeneous. Such structures formed from small density inhomogeneities that grow through gravitational interaction. The question now is where do these fluctuations arise from and the most accepted answer is inflation [68] (see [69] for a review). In this period, quantum fluctuations of the inflaton field are the seeds of the primordial density fluctuations in a exponentially expanding Universe. This mechanism further solves two problems that the Cosmic Microwave Background observations suggest: the flatness problem and the horizon problem.

The flatness problem states that if there is any curvature deviation from a flat geometry in the early Universe, that deviation would evolve to be a large curvature today, seen through the curvature density parameter, Ω_k. Since observationally $\Omega_k \approx 0$ today, it is required that $\Omega_k \sim 10^{-27}$ at early times to satisfy the closure relation of the density parameters. This early-time accelerated expansion of the Universe naturally drives Ω_k to zero due to the decrease of the comoving Hubble radius.

The horizon problem arises from the pattern present in the temperature maps of the cosmic microwave background. As we discuss in more detail in Sec. 2.2.1, the cosmic microwave background is a perfect black body radiation that was in thermal equilibrium before recombination. To justify the causal disconnected patches seen in the CMB, these patches must have been in contact before. Assuming it lasts long enough, an early-time accelerated expansion could thus drive the comoving horizon to be smaller and thus match the horizon size at recombination which is of the order of a few degrees on the sky.

Focusing now on how structures form, let us consider a scalar random field, as for instance the density contrast,

$$\delta(\vec{x}, t) \equiv \frac{\rho(\vec{x}, t) - \bar{\rho}(t)}{\rho(\vec{x}, t)} \,, \tag{2.24}$$

where $\rho(\vec{x}, t)$ is the density perturbation and $\bar{\rho}(t)$ is the mean cosmic background density. The density contrast $\delta(\vec{x}, t)$ is called statistically homogenous if all its moments are invariant under spatial translations. Also, it is called statistically isotropic if all its moments are invariant under spatial rotations. The density fluctuations δ are very small, $\delta \ll 1$. They have mean $\langle\delta\rangle = 0$ and we now define their variance. But first, we define the two-point correlation $\xi(\vec{r})$ between two fluctuations at $\vec{x}$ and $\vec{y}$ separated, which is written as

$$\xi(\vec{r}) \equiv \langle\delta(\vec{r})\delta(\vec{y} - \vec{r})\rangle \,, \tag{2.25}$$

where $\vec{r} = \vec{y} - \vec{x}$ and the brakets denote an ensamble average. This correlation function only depends on the norm r due to the statistical homogeneity and isotropy that we defined before. $\xi(\vec{r})$ contains the information about how likely it is to find another point within a distance r from the starting position. This quantity is very useful for the statistical

description of these small fluctuations. We now introduce the Fourier transform of the density field $\delta(\vec{k})$ in n dimensions defined as

$$\delta(\vec{x}) = \int \frac{\mathrm{d}^n k}{(2\pi)^n} \delta(\vec{k}) \exp(i\vec{k}\cdot\vec{x})\,, \qquad \delta(\vec{k}) = \int \mathrm{d}^n x\, \delta(\vec{x}) \exp(-i\vec{k}\cdot\vec{x})\,, \tag{2.26}$$

and define the variance of the fluctuations between two k-modes as

$$\langle \delta(\vec{k}_1)\delta(\vec{k}_2)\rangle = \int \mathrm{d}x\, \mathrm{d}y\, \langle \delta(\vec{x})\delta(\vec{y})\rangle \exp(-i\vec{k}_1\cdot\vec{x} + i\vec{k}_2\cdot\vec{y})\,. \tag{2.27}$$

From the correlation function definition $\xi(\vec{r})$, we introduce the power spectrum as

$$P(k) = \int \mathrm{d}^r \xi(\vec{r}) \exp(-i\vec{k}\cdot\vec{r})\,, \tag{2.28}$$

which means that the variance of the fluctuations is the Fourier transform of the correlation function,

$$(2\pi)^3 P(k)\delta_D(\vec{k}-\vec{k}') \equiv \langle \delta(\vec{k})\delta^*(\vec{k}')\rangle\,, \tag{2.29}$$

where δ_D is the Dirac delta function. The higher moments of the power spectrum are defined as

$$\sigma_j^2 \equiv \int_0^\infty \frac{\mathrm{d}k}{2\pi^2} k^{2j+2} P(k) \tag{2.30}$$

where σ_0^2 is the variance. It is often used a smoothing function to select a certain scale R when searching for objects with a specific size. This can be included in the power spectrum via a multiplication in Fourier space,

$$\sigma_R^2 = \int_0^\infty \mathrm{d}k \frac{k^2}{2\pi^2} P(k) W_R^2(k)\,, \tag{2.31}$$

where $W_R(k)$ is the normalized weight function. Typical choices of this function have a Gaussian shape or a top-hat filter. The two-point correlation can be generalized to a n-point correlation function.

To describe how dark matter perturbations rise and evolve, we need to introduce the Poisson equation

$$\Delta\Phi = 4\pi G\rho\,, \tag{2.32}$$

where Φ is the Newtonian potential, G is the gravitational constant and ρ is the density. The cosmic fluid follows the energy-conservation equation,

$$\partial_t \rho + \nabla(\rho\vec{v}) = 0\,, \tag{2.33}$$

and the Euler equation that describes conservation of the momentum,

$$\partial_t \vec{v} + (\vec{v}\cdot\nabla)\vec{v} = -\frac{\nabla p}{\rho} - \nabla\Phi\,. \tag{2.34}$$

Each quantity is dependent on the comoving coordinates, $\vec{x} = \vec{r}/a$ and time t which were omitted from the above expressions. Recovering the density contrast $\delta(\vec{x},t)$ from before, we perturb the above equations through $\delta\Phi, \delta\vec{v}$, and δp with respect to their background values $\rho_0, \vec{v}_0, p_0$, and Φ_0. It reads

$$\Delta\delta\Phi = 4\pi G\rho_0 a^2 \delta\,, \tag{2.35}$$

$$\dot{\delta} + \nabla \cdot \delta\vec{v} = 0\,, \tag{2.36}$$

$$\delta\dot{\vec{v}} + H\delta\vec{v} = -\frac{\nabla\delta p}{a^2\rho_0} + \frac{\nabla\delta\Phi}{a^2}\,, \tag{2.37}$$

as the linear perturbation equations in comoving coordinates. This set of equations is closed with the equation of state for density perturbations,

$$\delta p = c_s^2\rho_0\delta\,, \tag{2.38}$$

where c_s is the speed of sound. If we combine the divergence of the Euler equation with the time derivative of the continuity equation and substitute in the Poisson equation, this gives a differential equation for the density contrast for pressureless dark matter, $c_s = 0$, in Fourier space,

$$\ddot{\delta}(k,t) + 2H(t)\dot{\delta}(k,t) + \left(\frac{c_s^2k^2}{a^2} - 4\pi G\rho_m\right)\delta(k,t) = 0\,. \tag{2.39}$$

Using the Friedmann Eq. (2.7) and the explicit Hubble function at Eq. (2.13), we can rewrite the above relation into

$$\delta''(a) + \left(\frac{3}{a} + \frac{H'(a)}{H(a)}\right)\delta'(a) - \frac{3}{2}\frac{\Omega_m(a)}{a^2}\delta(a) = 0\,. \tag{2.40}$$

There is a decaying $D_-(a)$ and a growing $D_+(a)$ solution of this equation. This leads to the definition the linear growth factor $D_+(a)$ as,

$$\delta(a) = D_+(a)\delta_0\,, \tag{2.41}$$

normalized to be today, $D_+(a_0 = 1) = 1$. For $c_s \neq 0$, structures grow above the Jeans length, defined as $\lambda_J \equiv c_s\sqrt{\pi/(G\rho)}$. Finally, the growth rate is defined as

$$f(a) = \frac{\mathrm{d}D_+(a)}{\mathrm{d}a}\,, \tag{2.42}$$

which is often approximated by $f(a) \approx \Omega_m^\gamma(a)$ with $\gamma = 0.55$.

Inflation predicts an initial power spectrum to be a power-law with a spectral index as an exponent n_s, $P_{\mathrm{initial}}(k) \propto k^{n_s}$. The spectral index may be a running spectral index and depend on k. Most inflationary models predict a constant, nearly scale-free initial power spectrum. When inflation ends, the comoving Hubble radius increases again. The initial perturbations that were once super-horizon size enter the horizon. From these, the modes entering the horizon before the matter-radiation equality oscillate as expected from Eq. (2.39). On the other hand, the modes entering after the matter-radiation equality continue to grow as $\delta \propto a^2$. In general all modes grow as $\delta \propto a$ after matter-radiation equality, which supresses of the small scale modes. The initial power spectrum has an asymptotic,

$$P(k) = \begin{cases} k^{n_s} & \text{if } k \leq k_{\mathrm{eq}}\,, \\ k^{n_s-4} & \text{if } k > k_{\mathrm{eq}}\,, \end{cases} \tag{2.43}$$

where k_{eq} is the comoving wavenumber of the horizon at matter-radiation equality. A transfer function, $T(k,a)$, does transition between the two branches in a Boltzmann code

such as `CAMB` [70] or `CLASS` [50]. A Boltzmann code then produces the evolved linear power spectrum

$$P_{\rm lin}(k,a) = D_+^2(a)T^2(k,a)P(k)\,. \tag{2.44}$$

The above description is only valid for a small density contrast, $\delta \ll 1$. As the Universe expands, the linear perturbation theory breaks down and one needs to go to higher order in perturbation theory. From a simple expansion, the result has terms that couple different Fourier modes of the density field. A review on standard pertubation theory can be found in [71].

2.2. Cosmological observables

The precison era of Cosmology stems from technological advances that enabled highly accurate measurements of many probes. This section discusses these probes and the datasets used throughout this dissertation. The details of each method for different probes are essential, given the data tensions briefly aforementioned, to shed light on the hidden assumptions and other key features. We will describe the following observables: the cosmic microwave background, the weak gravitational lensing, the galaxy clustering probes and the Hubble expansion rate $H(z)$.

2.2.1. Cosmic microwave background

As the Universe expands and cools down from a hot dense plasma state, the conditions change such that light elements can form. This process is called primordial nucleosynthesis and happened when temperature of the Universe was around 0.1 MeV. The production of nuclei remains in thermal equilibrium until the interaction rate, τ, becomes larger than the expansion rate of the Universe. At this time, the environment is very hot and dense where photons are constantly scattered by electrons through Compton scattering. Once these collisions are not energetic enough to create new photons, the decoupling happens (also known as particle freeze-out). This means that as the density of free electrons decreases, it also reduces the probability of electron-photon interactions. The opacity eventually decreases since no more photons are created.

Given the large amount of photons with respect to baryons, the recombination process is delayed as the environment is ionized. When the temperature reaches around 0.3 eV, shortly after matter-radiation equality, electrons and protons combine to form hydrogen and helium. This yields the last scattering of photons before they travel freely into the Universe, unchanged by the baryon temperature. The corresponding hypersurface of this moment is known as the last-scattering surface and it is approximately determined by the Saha equation. Thus, these free photons are the remnant of these processes, currently with a frequency in the microwave part of the radio spectrum. This radiation is coming from all directions in the sky, independent of any other formed structure such as galaxies or stars, hence the name cosmic microwave background.

Already in 1948, Gamow [72] estimated the temperature of the CMB today, starting from the formation of deuterium. It was only later in 1965 that Penzias and Wilson [73] measured an unexpected constant signal corresponding to 3.5 K in their radio antenna. The CMB is well described by a thermal black body radiation spectrum with a temperature of $T_0 = 2.7255 \pm 0.0006$ K [74]. The properties of the CMB have been thereafter observed by several missions: first by the Cosmic Background Explorer (COBE) sattelite [75, 76], followed by the Wilkinson Microwave Anisotropy Probe (WMAP) collaboration [77], and recently the Planck collaboration [78].

In Sec. 2.1.6 discuss a few useful definitions that we now use. The relation between the temperature fluctuations and its power spectrum is

$$\langle a_{\ell m} a_{\ell',m'} \rangle = \delta_{\ell\ell'} \delta_{mm'} C_\ell \tag{2.45}$$

for a homogenous and isotropic field. Given its angular dependence, it is convenient to

express the CMB temperature in spherical harmonics, through

$$T(\theta,\phi)=\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}a_{\ell,m}Y_{\ell,m}(\theta,\phi)\qquad\text{with}\qquad a_{\ell,m}=\int_0^{2\pi}\mathrm{d}\phi\int_0^{\pi}\mathrm{d}\theta\sin\theta\,Y^*_{\ell,m}(\theta,\phi)\,T(\theta,\phi)\,. \tag{2.46}$$

From this, the temperature-temperature power spectrum is estimated as

$$C_\ell^{TT}=\frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell}|a_{\ell,m}|^2=\langle|a_{\ell,m}|^2\rangle\,, \tag{2.47}$$

where the quantity usually plotted is $\mathcal{D}_\ell^{TT}=\frac{\ell(\ell+1)}{2\pi}C_\ell^{TT}\approx\frac{k^3P(k)}{2\pi^2}$, as depicted in Fig. 2.1.

The improved observational resolution allowed to detect small temperature anisotropies at the level of 10^{-5}. These come from mainly three phenomena: the Sachs-Wolfe effect, the Baryon Acoustic Oscillations, and the Silk damping. Following the assumption that large-scale structures stem from early density fluctuations present when the CMB was formed, these should be related to fluctuations in the Newtonian potential, through the Poison equation. The Sachs-Wolfe effect [79] is the change in energy of a photon that went through a fluctuation of the gravitational potential. This means that the photon gains energy if the gravitational potential fluctuation is positive and, thus, there is a correspondence of cold spots with over-dense regions. This effect is visible on the large scales, specifically, it corresponds to the flat region at low multipole, ℓ, limited by the dashed line in Fig. 2.1.

Secondly, there are density fluctuations of dark matter that gravitationally interact with the baryon-photon plasma. The gravitational attraction of an overdensity is then counterbalanced by the pressure from the plasma. The successive push and pull produces sound waves, a signal called Baryon Acoustic Oscillations [80]. Since the pressure is dominated by the photons, the equation of state is $p=\frac{1}{3}\rho$. From this one can quantify the sound speed to be $c_s\approx\sqrt{\frac{p}{\rho}}$. This sets a limit to the largest length traveled by these sound waves, which defines the sound horizon r_s. These oscillations are visible on scales below the sound horizon, namely, the three highest peaks in Fig. 2.1.

Finally, while the recombination takes place, the number of photon collisions decreases, and so does the mean-free path of the photons increases. The Thomson scattering relates the mean-free path λ_C with the cross-section σ_T for given an electron number density n_e via $\lambda_C\approx\frac{1}{n_e\sigma_T}$. The diffusion of the photons after N scatterings is the mean free path becomes $\lambda_D\approx\sqrt{N}\lambda$. This leads to a diffusion scale that depends on the baryon-photon ratio and the matter-radiation ratio. Such effect is known as the Silk damping [81] that is visible on small scales. It can be seen for instance in Fig. 2.1 by a decreasing amplitude at large ℓ.

If the electrons receive radiation with a spatial quadrupole, then through Thomson scattering, the CMB photons will be linearly polarized. This effect is much smaller, at the 10% level regarding the main CMB signal, but its detection [82] associated with the BAO peaks can provide constraints on models of inflation. The combination of the polarization-polarization power spectrum C_ℓ^{EE}, the temperature-polarization cross power spectrum C_ℓ^{TE}, and the lensing power spectrum $C_\ell^{\phi\phi}$ composes the final constraints of this probe.

Fig. 2.1 shows the temperature anisotropies measured by the Planck collaboration [83], which are extremely well explained by the six-parameter ΛCDM model. Using a flat

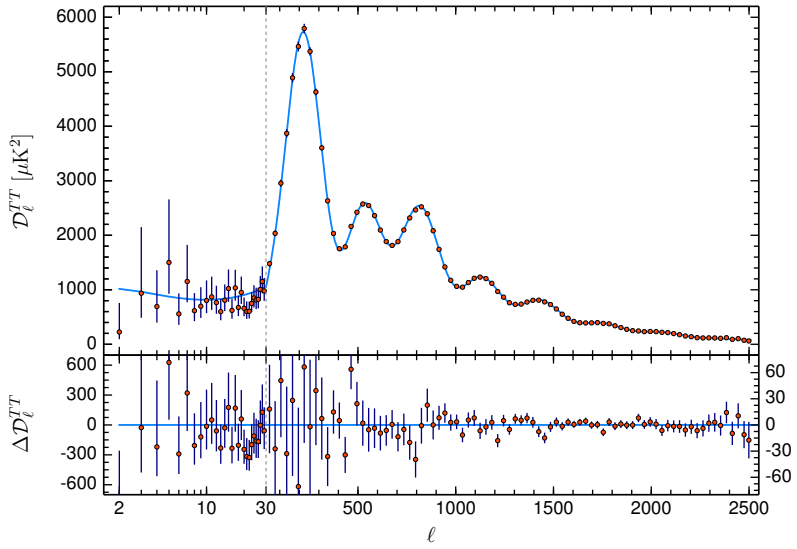

Figure 2.1.: Temperature anisotropies from the Planck collaboration [83]. The red dots with respective error bar correspond to the measurements and are plotted with the blue line representing the ΛCDM model at the best-fit cosmological paramaters found to fit this dataset. It is remarkable how well ΛCDM model can reproduce the CMB.

universe with a cold dark matter component and a cosmological constant, the blue solid line of Fig. 2.1 is a fitting curve with six parameters: the density of cold dark matter, $\Omega_c h^2$; the density of baryons, $\Omega_b h^2$, that incorporate hydrogen and helium as set by BBN; the angular scale of the acoustic oscillations, θ_*, which is written in terms of $100\theta_{MC}$; the optical depth of Thomson scattering from reionization, τ; the amplitude A_s, reported by $\ln(10^{10}A_s)$ and the spectral index n_s of a power-law spectrum of adiabatic perturbations. Furthermore, radiation is assumed to be made up of photons and low mass neutrinos. The remaining assumptions are the ones provided by the ΛCDM model as described in Sec. 2.1.3.

Although ΛCDM is an excellent fit, at several points in the CMB data reduction the ΛCDM model is assumed. Despite the model-dependency of the CMB, it is one of the cornerstones providing tight constraints on most of the cosmological parameters. Due to its accuracy, these became the reference values of cosmological parameters when testing the viability of any model. The results of the Planck collaboration [78] agree with simple single-field models of inflation. They find no evidence for primordial non-Gaussianity, isocurvature perturbations or cosmic defects.

2.2.2. Weak gravitational lensing

When looking to a galaxy very far away, it is likely that its image arrives at us distorted. Because light is deflected by matter inhomogeneities, this effect is called gravitational lensing. A comprehensive review on this subject is [84]. For a given source at distance D_s and a single lens at distance D_ℓ, the lens equation expresses the relation between the

observed position vector β and the true position of the source vector θ, given by

$$\beta = \theta - \alpha(\theta)\,, \tag{2.48}$$

where α is the scaled deflection angle vector. There are two different categories of this phenomenon: strong and weak gravitational lensing. The strong gravitational lensing generates large arcs and multiple images, as shown by the multiple solutions possible for Eq. (2.48). However, the weak lensing is a much smaller and frequent effect that can reveal properties of gravity as we shall explain.

Gravitational lensing can be represented by a distortion tensor with components of convergence κ and shear γ. The convergence κ is associated with the magnification effect and the shear γ expresses the stretching in the tangent direction. Specifically, one can rewrite the deflection angle α in terms of the lensing potential, Ψ, such as $\alpha = \nabla\Psi$. If the angular scale of the lens is much larger than the source, the lens mapping can be locally linearized as

$$\mathcal{A}(\theta) \equiv \frac{\partial\beta}{\partial\theta} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}\,, \tag{2.49}$$

where $\gamma \equiv \gamma_1 + i\gamma_2$ is the complex shear. The complex shear and the convergence are defined as

$$\gamma_1 = \frac{1}{2}(\partial_{11}\psi - \partial_{22}\psi)\,, \qquad \gamma_2 = \partial_{12}\psi\,, \qquad 2\kappa = \Delta\psi\,. \tag{2.50}$$

To measure the shear, the fundamental quantity is the geometric ellipticity ε of the galaxies. This can obtained by the reduced shear, $g \equiv \gamma/(1-\kappa)$, where in the weak lensing regime $\kappa \ll 1$, one recovers $\varepsilon = g \approx \gamma$. However, the galaxies measured are intrinsically elliptical, not circular, leading to the approximation of intrinsic ellipticity ε_s and lensed caused ellipticity g, $\varepsilon \approx g + \varepsilon_s$. It is fundamental that for large samples the intrinsic ellipticity averages out, $\langle\varepsilon_s\rangle \approx 0$. This is not a problem since the sky is full of distant faint galaxies and, typically, the angular scales are smaller than the variation of convergence and the shear of the lens. The measurements of ellipticities can be done by measuring the surface brightness using a forward modeling approach. Ellipticities can also be obtained in a model-independent way from quadrupole moments of the surface brightness.

To predict lensing effects produced by large-scale structures in a line-of-sight is not possible without the knowledge of the underlying matter distribution. To quantify this, the shear power spectrum $P_{\gamma\gamma}$ is obtained when measuring the galaxy ellipticities. Using linear theory, one can derive a relation of proportionality between the shear power spectrum $P_{\gamma\gamma}$ and the convergence power spectrum $P_{\kappa\kappa}$ [24, 85]. These spectra are also related to the matter power spectrum and cosmological parameters. Therefore, to measure of shapes of galaxies is another way to investigate the dynamics of the Universe.

Individual galaxy shapes do not provide enough information since the cosmic shear effect is small, plus galaxies are intrinsically elliptical. However, one can take the correlation function of these galaxies and its neighboring images in order to get a consistent distortion. The power spectrum of the convergence field, in the linear regime, can be written as

$$C_{ij}^{\kappa}(\ell) = \frac{9}{4}\int_0^\infty \mathrm{d}z \frac{W_i(z)W_j(z)H^3(z)\Omega_m^2(z)}{(1+z)^4} P\left(\frac{\ell}{r(z)}\right)\,, \tag{2.51}$$

where i,j are indexes refering to a tomographic approach with $n_{\rm bin}$ redshift bins. The conversion $k = \frac{\ell}{r(z)}$ is known as the Limber approximation [86], where $r(z)$ is the comoving distance. This approximation is valid for small angular separations and it simplifies the

calculations by avoiding Bessel functions to be integrated. It relies on the slowly-varying property of the applied function. Eq. (2.51) uses the window function $W(z)$ which is given by

$$W(z) = \int_z^\infty \mathrm{d}z' \left(1 - \frac{r(z)}{r(z')}\right) n(z')\,, \tag{2.52}$$

with $n(z)$ being the normalized source distribution function that is specific for each survey. The convergence power spectrum has additionally a shape noise contribution, $\sigma^2_{SN} = \frac{n_{\mathrm{bin}}\sigma^2_\epsilon}{\bar{n}}\delta_{ij}$, where σ^2_ϵ corresponds to the intrinsic ellipticity dispersion and $\bar{n}$ the mean number of sources in a steradian. There is the additional assumption that the same number of galaxies is contained in each bin.

Weak gravitational lensing can help us investigate the matter distribution. This can later be addressed in terms of the gravitational field, independently of the nature of the matter or its physical state. The geometric derivation of weak lensing also enables to understand the geometry of the Universe. These factors turn this probe into a vital contribution in cosmology, being connected to many studies about dark matter.

In addition, the CMB photons also suffer a small lensing effect. Previously we discussed how the CMB photons travel since the recombination epoch to us today and through that time also pass by matter inhomogeneities in large-scale structures [87–89]. The scattering of the deflection angle correlates with the typical size of the potential well, leading to a correlation length close to one of the acoustic peaks. This means that the BAO signal might be reduced due to the overall lensing of the CMB photons. For further details, a good review of this topic is found in [90].

2.2.3. Galaxy clustering

Another way to probe the Universe is by counting galaxies through the density contrast δ in terms of the density of galaxies δ_g. Galaxy clusters are the largest gravitationally bounded objects and their properties entail a lot of the cosmic history. Although, the estimators used in real data can be much more complicate the idea of galaxy clustering is simple: the galaxy number-density δ_g is obtained through the number count compared to the average number of objects in that cluster, which explicitly means

$$\delta_g \equiv \frac{n_g(k,z) - \langle n_g\rangle}{\langle n_g\rangle}\,. \tag{2.53}$$

The formation of galaxies is a complex and not well understood process that relies on the environment of the galaxy number-density δ_g, that generally relates to the underlying matter distribution via

$$\delta_g(k,z) = b(k,z)\delta(k,z)\,. \tag{2.54}$$

where $b(k,z)$ is the galaxy bias function that includes the uncertainties in our knowldge on galaxy formation. This relation is usually simplified to be scale independent in the linear regime at large scales. At small scales, more complicated models are required to match the matter power spectrum to the observed power spectrum. The endpoint of this probe is to obtain the matter power spectrum and decompose it in spherical harmonics in radial shells, in a similar approach as previously discussed for weak lensing [91–95]. The projected tomographic power spectrum is

$$C^g_{ij}(\ell) = \int_0^{\chi_H} \frac{\mathrm{d}\chi}{\chi^2} W^g_i(\chi) W^g_j(\chi) P\left(\frac{\ell}{\chi}, \chi\right)\,, \tag{2.55}$$

where a window function is used per tomographic bin W_i^g, specific for each survey specificities.

Galaxy clustering is, in principle, a straightforward probe for cosmology from the above description. In reality, large scale surveys need to be very careful in considering a few effects present in the matter power spectrum. From the theory, the power spectrum $P(k)$ depends on the cosmological parameters of, for instance, the ΛCDM model. This is later modified by several other effects as we shall describe next. The matter power spectrum is by definition the Fourier transform of the two-point correlation function of Eq. (2.54). In a forecast analysis, a model with a fixed set of parameters is the reference frame called the fiducial model. For such analysis, the full expression for the observed power spectrum is [96, 97]

$$P_{\rm obs}(k,\mu,z) = \frac{D^2_{A,\rm ref}(z)H(z)}{D^2_A(z)H_{\rm ref}(z)} b^2(z)(1+\beta(z)\mu^2)^2 \exp\big[-k^2\mu^2(\sigma_z^2/H(z)+\sigma_v^2(z))\big]\,. \quad (2.56)$$

In this expression, the fraction that includes the angular distance D_A^2 and the Hubble function $H(z)$ is linked to the geometrical change caused by the BAO. Different cosmological models produce distinct geometric changes in the BAO peak proportional to this fraction [98]. $b(z)$ is the galaxy bias function previously defined in Eq. (2.54). Typically, $b(z)$ is assumed to be a linear function of redshift but that assumption breaks down at very small and very large scales [99]. The redshift space distortions is associated with the function $\beta(z) = f(z)/b(z)$ with $b(z)$ being the galaxy bias and $f(z)$ the logarithmic growth rate of perturbations $f(z) = \mathrm{d}\ln D_+(z)/\mathrm{d}\ln a$. This is known as the Kaiser formula [100] and it expresses how distortions are created by the peculiar velocity divergences in redshift space. Furthermore, the Alcock-Paczyński test [101] is also included in Eq. (2.56), through the dependence on the angle μ and the scale k. If the cosmological parameters change, then μ and k are adjusted by a geometrical distance factor. This geometric test aims to detect the presence of a cosmological constant, which can be distinguished from the RSD if the dataset is accurate and large enough [102]. Finally, the observed power spectrum of Eq. (2.56) is damped by the spectroscopic errors σ_z and the non-linear peculiar velocity dispersion $\sigma_v(z)$. This dispersion $\sigma_v(z)$ is known as the Fingers of God effect and it corresponds to the first-order correction to the aforementioned Kaiser formula. At small scales, the Fingers of God effect suppress the power spectrum leading to a much higher uncertainty when attempting to reproduce the underlying theoretical function.

Baryon Acoustic Oscillations

As explained previously, before recombination, baryons and photons were tightly coupled. As the temperature drops and opacity decreases, these baryon perturbations are released as pressure waves. These leave an imprint at a specific scale on late-time matter clustering at the radius of the sound horizon,

$$r_s = \int_{z_d}^{\infty} \frac{c_s(z)}{H(z)} \mathrm{d}z\,, \quad (2.57)$$

where z_d is the redshift at drag epoch, right before recombination when photons and baryons decouple. This signature of the Baryon Acoustic Oscillations (BAO) appears as a precise peak in the galaxy correlation function [80]. These oscillations show up as wiggles in the matter power spectrum, at roughly $100h^{-1}$Mpc, making it a "standard ruler". This

feature can be measured through the line-of-sight and transverse directions constrain $H(z)$ and $d_A(z)$ separately. The sound horizon r_s relies on the speed of sound of the baryon-photon plasma and on time scales of the early universe and is thus not directly measured. Instead, the following ratios are estimated

$$\Delta\theta = \frac{r_s}{(1+z)d_A(z)}\,, \qquad \Delta z = \frac{r_s H(z)}{c}\,. \tag{2.58}$$

These ratios amount for an increase in the number of galaxy pairs for a specific angular separation $\Delta\theta$ in the line-of-sight and also an enhanced number of galaxy pairs with a redshift separation Δz in the radial direction. This method works as an inverse distance ladder approach where the distance calibration is done through the sound horizon r_s that can be measured by the CMB.

Redshift-Space Distortions

Another way to obtain these quantities is by the Alcock-Paczyński test, applied to the two-point statistics if redshift-space distortions (RSD) are well modeled. These appear as constraints on the product of two quantities, $f(z)\sigma_8(z)$. The first is the growth function $f(z)$ and the second is the normalization of the linear matter power spectrum for spheres of $8h^{-1}$Mpc, $\sigma_8(z)$. These quantities are defined as

$$f(z) \equiv \frac{\mathrm{d}\ln G(a)}{\mathrm{d}\ln a} \qquad \sigma_8(z) = 4\pi \int \frac{k^2 \mathrm{d}k}{(2\pi)^3} P(k,z)\,, \tag{2.59}$$

with the scale factor a defined as $a = 1/(1+z)$ and the $P(k,z)$ is the matter power spectrum as in Eq. (2.56). This is independent of any "standard ruler" and it can determine the ratio $\Delta z/\Delta\theta \approx (1+z)d_A(z)H(z)/c$. The combination of the peaks of BAO and the Alcock-Paczyński test is the approach used by the WiggleZ Dark Energy Survey [3]. They use measurements of the averaged galaxy correlation function and the two-dimensional power spectrum in tangential and radial Fourier bins. Similarly, [2] uses data from the Baryon Oscillation Spectroscopic Survey (BOSS) and Sloan Digital Sky Survey (SDSS) to obtain measurements of $H(z)$. The data from these last two references will be used later on and they correspond to the method 2 label on Tab. 3.2. Furthermore, the same BOSS survey also includes measures from the Lyman α forest [103, 104], also used in our work and labeled as method 3 in Tab. 3.2.

2.2.4. Hubble expansion rate $H(z)$ and Hubble parameter H_0

The discovery of the expansion of the Universe was coined by Edwin Hubble, in 1929, who found a relation between the distance and the radial velocity of nebulae [105]. This relation is known as the Hubble law,

$$v_r = H_0 d\,, \tag{2.60}$$

where the constant of proportionality H_0 corresponds to the current expansion rate, valid for small distances. The expansion of the Universe is usually directly linked to its different matter components. Thus, at any point in cosmic history, the expansion of the universe is given by

$$H(z) = H_0\sqrt{\Omega_\Lambda + \Omega_{k0}(1+z)^2 + \Omega_{m0}(1+z)^3 + \Omega_{r0}(1+z)^4}\,, \tag{2.61}$$

for a general non-flat ΛCDM model, where Ω_{m0} is the matter density parameter today, Ω_{r0} is the radiation density parameter today, Ω_Λ is the dark energy density parameter, and Ω_{k0} the curvature density parameter today. The curvature is taken as flat in ΛCDM, and generally, radiation is neglected. Since $\sum_i \Omega_{i0} = 1$, these considerations reduce the above equation to be dependent only on H_0 and Ω_{m0}, and thus, $H(z) = H_0\sqrt{\Omega_{m0}(1+z)^3 + 1 - \Omega_{m0}}$. Measuring the expansion rate of the Universe constitutes one of the crucial observational tests in Cosmology. Currently, there are several techniques to measure this and the main ones are outlined next.

Supernovae type Ia and Cepheid variable stars

The late-time accelerated expansion of the Universe was found independently by two teams, the High-Z Supernova Search Team [13] and the Supernova Cosmology Project [12]. This acceleration was measured through the light curves of supernova Type Ia (SNIa). This subtype of supernova is identified by having an absorption line of singly ionized silicon and the lack of a hydrogen line. Typically it stems from a binary system where one of the stars is a white dwarf. If this white dwarf is rotating slowing and accreting matter from its companion, the stability of this accretion is set by the Chandrasekhar limit that is 1.44 solar masses. Above this limit, the white dwarf no longer balances its own gravity with the electron degeneracy pressure, leading to an extremely bright supernova explosion. The absolute luminosity of these supernovae is roughly constant at the peak of highest apparent luminosity, making them known as "standard candles". What is measured is the apparent flux $\mathcal{F}$, the amount of light from a specific stellar object. This can be expressed as function of its luminosity L and its distance d_L, such as $\mathcal{F} = L/4\pi d_L^2$. We can later relate the apparent magnitude m, which is the measured brightness, with M the absolute magnitude, defined by being at a fixed distance of 10 pc, through

$$\mu \equiv m - M = 5\log_{10}\left(\frac{d_L}{10\text{pc}}\right), \tag{2.62}$$

where μ is the relative magnitude and color corrections might need to be added here for an accurate relation. This luminosity distance d_L is then associated to the content of a flat Universe through

$$d_L(z) = c(1+z)\int_0^z \frac{\mathrm{d}z'}{H(z')} \underset{\text{for } z\ll 1}{\approx} \frac{cz}{H_0}, \tag{2.63}$$

with the Hubble function $H(z)$ is given by Eq. (2.61). Thus, assuming a specific content of the Universe, one can infer H_0 for a very close SNIa or, in general, $H(z)$, since the redshift z is obtained by measuring the wavelength λ of the observed light through $z \equiv \frac{\lambda_0}{\lambda} - 1$. Clearly, from Eq. (2.62) it is required to know exactly the absolute magnitude M which cannot be disentangled from the apparent magnitude m within the same dataset. For this reason, the light curves need to be calibrated and there are several methods to do so [106–110]. In simple terms, one can add an extra parameter at the data analysis stage, to take the degeneracy of M with H_0 as well as other possible effects into account, but this naturally decreases the amount of information available leading to a higher uncertainty on this distance measurement.

Another way of calibrating the distance to the SNIa is to use the cosmic distance ladder. In [1, 111], the process is to use Cepheid variable stars in which brightness varies with a

well-defined period. This calibration process starts by measuring the distance to Cepheids that belong to the Milky Way through their parallax. After the brightness of the Cepheids is calibrated through comparison to other stars, these can be used as a reference to larger distances, and reaching thus the distances at which supernova exists in other galaxies. This allows for the calibration of the supernovae brightness. And, once going further to more distant galaxies that host both Cepheid variable stars and SNIa, the expansion of the Universe can then be measured. Comparing the distance of SNIa with its redshift, which is converted into its recession velocity, this cosmic ladder allows us to obtain the Hubble parameter today H_0 in a cosmological model-independent way. This calibration method used in [1, 111] is further supported by using Cepheids and detached eclipsing binaries at the Large Magellanic Cloud [112], and the distance to the galaxy NGC 4258 using its water masers at the nucleus [113]. These improvements led to the recent 1% accuracy of the H_0 determination [111], namely the last estimate, $H_0 = 74.03 \pm 1.42\,\mathrm{km\ s^{-1}Mpc^{-1}}$. A further discussion on the present discrepancies between the different methods can be found at Sec. 2.2.5.

Cosmic Chronometers

The cosmic chronometers method is a differential dating technique [114], which relies directly on the definition of the Hubble parameter writen as

$$H(z) = \frac{\dot{a}}{a} = -\frac{1}{(1+z)}\frac{\mathrm{d}z}{\mathrm{d}t}\,, \tag{2.64}$$

where a is the scale factor and the dot refers to the derivative with respect to z. This spectroscopic technique relies on the age difference between two passively-evolving galaxies, Δt, that were formed at the same time but have a different redshift, Δz. This ratio of $\Delta z/\Delta t$ allows to infer the above derivative $\mathrm{d}z/\mathrm{d}t$. Passively-evolving galaxies are galaxies that have a low star formation rate and an old stellar population. A careful selection of these galaxies provides a rather homogenous sample of galaxies with similar metallicity and low star formation rates. This measurement is independent of the metric or any cosmological model. Furthermore, to use this method for further distances, the passive evolution of these galaxies must be much slower than the age difference. Typically, this means that the best candidates are massive early-type galaxies which have most of their stellar mass formed very rapidly. Later on, an easy way to identify them was found, namely, through the spectral break at 4000 Å of wavelength λ in the galaxy light spectrum. This discontinuity arises from metal absorption lines that have their amplitude linearly related to the age and metal abundance of the stellar population. The valuable advantage of this technique is its differential nature that clears out systematical uncertainties or other effects from galaxy evolution. The majority of the measurements of $H(z)$ that we use are done through this method [115–120], presented with label 2 in Tab. 3.2. From these measurements, one can infer H_0 choosing a specific interpolation scheme and then read out the value at $z = 0$. This inference can be somehow biased if a value of H_0 is chosen and it is also dependent on the interpolation scheme and its underlying assumptions [121]. Interestingly, the cosmic chronometers dataset can support both a higher and a smaller value of H_0 given the choice of a star formation model as it is one of the key features of this technique [122].

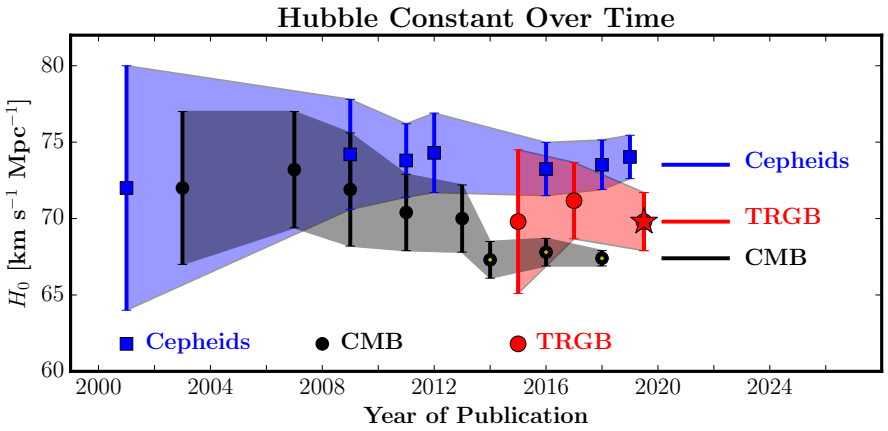

Figure 2.2.: Cepheids (blue), Tip of the Red Giant Branch (red) and Cosmic Microwave Background (black) are three examples of methods used to obtain H_0. Their uncertainty has decreased with time (year of publication) but a discrepancy between Cepheids and CMB measurements became clear. This figure is taken from [125].

Tip of the Red Giant branch

When the red giant stars stop the fusion of hydrogen and start fusing helium at their core, the star reaches a specific peak of brightness, usually known as the Helium flash. This produces a particular feature in the luminosity function of these very bright and red stars and they can be seen in galaxies that host SNIa within distances up to ~ 20 Mpc by the Hubble Space Telescope (HST). Among other benefits from this method such as little reddening effects or its isolated locations, these older stars are present in any type of galaxy, an advantage regarding Cepheids that are only present in late-type star-forming galaxies. Therefore, they can be used as an independent method taking the role of the Cepheids in the SNIa calibration. The estimated values of H_0 also reach a few percent accuracy, specifically, $H_0 = 69.6 \pm 1.9\,\mathrm{km\ s^{-1}Mpc^{-1}}$ [123] and $H_0 = 72.4 \pm 2.0\,\mathrm{km\ s^{-1}Mpc^{-1}}$ [124]. This offset comes from different estimations on the color corrections due to the dust of the Large Magellanic Cloud and calibration between ground-based photometry and the HST. In Sec. 2.2.5, we discuss the ongoing discrepancies between the different methods.

2.2.5. Tensions between different datasets

In the past decade, the amount of independent methods to measure H_0 has increased and, most importantly, the accuracy of these measurements has remarkably improved [126]. Better control over systematical uncertainties led to an error reduction from 10% to nearly 1% on the estimation of H_0, as seen in Fig. 2.2. Also, the fact that different methods were giving distinct values is not new, but the prevailing discrepancy is at the moment about 5σ of statistical significance. The H_0 tension is usually atributed to the disparity between the value measured by the CMB probe with the Planck mission [83], namely $H_0 = 67.4 \pm 0.5\,\mathrm{km\ s^{-1}Mpc^{-1}}$ and the local measurement from the Cepheids distances by the HST SH0ES [111], $H_0 = 74.03 \pm 1.42\,\mathrm{km\ s^{-1}Mpc^{-1}}$. However, recently a clear distinction between late-time and early-time probes appeared, as depicted in Fig. 2.3. This means that local measurements from Cepheid variable stars and strong lensing are on one side with

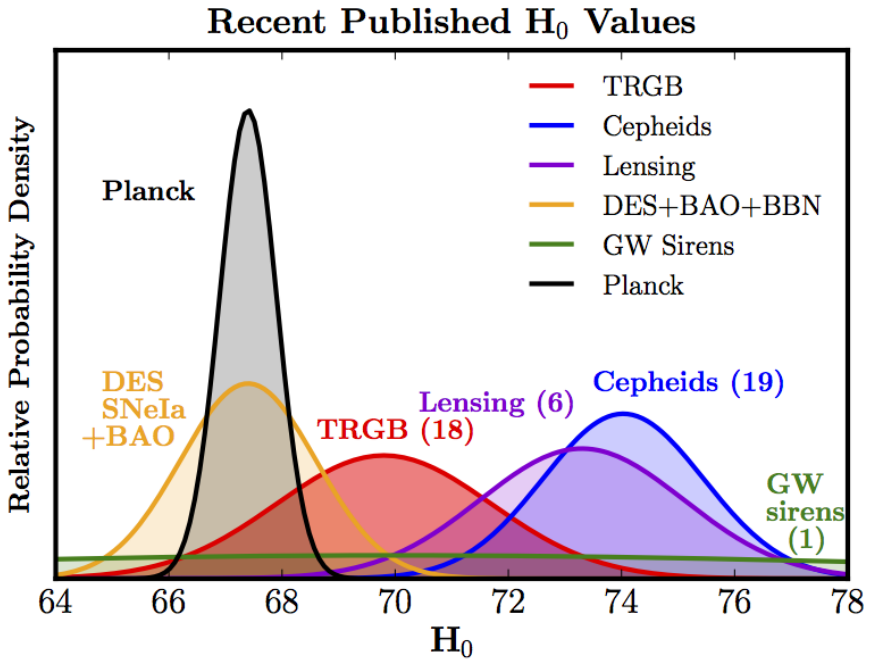

Figure 2.3.: Several recent different probes are not agreeing on the value of H_0. The DES+BAO+BBN and the Planck bounds correspond to probes of the early Universe. The remaining probes give information on the late Universe. Since there is evidence for both small and large values of H_0 from probes other than CMB and Cepheids, this might hint a possible crisis in cosmology. This figure is taken from [123].

larger values of H_0 and smaller values are given by CMB (both Planck and the previous WMAP mission), BAO and Big Bang Nucleosynthesis (BBN) on the other side. The TRGB value [123] is here in the middle of the tension with $H_0 = 69.6 \pm 1.9\,\mathrm{km\ s^{-1}Mpc^{-1}}$. However, this last value was updated from [125], namely $H_0 = 69.8 \pm 1.9\,\mathrm{km\ s^{-1}Mpc^{-1}}$, as [124] claimed that the extinction of the Large Magellanic Cloud was over estimated, finding then $H_0 = 72.4 \pm 2.0\,\mathrm{km\ s^{-1}Mpc^{-1}}$. This is one of the many examples of how the estimation of H_0 and its error have often been updated and scrutinized in the past three years, but unfortunately, its origin is still unknown. The two main explanations focus on unaccounted systematical errors or new physical models.

Ref. [123] states that the two main questions that need to be solved are the weak lensing effects on SNIa measurements and the local bubble possibility. The local bubble [127, 128] relies on the idea that our neighborhood is an underdense region which would then change the local uncertainty values. Through N-body simulations, they find that such an effect is too small to explain the tension, or any other local structures [129]. Regarding the weak lensing affecting SNIa, the effect is again almost negligible [130]. Assuming that the problem is not in the local measurements, the question remains about what are the alternatives if this tension persists. The six parameter ΛCDM model has not a very firm foundation on the nature of dark energy or dark matter. The possibilities of a new neutrino species [131], or even, an evolving dark energy equation of state, among other possibilities (see [132, 133] and references therein) are not yet enough to justify the present discrepancy and might even lead to further discordance with the other probes: BAO, WL, and BBN.

At low redshift, the predicted value for $f(z)\sigma_8$ from the ΛCDM model differs from the ones measured by a few weak lensing surveys, as shown later in Fig. 3.1. Usually this tension is easily identified on the (Ω_m, σ_8) parameter plane and for that reason, the values quoted are the quantity $S_8 \equiv \sigma_8\sqrt{\Omega_m/0.3}$. This discrepancy started a few years ago, between the CFTLenS [134] and Planck 2013 data release [135] of around $2-3\sigma$ level of statistical significance. Throughout the last years, the accuracy of both weak lensing and CMB surveys improved, including the new large weak lensing surveys such as the Kilo-Degree Survey (KiDS) [136–138] and the Dark Energy Survey (DES) [139]. Initially, these two surveys had also a difference between them which was solved by the addition of further data. Each survey shows about 2σ discrepancy with the Planck 2015 data release [140]. In the latest combination of both of these surveys, KiDS and DES [141], the tension with Planck 2018 [83] is at the 2.5σ level. This corresponds to $S_8 = 0.762^{+0.025}_{-0.024}$ for the weak lensing surveys against the $S_8 = 0.832 \pm 0.013$ from the CMB probe (TT,TE,EE+lowE+lensing). The origin of this discrepancy also remains unclear.

3. Model independent reconstruction of the anisotropic stress parameter η

In August 2017, the LIGO/VIRGO collaboration observed gravitational waves from a neutron star merger (GW170817) with its electromagnetic counterpart (GRB170817) detected by many other telescopes [51–55]. One of the most important consequences from this particular event is the constraint [51] on the difference between the speed of gravity c_T and the speed of light c to be

$$-3 \times 10^{-15} \leq \frac{c_T - c}{c} \leq +7 \times 10^{-16} \,. \tag{3.1}$$

From the several consequences of such measurement [56, 57], this constraint has a direct impact [58–61] on the four free frunctions in the Horndeski Lagrangian [65, 142, 143].

This restriction on the possible alternative theories to the ΛCDM model poses the question of how can one further constrain these theories. As introduced in Sec. 2.1.5, the anisotropic stress parameter η Eq. (2.20) describes deviations to GR. In the general Horndeski theory, and also in bimetric gravity [144], the anisotropic stress parameter corresponds to

$$\eta = h_2 \frac{1 + h_4 k^2}{1 + h_5 k^2} \,, \tag{3.2}$$

in Fourier space where h_i are time functions that depend on the four free functions in the Horndeski Lagrangian [65, 142, 143]. The grativational-wave event constraints h_2 equal to 1 since $h_2 = 1/c_T^2$ [65], but leaves all the other functions free. It is worth to note that also one obtains $\eta = 1$ in the limit of large sub-horizon scales and provided that the theory does contain at least one mass scale beside the Planck mass [67], . The general idea is that $\eta \neq 1$ indicates a deviation from standard gravity by a form of dark energy that does not behave as a perfect fluid.

Considering model-independent observables and linear structure formation, and assuming gravity remains universally coupled also when modified, one can build an estimate of η formed by three directly observable functions of redshift (see [145]). Given the current available data, it is interesting to revisit this estimator and analyze a possible deviation from the perfect fluid approach as a way to probe gravity.

These observables correspond to specific measured quantities like the dimensionless Hubble function $E(z)$, the growth rate of matter density perturbations times the normalization of the power spectrum $f\sigma_8(z)$, and the E_G statistics which relates the lensing potential and the growth rate of structure formation. The ratios of these observables are the key concepts for this work and they are described later in Sec. 3.1. Nonetheless, it is worth to mention that they incapsulate information from different cosmological probes such as galaxy clustering and weak lensing, as well as measurements of the Hubble function. We reconstruct these parameters as a function of redshift by using the most recent data available from several collaborations, presented in Sec. 3.2.

To reconstruct an unknown function and its derivative from sparse and noisy data is not trivial. Therefore, we use three different approaches to estimate these functions from the data. One of the methods is the binning scheme which is commonly used to group data in intervals. We also used the Gaussian Process method, a generalization of a Gaussian distribution, where instead of random variables, one has a distribution of random functions, connected by a specific correlation function. And lastly, we use a polynomial regression in which one assumes a linear model for the underlying function. Each of these methods has their advantages and caveats that are described in Sec. 3.3.

3.1. Model-independent observables

Previously, in Sec. 2.1.3, we presented the definition of the anisotropic stress paramter η, Eq. (2.20), as the ratio of the gravitational potentials. Also, we mentioned how matter background density and matter overdensity cannot be determined in a model-independent way. However, one can define model-independent observable quantities that do not depend on the assumptions of the initial conditions of the Universe, neither on the primordial power spectrum nor on nature of dark matter or the details of galaxy bias.

Following [145], these quantities are called A, R, L, and E, respectively denoting amplitude, redshift-space distortions, lensing, and the dimensionless Hubble function. They are defined as

$$\begin{aligned} A &= b\delta_m\,, & R &= f\delta_m, \\ L &= \Omega_{m0}\Sigma\delta_m\,, & E &= H/H_0\,. \end{aligned} \tag{3.3}$$

where b is the galaxy-matter linear bias, $f = \delta_m'/\delta_m$ is the growth rate where the prime is a derivative with respect to $\ln a$, and Ω_{m0} is today's matter fractional density. Both f and b can be functions of redshift and scale. In the following, we assume that scale dependence can be ignored as it is not provided by the available data. Thus, also η is assumed to be independent of scale in the observed range. From Eq. (3.2), it is possible to see that scale-independence sets in either at small scales $k \gg 1$ or at large scales $k \ll 1$ (but in this case $\eta \to 1$). The scale independence is present at all scales if $h_4 = h_5$ or if the theory does not contain a mass scale.

With these definitions from Eq. (3.3), it was shown in [65, 145] that one can obtain model-independent quantities where the effects of the shape of the primordial power spectrum and the galaxy bias are canceled out, namely

$$P_1 \equiv \frac{R}{A} = \frac{f}{b}, \tag{3.4}$$

$$P_2 \equiv \frac{L}{R} = \frac{\Omega_{\rm m0}\Sigma}{f}, \tag{3.5}$$

$$P_3 \equiv \frac{R'}{R} = f + \frac{1}{f}\frac{{\rm d}f}{{\rm d}a} = \frac{(f\sigma_8(z))'}{f\sigma_8(z)}\,, \tag{3.6}$$

with the prime We have defined $f\sigma_8(z)$ as

$$f\sigma_8(z) = \sigma_8 G(z) f(z)\,, \tag{3.7}$$

where σ_8 is the amplitude of the linear power spectrum defined in a spherical shell of radius 8 Mpc at redshift $z = 0$ and $G(z)$ is the growth function normalized to unity today, $\delta(z) = \delta_{m,0}G(z)$.

The continuity equation and the Euler equation together, relate the divergence θ of the peculiar velocities of galaxies propagating on geodesics to the gravitational potential Ψ

$$(a^2\theta)' = a^2k^2H\Psi\,, \tag{3.8}$$

in the linear structure formation regime. This allows us to write down the lensing and Poisson equations in Fourier space, respectively, with the definitions from Eq. (3.3), in the following way

$$-k^2(\Psi-\Phi) = \frac{3(1+z)^3L}{2E^2}\,, \tag{3.9}$$

$$-k^2\Psi = R' + R\left(2+\frac{E'}{E}\right). \tag{3.10}$$

The last equation is usually known as the equation of linear growth of matter perturbations. Dividing the lensing equation Eq. (3.9) by the equation for the growth of structure Eq. (3.10), we obtain the ratio of the gravitational potentials and the gravitational slip as a function of model-independent observables

$$\eta_{\rm obs} \equiv \frac{3P_2(1+z)^3}{2E^2\left(P_3+2+\frac{E'}{E}\right)} - 1\,. \tag{3.11}$$

We label this expression as $\eta_{\rm obs}$ to separate the theoretical prediction with this observationally motivated expression although $\eta_{\rm obs} = \eta$.

We reconstruct Eq. (3.11) using present data in a model-independent way. This is an interesting exercise since $\eta_{\rm obs}$ is constructed in a way that it is free from an initial power spectrum, galaxy bias, the density of matter or any assumptions about cosmic expansion. Not requiring a ΛCDM background or any other is an important asset in order to find deviations from the theoretical expectations.

The P_2 parameter can be related to the E_G statistics, defined in the cosmological literature (see [146] and references therein) as the expectation value of the ratio of lensing and galaxy clustering observables at a scale k

$$E_G = \left\langle \frac{a\nabla^2(\Psi-\Phi)}{3H_0^2f\delta} \right\rangle_k . \tag{3.12}$$

This relation simplifies to be $E_G = \Omega_m\Sigma/f$ for a flat ΛCDM cosmology. Since $\Sigma = 2$ for standard gravity, taking the Poisson equation as Eq. (2.22) and the definitions of the A, R, L, E variables as Eq. (3.3), the relation with P_2 is given by

$$P_2 = 2E_G\,. \tag{3.13}$$

The E_G statistics has been used several times as a test of modified gravity [146–148], but it is not a model-independent test *per se*. We will further explain the importance of this statement in Chapter 4 where we go a bit deeper into the definition of the E_G statistics.

3.2. Data

We reconstruct $\eta_{\rm obs}$ using $E(z)$, $P_2(z)$, and $P_3(z)$ that correspond to the data listed in Tab. 3.2, Tab. 3.4, Tab. 3.5, Tab. 3.3, Tab. 3.6, and Tab. 3.7. These datasets are also

Table 3.1.: Fiducial parameter values for our reference ΛCDM case, using Planck 2018 data from TT+TE+EE+lowE+lensing [83], except for H_0, where we use the local value from the HST collaboration [149] as explained in the main text.

Ω_{m0}	Ω_{DE}	Ω_b	n_s	σ_8	H_0[km/s/Mpc]
0.3153	0.6847	0.0493	0.9649	0.8111	73.45

shown in Fig. 3.1. We use the Hubble parameter $H(z)$ data to obtain $E(z)$ and $E'(z)$. For $P_2(z)$ we rescale the $E_G(z)$ data, while $P_3(z)$ is reconstructed from $f\sigma_8(z)$ and its derivative with respect to $\ln a$. Tab. 3.1 shows the cosmological parameters from the TT+TE+EE+lowE+lensing Planck 2018 best-fits [83], that we use to plot the ΛCDM curves for the different cosmological functions in Fig. 3.1. The details of the sources of the data are explained below.

In the present work, we only use the H_0 value to normalize the $H(z)$ measurements into the dimensionless quantity $E(z)$. Notice that H_0 is an observable quantity that can be estimated from local kinematics in a way that it is independent of cosmology and modified gravity. Therefore, for the normalization of the $E(z)$ measurements we need to choose a value of H_0, for instance from the latest results of the Planck collaboration [83] or the value obtained by the HST collaboration [149].

We choose to use the local value of H_0 determined by the HST collaboration in our calculations, which amounts to $H_0^{\rm HST} = 73.45 \pm 1.66$ [km/s/Mpc], because it is cosmology-independent as discussed previously. In Sec. 3.4 we compare and discuss the results using the Planck value. Thus, by construction, we have an extra data point at $z = 0$, namely, $E(z = 0) = 1$. The uncertainty on H_0 propagates to all $E(z)$ values, and we take this into account as detailed in the next section.

3.2.1. Hubble parameter data

Regarding the Hubble parameter measurements, we used the compilation of $H(z)$ data from [121] (see Tab. 3.2) which includes measurements from [115, 116, 118, 119], Baryon Oscillation Spectroscopic Survey (BOSS) [103, 104, 117] and the Sloan Digital Sky Survey (SDSS) [2, 120]. The several methods are described in more detail in Sec. 2.2.4.

In this compilation, the majority of the measurements were obtained using the cosmic chronometric technique, labeled as method 1 in Tab. 3.2. This method infers the expansion rate $\mathrm{d}z/\mathrm{d}t$ by taking the difference in redshift of a pair of passively-evolving galaxies. The remaining measurements were obtained through the position of the Baryon Acoustic Oscillation (BAO) peaks in the power spectrum of a galaxy distribution for a given redshift. This is labeled as method 2 in Tab. 3.2.

Additionally, we use the recent results from [1] where a compilation of Supernovae Type Ia from CANDELS and the CLASH *Multi-cycle treasury program* was analyzed, providing six measurements of the expansion rate $E(z)$, with considerably smaller error bars, compared to the other above mentioned techniques. These are listed in Tab. 3.3. In the original reference [1], the errors are not symmetric, therefore we recalculated symmetric errors, as the quadrature of the 1σ bounds on the left and right side of the central value.

The measurements from [103] and [104] are obtained using the BAO signal in the Lyman-α forest distribution alone or cross-correlated with Quasi-Stellar Objects (QSO) (for the details of the observational methods, we refer the reader to the original references). Ref-

erence [2] provides the covariance matrix of its three $H(z)$ measurements obtained from the radial BAO galaxy distribution. We display the covariance matrix in Tab. 3.4. To this compilation, we have added the results from the WiggleZ Dark Energy Survey [3] whose covariance matrix can be found in Tab. 3.5.

The cosmic chronometric technique provides measurements of H_0 that are independent of large-scale cosmology. Recent work [150] has shown that these data prefer a lower value for the H_0 value taking different data reconstruction methods. However, an upper value can also be found if a different model of stellar population synthesis is chosen when using the data from [117]. For our fiducial results, we fix our choice of the Hubble parameter to the HST measurement.

As previously mentioned, the data points of the Hubble parameter $H(z)$ have to be converted into the dimensionless expansion rate $E(z)$. This means that for each measurement $H_i = H(z_i)$, we compute

$$E_i = \frac{H_i}{H_0}\,, \tag{3.14}$$

so that the corresponding uncertainty is converted as well through

$$\delta E_i = \frac{\delta H_i}{H_0} - H_i \frac{\delta H_0}{H_0^2}\ . \tag{3.15}$$

The covariance of this random matrix is the expected value of the product of δE_i and δE_i, which is

$$\begin{aligned}\langle \delta E_i \delta E_j \rangle &= \left\langle \frac{\delta H_i}{H_0}\frac{\delta H_j}{H_0} \right\rangle + H_i H_j \left\langle \frac{\delta H_0}{H_0^2}\frac{\delta H_0}{H_0^2} \right\rangle \\ &= \frac{C_{ij}^{(H)}}{H_0^2} + E_i E_j \frac{\sigma_{H_0}^2}{H_0^2}\ ,\end{aligned} \tag{3.16}$$

where we have used the fact that errors on H_0 and H_i are uncorrelated, therefore $\langle \delta H_0 \delta H_i \rangle = 0$. $C_{ij}^{(H)}$ is the covariance matrix of our data on the Hubble function $H(z)$ and σ_{H_0} is the error on H_0. Eq. (3.16) amounts to adding an extra covariance matrix to our data covariance matrix.

3.2.2. E_G data

We use the E_G data compiled in Tab. 3.6. These measurements are obtained using redshift-space distortions and galaxy-galaxy lensing as described in Sec. 2.2.2 and Sec. 2.2.3. Our compilation includes the measurements of the Red Cluster Sequence Lensing Survey (RCSLenS) [151] where the analysis combines the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS), the WiggleZ Dark Energy Survey and the Baryon Oscillation Spectroscopic Survey (BOSS). We also use the results of the VIMOS Public Extragalactic Redshift Survey (VIPERS) [147]. And the remaining measurements available are the results from the joint analysis of KiDS+2dFLenS+GAMA [138].

These sources provide measurements in real space within the scales $3 < R_p < 60h^{-1}$Mpc and in the linear regime, which is the one we are interested in. They have been obtained over a relatively narrow range of scales λ meaning that we can consider them relative to the $k = 2\pi/\lambda$-th Fourier component, as a first approximation. In any case, the discussion about the k-dependence of η is beyond the scope of this work, so the final result can

be seen as an average over the range of scales effectively employed in the observations. Moreover, in the estimate of E_G, based on [146], one assumes that the redshift of the lens galaxies can be approximated by a single value. With these approximations, indeed E_G is equivalent to $P_2/2$, otherwise, E_G represents some sort of average value along the line of sight. We caution that these approximations can have a systematic effect both on the measurement of E_G and on our derivation of η. In the next chapter, we quantify the level of bias possibly introduced by these approximations in our estimate. For further details and discussion, see references [147, 151] as well as Chapter 4.

In Fig. 3.1, a large discrepancy between ΛCDM and the data points is visible, which was also noted in [138]. In fact, they further notice that if one allows the cosmological parameters to vary, the data prefer a larger value for Ω_m. Since the data is correlated with $f\sigma_8$ data, one can see this as another perspective on the '$f\sigma_8$ tension.

3.2.3. $f\sigma_8$ data

The variable P_3 is recovered through the reconstruction $f\sigma_8(z)$ and its derivative as a function of redshift. A compilation of the available data for $f\sigma_8(z)$ can be found in Tab.3.7. This quantity is obtained through measurements of the redshift-space distortions (RSD) in the two point-correlation function of a galaxy survey, described in Sec. 2.2.3.

Our compilation includes measurements from the 6dF Galaxy Survey [152], the Subaru FMOS galaxy redshift survey (FastSound) [153], WiggleZ Dark Energy Survey [3], VIMOS-VLT Deep Survey (VVDS) [154], VIMOS Public Extragalactic Redshift Survey (VIPERS) [147, 155–157] and the Sloan Digital Sky Survey (SDSS) [2, 158–164]. Other works in the literature which perform RSD measurements only report $f\sigma_8(z)$ values indirectly, i.e. report f/b and $b\sigma_8$ values. Since we would have to assume something on the bias or the σ_8 relation, these works are not considered here, e.g. [165] and [166]. Furthermore, for numerical reasons, before applying any reconstruction method, we use the natural logarithm of the data, i.e. $\ln f\sigma_8(z)$, which allows us to compute the P_3 observable as a simple derivative with respect to $\ln a$.

Table 3.2.: $H(z)$ measurements compiled by [121] with the respective original references.

z	$H(z)$ (km/s/Mpc)	$\sigma_{H(z)}$ (km/s/Mpc)	Reference	Method	z	$H(z)$ (km/s/Mpc)	$\sigma_{H(z)}$ (km/s/Mpc)	Reference	Method
0.07	69	19.6	[120]	1	0.510	90.4	1.9	[2]	2
0.09	69	12	[119]	1	0.593	104	13	[115]	1
0.12	68.6	26.2	[120]	1	0.600	87.9	6.1	[3]	2
0.17	83	8	[119]	1	0.610	97.3	2.1	[2]	1
0.179	75	4	[115]	1	0.680	92	8	[115]	1
0.199	75	5	[115]	1	0.730	97.3	7	[3]	2
0.2	72.9	29.6	[120]	1	0.781	105	12	[115]	1
0.27	77	14	[119]	1	0.875	125	17	[115]	1
0.28	88.8	36.6	[120]	1	0.880	90	40	[118]	1
0.352	83	14	[115]	1	0.900	117	23	[119]	1
0.38	81.5	1.9	[2]	2	1.037	154	20	[115]	1
0.3802	83	13.5	[117]	1	1.300	168	17	[119]	1
0.4	95	17	[119]	1	1.363	160	33.6	[116]	1
0.4004	77	10.2	[117]	1	1.430	177	18	[119]	1
0.4247	87.1	11.2	[117]	1	1.530	140	14	[119]	1
0.44	82.6	7.8	[3]	2	1.750	202	40	[119]	1
0.4497	92.8	12.9	[117]	1	1.965	186.5	50.4	[116]	1
0.4783	80.9	9	[117]	1	2.340	222	7	[103]	3
0.480	97	62	[118]	1	2.360	226	8	[104]	3

Table 3.3.: $E(z)$ measurements from [1]. The error of the last measurement is not symmetric, therefore, it was recalculated as the quadrature of the 1σ bounds on the left and right side of the central value.

z	$E(z)$	$\sigma_{E(z)}$
0.07	0.997	0.023
0.2	1.111	0.020
0.35	1.128	0.037
0.55	1.364	0.063
0.9	1.52	0.12
1.5	2.67	0.675

Table 3.4.: Covariance matrix for the $H(z)$ data from [2].

z	Covariance matrix		
0.38	3.65	1.78	0.93
0.51	1.78	3.65	2.20
0.61	0.93	2.20	4.45

Table 3.5.: Covariance matrix for the $H(z)$ data from [3].

z	Covariance matrix		
0.44	0.0064	0.0025704	0
0.60	0.0025704	0.003969	0.00254016
0.73	0	0.00254016	0.005184

Table 3.6.: $E_G(z)$ data set. First column: redshift, second column: the E_G value and third column is the corresponding error. The fourth column shows the considered interval in real space that was used to obtain each data point and the last column points to the reference in the literature.

z	$E_G(z)$	$\sigma_{E_G(z)}$	Scale ($h^{-1}Mpc$)	Reference
0.267	0.43	0.13	$5 < R_p < 40$	[138]
0.305	0.27	0.08	$5 < R_p < 60$	[138]
0.32	0.40	0.09	$R_p > 3$	[151]
0.32	0.48	0.10	$R_p > 10$	[151]
0.554	0.26	0.07	$5 < R_p < 60$	[138]
0.57	0.31	0.06	$R_p > 3$	[151]
0.57	0.30	0.07	$R_p > 10$	[151]
0.60	0.16	0.09	$3 < R_p < 20$	[147]
0.86	0.09	0.07	$3 < R_p < 20$	[147]

Table 3.7.: $f\sigma_8(z)$ data with the corresponding redshift and error. The fourth column points to the reference in literature.

z	$f\sigma_8(z)$	$\sigma_{f\sigma_8(z)}$	Reference	z	$f\sigma_8(z)$	$\sigma_{f\sigma_8(z)}$	Reference	z	$f\sigma_8(z)$	$\sigma_{f\sigma_8(z)}$	Reference
0.067	0.423	0.055	[152]	0.44	0.416	0.080	[3]	0.727	0.296	0.077	[155]
0.15	0.49	0.15	[158]	0.51	0.458	0.038	[2]	0.73	0.437	0.072	[3]
0.17	0.51	0.06	[154]	0.55	0.444	0.038	[162]	0.80	0.47	0.08	[156]
0.25	0.3512	0.0583	[159]	0.57	0.488	0.060	[163], [164]	0.85	0.45	0.11	[157]
0.30	0.366	0.067	[160]	0.60	0.390	0.063	[3]	0.86	0.46	0.09	[147]
0.35	0.445	0.097	[161]	0.60	0.441	0.071	[160]	0.86	0.48	0.10	[147]
0.37	0.4602	0.0378	[159]	0.60	0.48	0.11	[147]	1.36	0.482	0.116	[153]
0.38	0.497	0.045	[2]	0.60	0.48	0.12	[147]				
0.40	0.394	0.068	[162]	0.61	0.436	0.034	[2]				

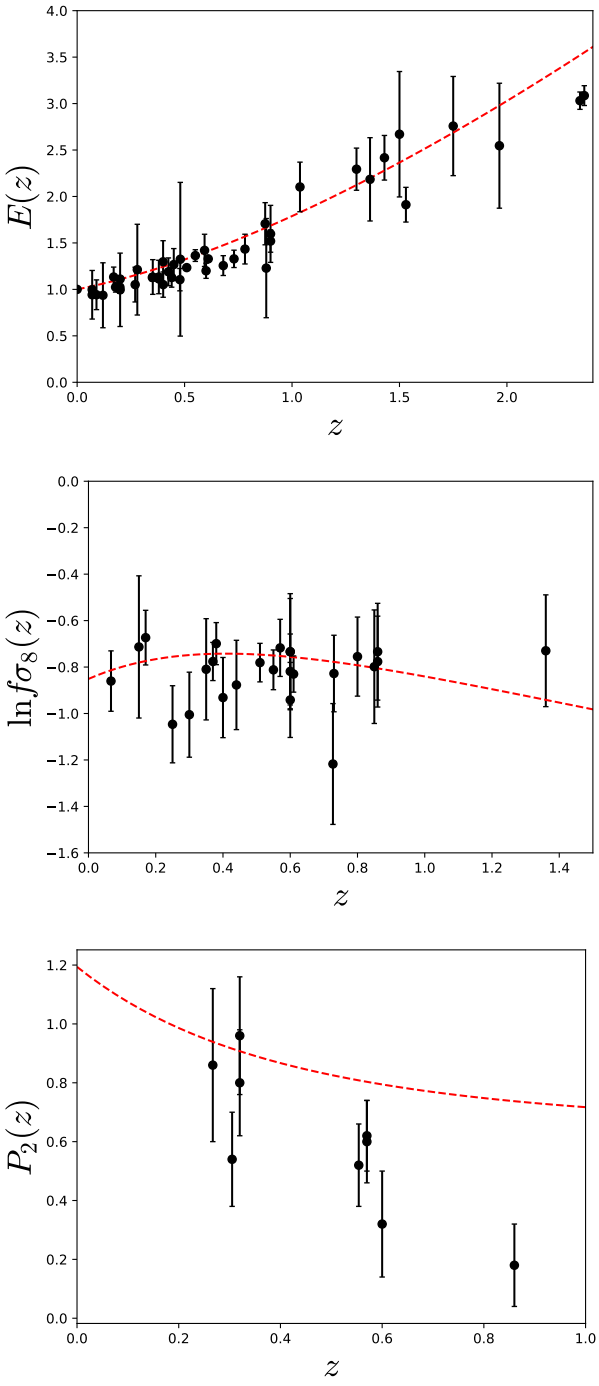

Figure 3.1.: The aforementioned data sets (black dots) displayed with the resepctive ΛCDM prediction as a function of redshift (dashed red line), using Planck 2018 values for the cosmological parameters as on Tab. 3.1. **Top panel:** $E(z)$ data from [121] where part of the data was rescaled using the H_0 from HST collaboration. **Middle panel:** Plot of the natural logarithm of the $f\sigma_8$ data points from Tab. 3.7. **Bottom panel:** Data set for P_2, obtained using E_G data from Tab. 3.6.

3.3. Data reconstruction methods

The need for data reconstruction of each observable stems from different data sets having different redshift values. In addition, we need to take derivatives for the case of $E(z)$ and $f\sigma_8(z)$. Therefore, we require to have a reliable method that allows us to access the underlying function of redshift. There is no easy answer to this problem since different methods include more or less model assumptions, require few or many parameters and these can severely impact the final reconstructed function as well as the final errors. For this reason, we consider and compare three methods to obtain $\eta_{\rm obs}$: binning, Gaussian Process, and generalized polynomial regression.

3.3.1. Binning

One way to reorganize the data in order to see its underlying behavior is to assemble the data into bins. This method consists of dividing the data into particular intervals (bins), which in this case are intervals of redshift. Then for each of these intervals, the average of the data contained in that bin is calculated. Denoting $s_k \equiv s(z_k)$ as a generic data value, with dependent variable s, located at the point z_k with error σ_k^s, the binning procedure is done by applying the following formula

$$\bar{s}_i = \frac{\sum_k^{N_i} s_k \left(\sigma_k^s\right)^{-2}}{\sum_k^{N_i} \left(\sigma_k^s\right)^{-2}}, \qquad \sigma_i^{\bar{s}} = \frac{1}{\sqrt{\sum_k^{N_i} \left(\sigma_k^s\right)^{-2}}} \quad , \tag{3.17}$$

where N_i is the number of data points inside the bin i, $\bar{s}_i$ is the new value of the dependent variable at the center of the bin z_i. The bin z_i is simply $z_i = (z_{k+1} - z_k)/2$, the arithmetic mean between the upper and lower borders of the bin. At this point, the error is updated to be $\sigma_i^{\bar{s}}$. This means that we are converting the information of the subset of data contained in a specific bin into one unique data point by taking the weighted average for the data values and the data errors over all points contained in that interval. The square of the new error at the center of the bin, $(\sigma_i^{\bar{s}})^2$, is the mean of the errors squared from all the N_i points contained in the bin with index i.

For our final result $\eta_{\rm obs}$, we also need to compute the derivatives of the data for the functions $E(z)$ and $\ln(f\sigma_8(z))$, at the exact same redshifts as for the other functions. Therefore, we need to bin the original data in different bins centered at new points z_j, such that by using finite differences we can compute the derivative of the dependent variable and its associated error at the z_i in the following form

$$\bar{s}_i' = -(1+z_i)\frac{\bar{s}(z_{j+1}) - \bar{s}(z_j)}{\Delta z_j} \, , \qquad \sigma_i^{\bar{s}'} = (1+z_i)\frac{1}{\Delta z_j}\sqrt{(\sigma_{j+1}^{\bar{s}})^2 + (\sigma_j^{\bar{s}})^2} \, , \tag{3.18}$$

where $\Delta z_j = z_{j+1} - z_j$. Note that a prime denotes a derivative with respect to $\ln a$.

Our observable $\eta_{\rm obs}$ is estimated as in Eq. (3.11) through $E(z)$, P_2, P_3, and $E'(z)$, which we denote generally by $y^{(1)}$, $y^{(2)}$, $y^{(3)}$ and $y^{(4)}$, respectively. Consequently, to calculate the final error on $\eta_{\rm obs}$, we use standard error propagation, assuming no correlation among the $y^{(i)}$ variables, so that the error $\sigma_i^{\eta_{\rm obs}}$ at the redshift z_i is specifically

$$(\sigma_i^{\eta_{\rm obs}})^2 = \sum_{\alpha=1}^{4} \left(\sigma_i^{y^{(\alpha)}} \frac{\partial \eta_{\rm obs}(z_i)}{\partial y^{(\alpha)}} \right)^2 , \tag{3.19}$$

where we also assume that the bins are large enough, such that the correlation among the bins is negligible. In this way, Eq. (3.11) and its estimated error can be evaluated at the centers of the bins z_i. However, the maximum number of final bins N_i is constrained by the number of data points available for the smallest data set among the $y^{(\alpha)}$ functions, which might be a disadvantage when working with datasets of a few measurements as it is the case of the E_G measurements. Also, the accuracy of this method is very sensitive to the bin size and thus it does not capture high-frequency modes in the data. We present results on the binning method with more detail in Sec. 3.4.

3.3.2. Gaussian Process

Another way of reconstructing a continuous function from a dataset is using the method of Gaussian Process (see [167] for a comprehensive description). A Gaussian Process (GP) can be regarded as the generalization of Gaussian distributions to the space of functions since it provides a probability distribution over continuous functions instead of a distribution over a random variable.

This method has been used several times in cosmology, especially for the determination of the equation of state of dark energy w and the Hubble function $H(z)$ (see, for example, [121, 150, 168–170]). It became recently a popular method for data reconstruction as it does not assume a concrete parametrization for the data itself. Instead, one assumes Gaussian distributed data and errors as well as a kernel function.

Considering a dataset $\mathcal{D} = \{(x_i, y_i)|i = 1, ...n\}$ of n observables where x_i are deterministic variables and y_i random variables, the goal is to obtain a continuous function $f(x)$ that best describes the dataset. A function f evaluated at a point x is a Gaussian random variable with mean μ and variance Var(f). The $f(x)$ values depend on the function value evaluated at another point x'. The relation between the value of the function at these two points is given by a covariance function $\mathrm{cov}(f(x), f(x')) = k(x, x')$, which evaluated at $x = x'$ gives the variance $\mathrm{Var}(f(x)) = k(x, x)$. Thus, the distribution of functions at the point x is characterized by (for more details, see [168])

$$\mu(f(x)) = \langle f(x) \rangle \ , \qquad k(x, x') = \langle (f(x) - \mu(x))(f(x') - \mu(x')) \rangle \ , \tag{3.20}$$

where the brakets correspond to the expected value.

The covariance function $k(x, x')$ is in principle arbitrary. Since we are interested in reconstructing the derivative of the data, we need to choose a differentiable function. In this work we choose a Gaussian covariance function

$$k(x, x') = \sigma_f^2 \exp\left[- \frac{(x - x')^2}{2\ell_f^2} \right] , \tag{3.21}$$

is the covariance function that we choose in this work, as it is the most common and it has the least number of parameters. In Sec. 3.4, we discuss how this assumption does not change considerably our results. This function depends on the hyperparameters σ_f and ℓ_f, which adjusts the shape of the covariance function and acts as a form of prior on the set of possible functions that we can obtain with the GP method. The hyperparameter ℓ_f can be considered as the typical correlation length scale of the independent variable, while the signal variance σ_f can be thought of as the typical variation scale of the dependent variable.

In a Gaussian Process using real data (x_i, y_i) where $y_i = f(x_i)+\epsilon_i$, the errors are assumed to be Gaussian and the observations to be scattered around the underlying function. The noise ϵ_i is Gaussian with a covariance matrix C, which needs to be taken into account for the joint likelihood function. This means that the reconstruction itself depends on the number and quality of the available data.

Following a Bayesian approach, one can compute the joint likelihood function for the data and the reconstructed function. Thus, for a Gaussian prior for both the data and the random functions, one can marginalize over the space of functions f and obtain the logarithm of the marginal likelihood as (see [168])

$$\begin{aligned}\ln \mathcal{L} = &-\frac{1}{2}\sum_{i,j=1}^{N}\left\{\left[y_i-\mu(x_i)\right]\left[k(x_i,x_j)+C(x_i,x_j)\right]^{-1}\left[y_j-\mu(x_j)\right]\right\} \\ &-\frac{1}{2}\ln\left|k(x_i,x_j)+C(x_i,x_j)\right| - \frac{N}{2}\ln 2\pi \ .\end{aligned} \tag{3.22}$$

Maximizing the logarithm of the marginal likelihood gives then the optimal hyperparameters σ_f and ℓ_f. In a full Bayesian approach, one should marginalize over the hyperparameters, using Monte Carlo Markov chain (MCMC) algorithms, to obtain the fully marginalized posterior distribution on the reconstructed function. As suggested in [168], we assume that the probability distribution of the hyperparameters is sharply peaked, which allows us to take them out of the integration and effectively fix them to their optimal values.

The Gaussian Process algorithm is implemented in a publicly available python code, named GaPP [168]. The GaPP code computes the continuous function of a given dataset and its derivatives up to third order, for a multi-dimensional dataset. It also takes into account correlated errors in the data and allows one to choose among different covariance functions, also known as kernel functions. For the case of the Gaussian kernel function as described above, the σ_f and ℓ_f parameters are optimized by the GaPP code through the maximization of the logarithm of the marginal likelihood function in Eq. (3.22).

Also, for the case of reconstructing the derivative of the data, there is a covariance between the reconstruction of f and f', which should also be determined by a Monte Carlo sampling. GaPP takes a first-order approximation and uses statistical error propagation which is valid for small errors. These approximations may have an impact on the final constraints of this work, particularly as underestimated errors on the reconstructed function as discussed in the original reference [168].

For each of the data sets, we use the GaPP code to reconstruct the underlying function and its derivative where we did not specify any prior on the hyperparameters or the mean function of the Gaussian Process to remain agnostic towards these choices. The details of our approach using this code concerning the chosen hyperparameters and covariance functions are further discussed in Sec. 3.4.

3.3.3. Polynomial regression

As a third reconstruction method, we use a generalized polynomial regression, a widely used method to obtain model parameters from data. Since we want to do this as model-independently as possible, we do not impose a priori any polynomial order for the reconstruction, but we let the data decide which is the maximum possible order. In the

following, we describe the standard method of polynomial regression. Nevertheless, several details need to be correctly accounted for when differentiating data with correlated errors. Thus we discuss the method in detail in Appendix A.

We start by assuming that we have N data points y_i, one for each value of the independent variable x_i (which are *not* random variables) and that

$$y_i = f_i + e_i \ , \tag{3.23}$$

where e_i are errors (random variables), which are assumed to be distributed as Gaussian variables. Here f_i are theoretical functions that depend linearly on parameters A_α

$$f_i = \sum_\alpha \bar{A}_\alpha g_{i\alpha} \ , \tag{3.24}$$

where $g_{i\alpha}(x_i)$ are functions of the variable x_i. This is the definition of a linear model. Defining the matrix of basis functions as G and the data vector as D in the following way (always summing over repeated Latin indexes)

$$G_{\alpha\beta} \equiv g_{\beta i} C^{-1}_{ij} g_{\alpha j} \ , \tag{3.25}$$

$$D_\alpha \equiv y_i C^{-1}_{ij} g_{\alpha j} \ , \tag{3.26}$$

where C_{ij} is the data covariance matrix. We see that the linear model can be written as

$$\mathbf{GA} = \mathbf{D} \ . \tag{3.27}$$

We are interested in finding the coefficients $\boldsymbol{A} = \{A_0, A_1, ...\}$ of the model. For that, we invert the above equation and solve for $\boldsymbol{A}$,

$$\bar{\mathbf{A}} = \mathbf{G}^{-1}\mathbf{D} \ , \tag{3.28}$$

which is also known as the *normal equation*.

If the prior is uniform in an infinite range (improper prior), the parameters in the linear problem have a Gaussian posterior with mean $\bar{\mathbf{A}}$ and correlation matrix given by the inverse of its Fisher matrix. Since in the linear problem the data covariance matrix does not depend on the parameters, we have the Fisher matrix

$$F_{\alpha\beta} \equiv C^{-1}_{ij} \frac{\partial f_i}{\partial \bar{A}_\alpha} \frac{\partial f_j}{\partial \bar{A}_\beta} = C^{-1}_{ij} g_{\alpha i} g_{\beta j} = G_{\alpha\beta} \ . \tag{3.29}$$

Once the coefficients are known, we can obtain the data values on any point x_A, even if not present in the data. For that, we use Eq. (3.24) and evaluate it at x_A, namely $f_A = \sum_\alpha \bar{A}_\alpha g_{A\alpha}$, where $g_{A\alpha}$ means the function g_α evaluated at x_A, with an error $\sigma^2_A = F^{-1}_{\alpha\beta} g_{A\alpha} g_{A\beta}$. We can select a number of arbitrary points $x_{A,B,C}$ and obtain the error matrix for the reconstructed function at these points as

$$C_{AB} = F^{-1}_{\alpha\beta} g_{A\alpha} g_{B\beta} \ . \tag{3.30}$$

In our particular case, we have three datasets $(y^{(0)}, y^{(1)}, y^{(2)}) = (\ln(fs_8(z)), E(z), E_G(z))$ and we wish to estimate the error on a function $\eta_{\rm obs}(y^{(1)}, y^{(2)}, y^{(3)}, y^{(4)})$, where $y^{(4)} = y^{(1)\prime}$ and $y^{(3)} = y^{(0)\prime}$ where a prime denotes, as already mentioned, a derivative with respect to $\ln a$. The remaining details are in Appendix A with the exception of the choice

of the order of the polynomial. In principle, the order is arbitrary, up to the number of data points for each data set. However, it is clear that with too many free parameters the resulting χ^2 will be very close to zero, which is statistically unlikely. Also, too many parameters render a numerical Fisher matrix computationally unstable (producing, e.g., a non-positive definite matrix) and the polynomial function is wildly oscillating. On the other hand, too few parameters restrict the allowed family of functions. Therefore, we select the order of the polynomial function by choosing the polynomial degree for which the reduced chi-squared $\chi^2_{\rm red} = \chi^2/(N-P)$, is closest to 1 and such that the Fisher matrix is positive definite. Since our datasets contain data points from different experiments, there are some data points located at the same redshift or very close to each other, with different values of the dependent variable. In the case of a perfect fit, the polynomial would go through all points leading to spurious oscillations. For this reason, we take the weighted average of data points that are closer than $\Delta z = 0.01$ in redshift, before using them as an input into the polynomial regression algorithm.

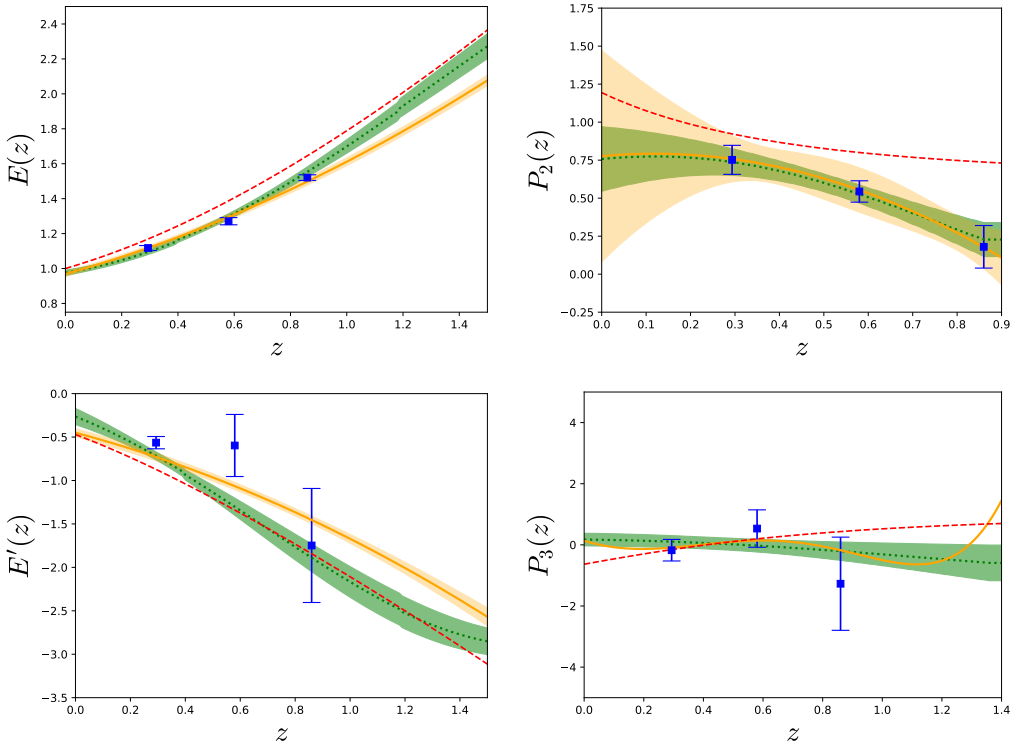

Figure 3.2.: Comparison of the three reconstruction methods for each of the model-independent variables. The binning method in blue squares with error bars, Gaussian Process as a green dotted line with green bands, polynomial regression as a solid yellow line with yellow bands. All of them depicting the 1σ uncertainty. **Left panel:** Plot of the reconstructed $E(z)$ function on the top and its derivative $E'(z)$ on the bottom. **Right panel:** Plot of the reconstructed $P_2(z)$ function on the top and the reconstructed $P_3(z)$ function on the bottom. For each case, we show the theoretical prediction of our reference ΛCDM model as a red dashed line.

3.4. Results

We now discuss the results of the final observable $\eta_{\rm obs}$ for each of these methods. In comparison, the binning method contains the least number of assumptions compared to the polynomial regression or the Gaussian Process method. It is essentially a weighted average over the data points and its error bars at each redshift bin. To calculate P_3 and E' implies that we take derivatives using the finite difference method. The caveat of this approach is that it introduces correlations among the errors of the function and its derivatives that we cannot take into account with this simple binning method. Furthermore, for the binning method, we do not take into account possible non-diagonal covariance matrices for the data, which we do for polynomial regression and the Gaussian Process reconstruction.

Fig. 3.2 shows the reconstructed functions obtained by the binning method, the Gaussian Process, and polynomial regression, alongside with the theoretical prediction of the standard ΛCDM model. In all cases, the error bars or the bands represent the 1σ uncertainty.

In the binning method, the number of bins is limited by the maximum number of existing data redshifts from the smallest data set corresponding to one of our model-independent observables. In this case, this is the quantity E_G, for which we have effectively only three redshift bins. Looking at Tab. 3.6 and comparing with Fig. 3.1, we can see that there are nine E_G data points, but they are very close to each other in redshift, due to being measured by different collaborations or at different scales in real space for the same z. As explained in the data section above, we regard this data as an average over different scales, assuming that non-linear corrections have been correctly taken into account by the respective experimental collaboration. Thus, this leaves us with three possible redshift bins, centered at $z_1 = 0.294$, $z_2 = 0.580$ and $z_3 = 0.860$, all of them with an approximate bin width of $\Delta z \approx 0.29$. At these redshifts, we obtain $\eta_{\rm obs}(z_1) = 0.48 \pm 0.45$, $\eta_{\rm obs}(z_2) = -0.03 \pm 0.34$ and $\eta_{\rm obs}(z_3) = -2.78 \pm 6.84$. These values and the estimate of the intermediate model-independent quantities are shown in Tab. 3.9.

Regarding the Gaussian Process method, we have computed the dimensionless Hubble function and its derivative, $E(z)$ and $E'(z)$, with the *dgp* module of the GaPP code. Using the data of Tab. 3.3 and its correlation matrix, we reconstructed $E(z)$ and $E'(z)$ for the correspondent redshift interval using the Gaussian function as the covariance function and initial values of the hyperparameters $\theta = [\sigma_f = 0.5, \ell_f = 0.5]$ that later are estimated by the code. The same procedure was done for the $P_2(z)$ data, obtained by Eq. (3.13) using Tab. 3.6. We obtain for $E(z)$ and $E'(z)$ the hyperparameters $\sigma_f = 2.12$ and $\ell_f = 2.06$ and for the P_2 function, $\sigma_f = 0.58$ and $\ell_f = 0.67$.

For the $P_3(z)$ observable, the hyperparameters obtained by the GaPP code led to a very flat and unrealistic reconstruction, that suggested us to take another approach for obtaining the optimal hyperparameters. Therefore, we have performed further tests of the GaPP code with the $E(z)$ data, $f\sigma_8$, and $\ln(f\sigma_8(z))$. This approach consists in sampling the logarithm of the marginal likelihood of a grid of chosen hyperparameters and find the values that maximize it, instead of letting the code determine the best-fit hyperparameters. We chose a grid of the hyperparameters σ_f, ℓ_f with 300 logarithmically spaced values, from 0.01 to 2 for σ_f and from 0.01 to the maximum redshift of the dataset for ℓ_f. Remember that the hyperparameter ℓ_f constrains the typical scale on the independent variable z. For this reason, we impose that ℓ_f needs to be smaller than the redshift range of the data as an additional prior, which was not guaranteed by the default GaPP code. Then we chose

the pair of hyperparameters corresponding to the maximum of the log-marginal likelihood Eq. (3.22).

For the $E(z)$ data, we find that there is no significant change between the reconstructions given by the default GaPP optimization routine and our grid approach. However, for the $f\sigma_8(z)$ and $\ln(f\sigma_8(z))$ data visibly different reconstructions arise. In our work, we used the $\ln(f\sigma_8(z))$ data and thus we show in Fig. 3.3 the reconstructed functions with its uncertainty band in green for the different hyperparameter choices plotted together with the data used as black error bars and the theoretical expectation from the ΛCDM model for comparison. On the left side, we see that GaPP estimates a best-fit correlation length of $\ell_f = 288$, which yields a very flat reconstruction of $\ln(f\sigma_8(z))$. On the right side of Fig. 3.3, we see that by setting a prior on the correlation length, that is from 0.01 to 1.36, we recover a function that follows much better the general data trend with $\ell_f = 1.36$ than the hyperparameters set by the GaPP optimization routine.

Therefore, for the $\ln(f\sigma_8(z))$ data using our grid approach, we obtain $\sigma_f = 0.549$ and $\ell_f = 1.361$. Its reconstructed derivative P_3 is in the lower right panel of Fig. 3.2. The function remains relatively flat, compared to the one given by other methods, but this approach gives more reasonable results than the ones given by the default GaPP code. We acknowledge that the grid approach also has its limitations and that the ideal scenario would be to run a specially designed Monte Carlo Markov chain algorithm, not only for this specific dataset but for all of them as well as for the choice of the kernel function.

Regarding the choice of the kernel function, several functions were compared, each of them with a different number of parameters to see the impact on the output. We tested the Gaussian kernel with two parameters, (σ_f, ℓ_f); the rational quadratic kernel with three parameters and the double Gaussian kernel with four parameters (see the original reference for the explicit implemented formula [168]). We performed tests using the $H(z)$ data obtained with the cosmic chronometer technique and the $f\sigma_8(z)$ data. Our tests show that the different choices shift the reconstructed function up to 6% on its central value compared to the Gaussian kernel function. This happens for $H(z)$ while the effect is negligible for $f\sigma_8(z)$. Taking into account the above choices and procedure, we report that with the Gaussian Process method we obtain $\eta_{\rm obs}(z_1) = 0.38 \pm 0.23$, $\eta_{\rm obs}(z_2) = 0.91 \pm 0.36$ and $\eta_{\rm obs}(z_3) = 0.58 \pm 0.93$.

For the polynomial regression method, we find $\eta_{\rm obs}(z_1) = 0.57 \pm 1.05$, $\eta_{\rm obs}(z_2) = 0.48 \pm 0.96$ and $\eta_{\rm obs}(z_3) = -0.11 \pm 3.21$. Note that we applied the criteria of a $\chi^2_{\rm red}$ closest to one and a positive definite Fisher matrix to chose the order of the polynomial for each of the datasets. These criteria led to a choice of a polynomial of order 3 for the $E(z)$ and $E_G(z)$ data and order 6 for the $\ln(f\sigma_8(z))$ data. These polynomials illustrated in Fig. 3.2 as solid yellow lines, plotted with the corresponding 1σ uncertainty bands. The higher order of the polynomial of $\ln(f\sigma_8(z))$ explains the "bumpiness" of the reconstruction of P_3, leading to larger errors on this observable in comparison to the GP method. This "bumpiness" also makes the reconstruction generally more conservative. Although we expect a smoother reconstruction for this quantity, this actually captures the behavior of the dataset. This means that, in comparison with Fig. 3.1, the reconstruction of the underlying function is very dependent on the assumption of the shape of the function of redshift or a specific model like ΛCDM. This expectation of a smooth function can sometimes be misleading when one aims for understanding the meaning of the data.

Fig. 3.4 shows the reconstructed $\eta_{\rm obs}$ as a function of redshift with the three different

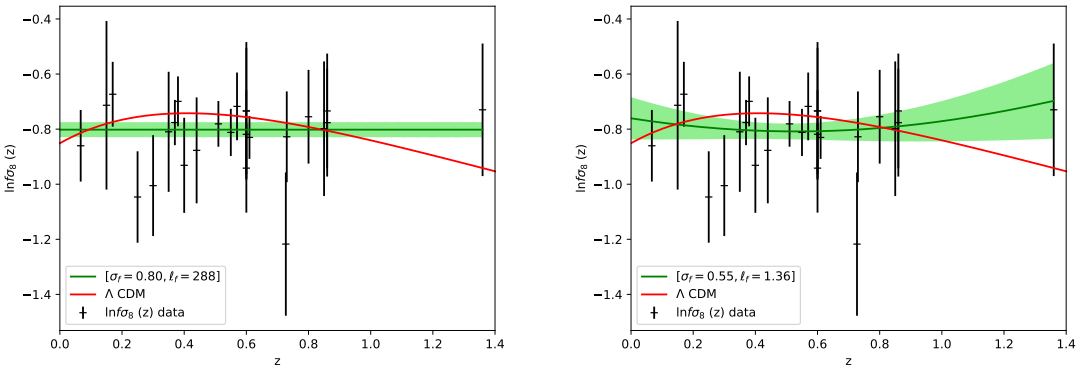

Figure 3.3.: Reconstrunction of the $\ln(f\sigma_8(z))$ data by the Gaussian Process method using the best-fit hyperparameters by the GaPP code (left panel) and using a grid in hyperparameter space with a prior on ℓ_f (right panel).

methods, again with GP in green dashed line, polynomial regression in a yellow solid line and the binning method in blue squares with error bars. We conclude that, in general, the methods are consistent with each other within their 1σ uncertainties with the exception of the second bin between the binning and the GP methods. Also, in most bins the results are consistent with the standard gravity scenario with the exception of the first bin for the binning and the Gaussian processes methods. These exceptions may come exactly from the different nature of these two methods. On one side, the binning method gives priority to the weighted average of the quantity as well as its correspondent weighted average error. This process is also very biased by the number of bins, which in our case are just 3, meaning one cannot infer anything else besides a generic trend. On the other side, the Gaussian shape assumed in the GP leads to a coherent and consistent smooth behavior throughout the range in question. As pointed out before, the polynomial regression is sensitive to the "bumpiness" of the data, ending up gathering both the behaviors from the other two methods. Nevertheless, this figure shows clearly how different data reconstruction methods can lead to different conclusions about the reconstructed function as well as the consistency of the analyzed data set.

We find that the error bars of the Gaussian Process reconstruction are generally smaller than the other methods, such that at the lowest redshift, GP is not compatible with $\eta_{\rm obs} = 1$ at nearly 2σ, while in the case of the binning method at the intermediate redshift, $z = 0.58$, the tension is nearly 3σ. These mild tensions can come from the already known data tensions of these datasets but also may be method dependent, which does not allow for strong statements on deviations of General Relativity.

As previously mentioned in Sec. 3.2.1, we need to choose the H_0 value in order to obtain the dimensionless Hubble function $E(z)$. Our model-independent estimate for η requires the dimensionless Hubble function, therefore, part of our data set needs to be converted from $H(z)$ into $E(z)$. However, as already described in Sec. 2.2.5, there is a statistically significant tension between the values measured by different probes, namely the value from the 2018 results of the Planck collaboration [83], which is $H_0^{\rm Planck} = 67.36 \pm 0.54\,[\rm km/s/Mpc]$, and the value from HST collaboration [149], which is $H_0^{\rm HST} = 73.45 \pm 1.66\,[\rm km/s/Mpc]$. The estimate of η shifts with the choice of H_0 as shown in Tab. 3.8. There is no significant change on the mean value and uncertainty. However, for the last

bin of the binning case, the uncertainty increases by a factor of 10. While this result is compatible with the other bins, it shows how sensitive the binning method is. Since our aim is to have a model-independent estimate of η, we chose the HST collaboration H_0 value due to being independent of a cosmological model.

Table 3.8.: The reconstructed $\eta(z)$ using different values of H_0 to normalize the $H(z)$ data at three different redshifts $z = (0.294, 0.58, 0.86)$ with its respective 1σ errors, for each of the reconstruction methods.

Method	Choice of H_0	$\eta(z)$		
		$z_1 = 0.294$	$z_2 = 0.58$	$z_3 = 0.86$
Binning	H_0 HST	0.48 ± 0.45	-0.03 ± 0.34	-2.78 ± 6.84
	H_0 Planck 2018	0.56 ± 0.54	-0.14 ± 0.32	-6.75 ± 75.64
GaPP	H_0 HST	0.49 ± 0.25	0.94 ± 0.33	0.27 ± 0.67
	H_0 Planck 2018	0.31 ± 0.22	0.72 ± 0.33	0.36 ± 0.79
Linear Regression	H_0 HST	0.57 ± 1.05	0.48 ± 0.96	-0.11 ± 3.21
	H_0 Planck 2018	0.51 ± 1.07	0.37 ± 0.93	-0.18 ± 3.11

Table 3.9.: The reconstructed model-independent variables $E, E', P_2, P_3, \eta(z)$ with their 1σ errors, at three different redshifts $z = (0.294, 0.58, 0.86)$, for each of the reconstruction methods. The polynomial regression method is compatible with the ΛCDM scenario while the other two methods show some tension at lower redshift.

Method	Parameter	Redshift bins			Weighted mean
		$z_1 = 0.294$	$z_2 = 0.58$	$z_3 = 0.86$	
Binning	$E(z)$	1.12 ± 0.01	1.27 ± 0.02	1.51 ± 0.02	
	$E'(z)$	-0.56 ± 0.07	-0.60 ± 0.36	-1.75 ± 0.66	
	$P_2(z)$	0.75 ± 0.10	0.54 ± 0.07	0.18 ± 0.14	
	$P_3(z)$	-0.17 ± 0.35	0.53 ± 0.61	-1.27 ± 1.52	
	$\eta_{\text{obs}}(z)$	0.48 ± 0.45	-0.03 ± 0.34	-2.78 ± 6.84	0.15 ± 0.27
Gaussian Process	$E(z)$	1.10 ± 0.01	1.30 ± 0.02	1.55 ± 0.03	
	$E'(z)$	-0.73 ± 0.05	-1.30 ± 0.10	-1.89 ± 0.16	
	$P_2(z)$	0.74 ± 0.09	0.53 ± 0.06	0.23 ± 0.11	
	$P_3(z)$	-0.10 ± 0.20	-0.03 ± 0.21	-0.21 ± 0.30	
	$\eta_{\text{obs}}(z)$	0.38 ± 0.23	0.91 ± 0.36	0.58 ± 0.93	0.53 ± 0.19
Polynomial Regression	$E(z)$	1.12 ± 0.01	1.29 ± 0.02	1.50 ± 0.02	
	$E'(z)$	-0.73 ± 0.04	-1.06 ± 0.04	-1.45 ± 0.04	
	$P_2(z)$	0.76 ± 0.15	0.55 ± 0.15	0.18 ± 0.14	
	$P_3(z)$	-0.09 ± 0.80	0.14 ± 0.78	-0.17 ± 3.02	
	$\eta_{\text{obs}}(z)$	0.57 ± 1.05	0.48 ± 0.96	-0.11 ± 3.21	0.49 ± 0.69

Finally, we combine the estimates at three redshifts of Tab. 3.9 into a single value. Assuming a constant η_{obs} in the entire observed range and performing a simple weighted average, we find $\eta_{\text{obs}} = 0.15 \pm 0.27$ (binning), $\eta_{\text{obs}} = 0.53 \pm 0.19$ (Gaussian Process) and $\eta_{\text{obs}} = 0.49 \pm 0.69$ (polynomial regression).

In conclusion, the Gaussian Process method yields the smallest error. It is sometimes presented as model-independent which is not fully the case. The Gaussian assumption can have a big impact on the reconstruction if the data is not exactly Gaussian. Also, the kernel function depends on a small number of parameters, which can be complicated

to constrain and deeper understanding of the implications of the choice of this function is needed. Both GP and the binning method, taken at face value, would rule out standard gravity. However, not taking into account the correlation induced by the finite differences may lead to a smaller overall error. Overall, we think the polynomial regression method provides a good compromise between the least number of assumptions and correcting accounts for correlations on the data and its derivative. Therefore, it is considered our main result.

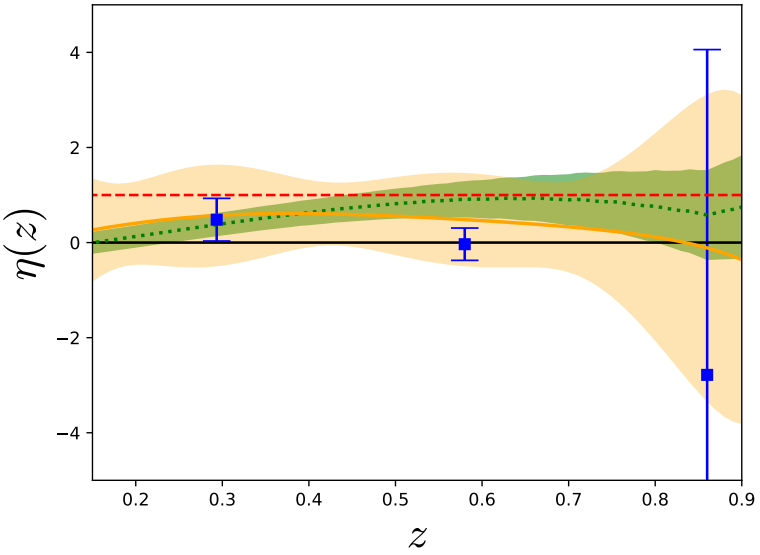

Figure 3.4.: Plot of the reconstructed $\eta_{\rm obs}$ as a function of redshift, using the binning method (blue squares), Gaussian Process (green dotted line) and polynomial regression (yellow solid line). The corresponding error bands (error bars for the binning method), represent the 1σ estimated error on the reconstruction. As a reference, the dashed shows red line the value in standard gravity.

3.5. Summary

The gravitational-wave event gave more evidence for standard gravity and constrained, or even ruled out, a few modified gravity theories. Despite that, further consistency checks and tests of deviations of General Relativity are needed. Given the current available data, the model-independent expression of the anisotropic stress η becomes an interesting task to carry out. In this work, we estimated η as proposed in [65]. If a strong deviation would be detected, that would imply the perfect fluid approach does not hold, giving rise to the need to modified gravity or other assumptions usually done in ΛCDM.

We used various datasets ranging from cosmic chronometers data as well as some Supernovae Ia measurements for the Hubble parameter $H(z)$, alongside weak lensing and redshift space distortions for $f\sigma_8(z)$ and E_G(z) to estimate the anisotropic stress $\eta_{\rm obs}$. No assumptions of background cosmology, galaxy bias, initial conditions, and matter abundance were made. A model-independent estimation is a tough task since there is no unique way of performing it [171–173]. We obtained this estimate using three distinct data reconstruction methods: binning, Gaussian Process and polynomial regression. Each makes

different assumptions which led to distinctive results for the value of $\eta_{\rm obs}$. We took into account these assumptions when analyzing our results as there is no unique and universal method for data reconstruction. The polynomial regression method reveals to be the most conservative in a good compromise with a small number of assumptions. We find $\eta_{\rm obs} = 0.44 \pm 0.92$ at $z = 0.294$, $\eta_{\rm obs} = 0.42 \pm 0.89$ at $z = 0.58$, and $\eta_{\rm obs} = -0.14 \pm 3.01$ at $z = 0.86$.

Generally, the reconstruction methods agree with each other. All results are compatible for the first bin but with a trend for a lower value. In the second bin, the binning method is 1.5σ away from the GP. For the last and farthest bin, the errors are much larger as the data scatter strongly from the standard scenario. In some cases, the standard gravity value $\eta = 1$ is two or even three sigma away from our result, but it is in every bin compatible for at least two of the three methods. These results are not statistically strong to claim deviations of GR. The deviations found can result from the assumptions of the reconstruction methods or from the data tensions present between CMB data and low redshift data.

4. Review of the E_G statistics

Testing gravity in a model-independent way is of utter importance to find a theory that describes the Universe, among several theories. But it is a highly challenging task since most of the observables rely on simple ingredients such as, for instance, the distance to a galaxy, which at cosmological scales depending on the considered model. In the last chapter, we explain how the estimate $\eta_{\rm obs}$ is a good example to test the perfect fluid assumption in General Relativity. This chapter is about the E_G statistics, an estimator build to be *a discriminating probe of gravity*. As such, several times the E_G statistics was used or measured as a test of modified gravity [146–148, 174]. The E_G statistics tries to capture a modified gravity signature through a ratio of information from galaxy clustering and weak lensing obeservables that are sensitive to particular features of a theory of gravity but it is not a model-independent test. A *smoking gun* for gravity should be independent of details from the most accepeted model in order to correctly evaluate the properties of gravity. Moreover, the theoretical definition of E_G is only equivalent to the observational definition under a flat ΛCDM cosmology scenario. In this chapter, we will explore the definition of the E_G statistics and its assumptions. Our goal is to check if the equivalence between definitions holds for possible modified gravity models and if a scale dependence exists.

The E_G statistics was first proposed in 2007 by Zhang *et al.* [148]. The idea behind this statistics is the following. Given the perturbed Friedman-Lemaître-Robertson-Walker metric of Eq. (2.19), lensing is sensitive to the term $\nabla^2(\Phi - \Psi)$ along the line of sight. In the case of General Relativity and given that there are no anisotropic stresses, we get $\Phi = -\Psi$ and lensing sensitive to $\nabla^2\Phi$. This is proportional to the fractional overdensity δ through the Poisson equation, making lensing directly related to δ. Testing this prediction allows distinguishing between models of gravity as it does not hold for all modified gravity scenarios. In general, to discriminate models is not easy since modified gravity usually tries to mimic the expansion history of the universe $H(z)$ from ΛCDM. It can be possible to differentiate models with gravitational slip because that lead to a different growth of structure and effects from gravitational lensing. This is the reason why this E_G quantity can be a good test for finding the model that describes the dynamics of the universe.

4.1. Theoretical definition of E_G

The E_G statistics links the gravitational potentials related to the gravitational lensing with the matter overdensity. Consequently, if the same population of galaxies is being used, the galaxy bias cancels out and there is no dependence on initial matter fluctuations. The E_G statistics is defined by

$$E_G = \left\langle \frac{a\nabla^2(\Psi - \Phi)}{3H_0^2 f\delta} \right\rangle_k . \tag{4.1}$$

For the ΛCDM model, the expected value simplifies to $E_G = \Omega_{m0}/f$.

However, it is not a model-independent test because Ω_{m0} and f are not observable quantities. First, Ω_{m0} can be obtained through Supernovae Type Ia by first measuring $H(z)$. Then, one needs to assume a flat space and an equation of state for matter and dark energy. These assumptions may not hold for modified gravity models. Furthermore, the "dark degeneracy" discussed, for instance, in [175], shows that the separation between a matter component and dark energy component is unavoidably model dependent.

Second, the growth rate f is estimated by solving the differential equation of the perturbation growth for matter. This requires initial conditions, that are normally taken to be cold dark matter at high redshift. Again, this assumption is not necessarily true in modified gravity, for example, in the original Brans-Dicke model.

Consequently, any discrepancy between the theoretical value of E_G to its observed value does not necessarily mean the detection of modified gravity or deviation from ΛCDM. This contrast can be due to a different Ω_{m0} or different initial conditions. It can also be due to a genuine signature of a non-standard modified gravity parameter Σ. Therefore, E_G can be applied to investigate specific models by comparing each particular prediction with observations, for instance, ΛCDM model, but not to directly probe gravity. This is a contrasting scenario with respect to η_{obs} Eq. (3.11) that is model-independent by estimating directly η without assuming a value of Ω_{m0} nor for f. Thus, if one finds $\eta_{\rm obs} \neq 1$ with data, then ΛCDM and any model of standard gravity where dark matter is a perfect fluid are ruled out.

Finally, we would like to point out the cautionary remark done in [138], namely that their results about E_G should not be employed until the tension between Ω_{m0} in different observational datasets is resolved. This problem does not arise at the estimation of the anisotropic stress parameter $\eta_{\rm obs}$, on the Chapter 3.

On the initial proposal [148], the authors assumed LAMOST, SKA and LSST specifications with WMAP best-fit parameters and compared ΛCDM, flat DGP and $f(R)$ gravity models. They claim that the differences are substantial and, therefore, future measurements of E_G would be able to detect a deviation from General Relativity. Nevertheless, they note that the accuracy of the measurements of the distance D and H can produce shifts in E_G. These shifts are negligible if D is measured with a 1% accuracy but the same accuracy in H translates in an up to 3% fractional error in E_G. Later on, the cosmic variance errors dominate at small scales and this estimate also depends on the survey parameters of the fraction of the sky coverage and the redshift of the lensing sources. These assumptions in sum may undermine the overall accuracy of this quantity to indeed be a test for gravity.

The first measurement of E_G came in 2010 [176] and the second measurement in 2015 [151]. These measurements follow a slightly different definition of E_G, whereas $\beta = f/b$. At linear scales and if the bias b is constant, the definition of Eq. (4.1) is recovered. However, these assumptions and other parameters led to further exploration of the difference between definitions by [146]. They found that these definitions are indeed equivalent in the case of a flat ΛCDM cosmology, being often the reference for more recent E_G measurements. However, they notice four potential sources of theoretical uncertainty. These sources are the distribution of source galaxies P_s, the projection length δ_ℓ, the cut-off scale R_0, and the bias b, from which only P_s does not affect E_G. The equivalence between these two definitions appears to be dependent on specific conditions such as the previous functions which may invalidate the E_G measurements and the conclusions drawn from them. In other words, since we don't have any knowledge about the bias function and the E_G definitions

are only equivalent for flat ΛCDM, then E_G might not be the correct way to distinguish models. For this reason, we aim to compute E_G and understand better what it is as well as the assumptions behind it.

4.2. Towards an observationally motivated general definition

Our goal in this section is to revisit the E_G defintions and see under which conditions one can extract

$$E_{G0} \equiv \frac{\Omega_{m0}\Sigma}{2f} = \frac{P_2}{2}\,. \tag{4.2}$$

We define the lensing modified gravity parameter $\Sigma = Y(1+\eta)$ which is 2 in standard gravity.

Let us assume that we have a population of lenses in a redshift range Δz_ℓ and a population of sources at a higher redshift range Δz_s. Following the approach of [146], we assume that the lenses are at a known z_ℓ and the sources are confined in a small range z_s. The definition of [146] is

$$E_G = \frac{bY_{gm}(R)}{fY_{gg}(R)}\,, \tag{4.3}$$

where b is the bias function and f the growth rate function that are evaluated at the average lens redshift z_ℓ. We use $8\pi G = 1$. The galaxy-lens correlation function in the numerator is integrated over a ring on the sky with radius from R_0 (the cut-off scale) to R,

$$\begin{aligned} Y_{gm}(R) =& \frac{\rho_{c0}\Omega_{m0}}{4}\int_0^\infty \frac{\mathrm{d}z}{H(z)} r(z_\ell)r(z)(1+z_\ell)(1+z)\int_z^\infty \frac{\mathrm{d}z'}{H(z')} W_s(z')\frac{[r(z')-r(z)][r(z')-r(z_\ell)]}{r(z')^2} \\ &\times \int_{R_0}^R \mathrm{d}R' K(R,R')\xi_{gm}(R',z_\ell,z)\,, \end{aligned} \tag{4.4}$$

where ρ_{c0} is the critical density today, $r(z)$ is the comoving distance and R is a projected distance on the sky. Taking a realistic normalized redshift distribution of sources $P_s(z,z_0)$, namely,

$$P_s(z,z_0) = \frac{z^\alpha}{N_0} e^{-\left(\frac{z}{z_0}\right)^\beta}\frac{H(z)}{c}, \tag{4.5}$$

where we set $\alpha = 2$, $\beta = 1.5$, $z_0 = 0.9/1.412$, and N_0 is the normalization factor as $\int(\mathrm{d}z/H(z))P_s(z) = 1$. We also use $c = 1$. The window function W_s is then defined as

$$W_s(z) = \begin{cases} \dfrac{P_s(z)}{\int_{z_\ell}^\infty \frac{\mathrm{d}z_s}{H(z_s)}P_s(z_s)\left[\frac{r_\ell(r(z_s)-r_\ell)(1+z_\ell)}{r(z_s)}\right]^2}\,, & z \geq z_\ell, \\ 0, & z < z_\ell, \end{cases} \tag{4.6}$$

Note that we use a slightly different notation then [146]. We also need the kernel function $K(R,R')$ described by

$$K(R,R') = \frac{2}{R^2}R' - \delta_D(R'-R) + \left(\frac{R_0}{R}\right)^2\delta_D(R'-R_0). \tag{4.7}$$

The denominator of Eq. (4.3) is the galaxy-galaxy lensing that is expressed by

$$Y_{gg}(R) = \rho_{c0}\int_{z_\ell-\Delta_\ell}^{z_\ell+\Delta_\ell}\frac{dz}{H(z)}\int_{R_0}^R \mathrm{d}R' K(R,R')\xi_{gg}(R',z_\ell,z)\,, \tag{4.8}$$

where Δ_ℓ is some interval around z_ℓ beyond which we can neglect the clustering effect. For this, we use the Fourier transform of the power spectrum $P(k)$, which is the matter correlation function ξ_{mm} that depends on the distance $d = \sqrt{R^2 + [r(z) - r(z_\ell)]^2}$. We can rewrite as

$$\xi_{mm}(r) = \frac{1}{2\pi^2}\int_0^\infty k^2 e^{-k^2} P(k)\frac{\sin(kr)}{kr}\mathrm{d}k, \tag{4.9}$$

and if the kernel function is taken into account, then we get

$$\hat{\xi}_{mm} = \frac{1}{2\pi^2}\int_0^\infty \mathrm{d}k k^2 e^{-k^2} P(k)\hat{J}_0(k, R, z, z_\ell). \tag{4.10}$$

With this definition, the integral can be done analytically,

$$\begin{aligned}\hat{J}_0(k, R, z, z_\ell) &\equiv \int_{R_0}^{R} \mathrm{d}R' K(R, R') J_0(k, R', z, z_\ell) \qquad (4.11)\\ &= \tfrac{2}{R^2}\textstyle\int_{R_0}^{R} \mathrm{d}R' R' J_0(k, R', z, z_\ell) - J_0(k, R, z, z_\ell) + \left(\frac{R_0}{R}\right)^2 J_0(k, R_0, z, z_\ell)\\ &= \tfrac{2}{R^2}\left[\tfrac{\cos(k\sqrt{R_0^2+[r(z)-r(z_\ell)]^2})}{k^2} - \tfrac{\cos(k\sqrt{R^2+[r(z)-r(z_\ell)]^2}}{k^2}\right] - J_0(k, R, z, z_\ell) + \left(\tfrac{R_0}{R}\right)^2 J_0(k, R_0, z, z_\ell)\end{aligned}$$

where

$$J_0(k, R, z, z_\ell) = \frac{\sin\left(k\sqrt{R^2 + [r(z) - r(z_\ell)]^2}\right)}{k\sqrt{R^2 + [r(z) - r(z_\ell)]^2}}. \tag{4.12}$$

Typically, b, G, and Σ are k-independent functions but that is not necessarily true. Gathering the previous expressions, the general expression for $E_G(R)$ considering that b, G can be scale dependent, and including Σ, is

$$E_G = \frac{\Omega_{m0}\int_0^\infty \frac{\mathrm{d}z}{H(z)}\hat{\xi}_{gm}(R, z_\ell, z) r(z_\ell) r(z)(1+z_\ell)(1+z)\int_z^\infty \frac{\mathrm{d}z'}{H(z')} W_s(z')\frac{[r(z')-r(z)][r(z')-r(z_\ell)]}{r(z')^2}}{2f(z_\ell)\int_{z_\ell-\Delta_\ell}^{z_\ell+\Delta_\ell}\frac{\mathrm{d}z}{H(z)}\hat{\xi}_{gg}(R, z_\ell, z)}, \tag{4.13}$$

or, rewriting as $\alpha(R) \equiv \frac{E_G}{E_{G0}}$, we seek to obtain $\alpha = 1$ in our calculations through the formula

$$\alpha(R) = \frac{\int_0^\infty \frac{\mathrm{d}z}{H(z)}\hat{\xi}_{gm}(R, z_\ell, z) r(z_\ell) r(z)(1+z_\ell)(1+z)\int_z^\infty \frac{\mathrm{d}z'}{H(z')} W_s(z')\frac{[r(z')-r(z)][r(z')-r(z_\ell)]}{r(z')^2}}{\Sigma(z_\ell)\int_{z_\ell-\Delta_\ell}^{z_\ell+\Delta_\ell}\frac{\mathrm{d}z}{H(z)}\hat{\xi}_{gg}(R, z_\ell, z)}, \tag{4.14}$$

where

$$\hat{\xi}_{gm}(R, z_\ell, z) = \frac{1}{2\pi^2}\int_0^\infty \mathrm{d}k G(z,k)G(z_\ell,k)b(z,k)\Sigma(z,k)k^2 e^{-k^2}P(k)\hat{J}_0(k, R, z, z_\ell), \tag{4.15}$$

$$\hat{\xi}_{gg}(R, z_\ell, z) = \frac{1}{2\pi^2}\int_0^\infty \mathrm{d}k G(z,k)G(z_\ell,k)b(z,k)b(z_\ell,k)k^2 e^{-k^2}P(k)\hat{J}_0(k, R, z, z_\ell). \tag{4.16}$$

These functions can be approximated in order to become computationally more stable and efficient. We use

$$\hat{\xi}_{gm}(R, z_\ell, z) = b_0(z_\ell) f(R, \Delta, 0) + b_1(z_\ell) f(R, \Delta, n), \tag{4.17}$$

$$\hat{\xi}_{gg}(R, z_\ell, z) = b_0(z)b_0(z_\ell) f(R, \Delta, 0) + b_0(z)b_1(z_\ell) f(R, \Delta, n) + b_0(z_\ell)b_1(z) f(R, \Delta, n) + b_1(z)b_1(z_\ell) f(R, \Delta, 2n), \tag{4.18}$$

where

$$f(R,\Delta,q)=\frac{1}{2\pi^2}\int_0^\infty \mathrm{d}k k^{2+q}e^{-k^2}P(k)\hat{J}_0(k,R,\Delta)\,, \tag{4.19}$$

is evaluated for $q=0,n,2n$.

4.3. Computational details and results

The goal is to find the parameter region where $\alpha(R)\approx 1$. Our procedure in order to compare E_{G0} with Eq. (4.2) and Eq. (4.13) is the following:

- Use the cosmological functions given by ΛCDM: growth function $G(z)$, Hubble function $H(z)$, distance $r(z)=\int\frac{\mathrm{d}z}{H(z)}$, assuming $H_0=1$. Here we assume an universe with matter and dark energy, with $\Omega_m=0.3$ and a common approximation for the growth function $f(z)=\Omega_m(z)^\gamma$ where $\gamma=0.55$ for ΛCDM;

- Take the source distribution function P_s of Eq. (4.5) and chose the bias function $b(z,k)$ and lensing function $\Sigma(z)$. Alternatively, a simulated power spectrum from a Boltzman code can be a good choice for $P_s(z)$;

- Evaluate the correlation function that takes into account the growth function, bias function and kernel function $\hat{\xi}_{gm}$ and $\hat{\xi}_{gg}$ using Eq. (4.17) and Eq. (4.18);

- Compute Eq. (4.14), which should be around 1.

We identify the relevant parameter space given by

$$R_0, z_\ell, \Delta_\ell, \tag{4.20}$$

plus any parameter that enters in $P(k)$ if not those defined in Eq. (4.5) (see [146] which compare 3 different source distribution functions). Also, it depends on the cosmological parameters through $H(z)$ and the distance definition as well as in the growth function $G(z)$ if an approximation is used or not. Finally, the definition of $\Sigma(z), b(z,k)$ may introduce extra parameters too. In order to directly compare with Fig. 4.1 from Leonard *et al.* [146], we chose $z_\ell=0.5$, $R_0=1.5$ Mpc$/h$ and Δ_ℓ such that $r=\int_{z_\ell}^{z_\ell+\Delta_\ell}\frac{\mathrm{d}z}{H}=500$ Mpc$/h$, or simply $\Delta_\ell\approx H(z_\ell)500$. We also chose to have a z, k-dependent bias function such as

$$b(z,k)=b_0(z)+b_1(z)k^n \tag{4.21}$$

with $n=1.28$ and b_0, b_1 values from [177].

The results of this computation are currently work in progress. Nevertheless, the last results are depicted in Fig. 4.2 with the $\alpha(R)$ function from Eq. (4.14). There is an overall factor missing, moving the results for about a factor of 2. It is important to notice the growing effect at smaller distances which overlaps with the yellow area that represents the distances to which E_G was measured as in Tab. 3.6.

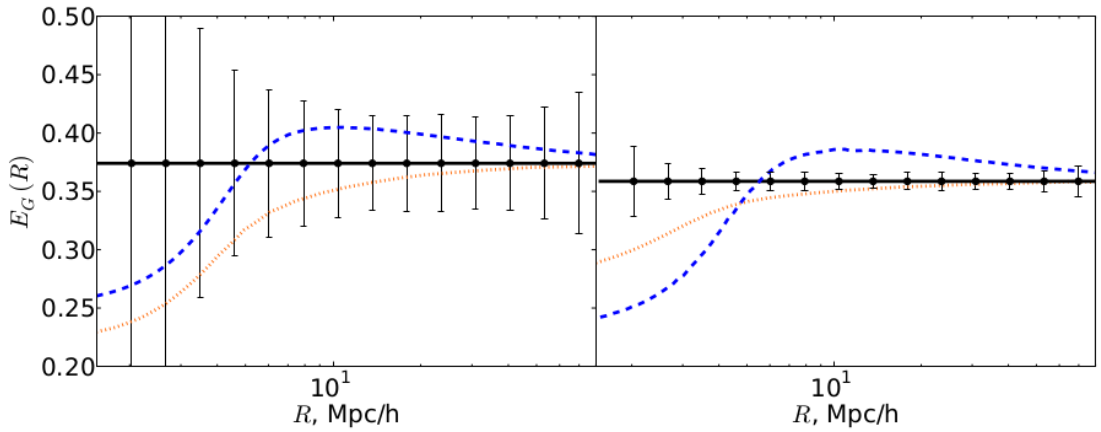

Figure 4.1.: Graphics from Leonard *et al.* [146] that shows E_G(R) with a scale-dependent bias function. The left side panel and right side panel correspond to different specifications of next-generation instruments. For the left side panel, it corresponds to the DESI instrument for the galaxy-galaxy and lens-galaxy cross-correlations and source galaxies from a DETF4 type survey, with $z_\ell = 0.8$. On the right side panel, these cross-correlations come from stage 2 of SKA and source galaxies from LSST, with $z_\ell = 1.0$.

Several numerical technicalities can interfere with the final result. For instance, the integral in Eq. (4.9) is a highly oscillating function. For this reason, we redefined the expression such that the integral can be performed analytically. We use the last line of Eq. (4.11). After this upgrade and given that we would like to take into account possible scale dependence, we further need the approximation done with Eq. (4.17) and Eq. (4.18). Finally, the window function at Eq. (4.6) is also rewritten in a way it is faster to integrate and compute. All of these approximations, in particular, each integral involved, can also have further optimization options that led to very different final results in the code written in Mathematica.

In the literature, the sources of the E_G measurements [138, 147, 151] chose different values of R_0, that is, $2.0, 1.0, 1.5\,h^{-1}$Mpc respectively. Each of these surveys checked the dependence of their results on the choice of this cut-off scale R_0 between $1.0 < R_0 < 3.0\,h^{-1}$Mpc but did not find any significant difference within this range of values. Their analysis is tomographic in real space where the narrow lens approach is assumed. Thus, the reported redshift is the average of the redshift bin the measurement is done and corresponds to z_ℓ. As previously noted, these considerations may have an impact on the final estimate of E_G.

4.4. Summary

In this chapter we explored the motivation and the definitions of the E_G statistics. This estimator gathers the sensitivity of the weak lensing and galaxy clustering with the aim of distinguish models of modified gravity. We described both the theoretical and the observational definitions of this estimator. It depends on a redshift distribuition of sources $P_s(z)$, which can be taken from a more realistic measurement scenario such as SKA or

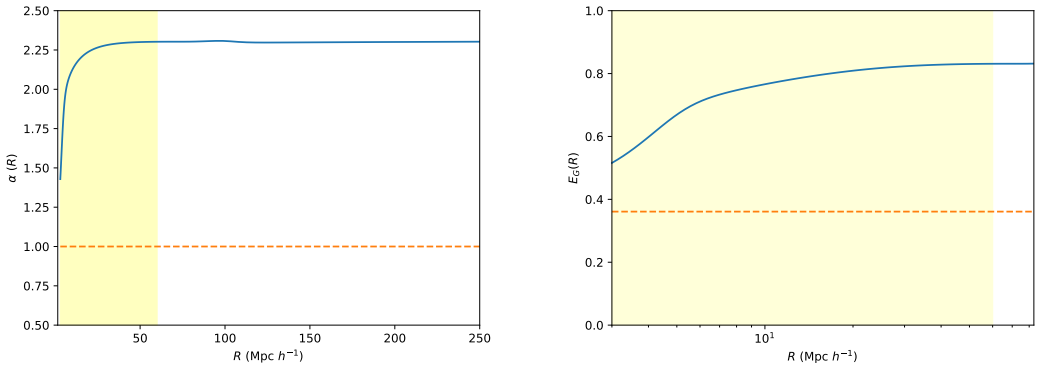

Figure 4.2.: Left side: Preliminary results of Eq. (4.14) as a function of the distance R, where we set $z_\ell = 0.8$ and $\delta_\ell = 500\text{Mpc}/h$, represented by the solid line. The expected value for this quantity is one, as the dashed line represents. The yellow region corresponds to the region of the measurements compiled in Tab. 3.6. Right side: E_G(R) as given by Eq. (4.13) as the solid line and the expected ΛCDM value from Eq. (4.2) as the dashed-line.

DESI instruments. The cutoff-scale R_0 is the scale below which information is discarted and the projection length δ_ℓ at the galaxy-galaxy cross-correlation. It also depends on the bias function $b(z, k)$ which is usually taken as linear but in reality its behaviour is unknown.

This estimator has a simple expression in a ΛCDM scenario but it is not straightforward to obtain it observationally. The geometry of the sources and lensed galaxies need to be carefully accounted as well as an extension to alternative theories of gravity. We identified the relevant parameter space and proceeded to implement these equations numerically, although it is still work in progress. For now, we can recover the increasing slope for lower distances but there is an overall factor of 2 missing, in case of comparison with the work by Leonard et al. [146]. These would be the details missing also if this expression holds for all distances as it is meant for but as it is our aim to test.

We would like to point that a correct calculation of this estimator with a comparison for different theories could indeed help assess properties of gravity. It should not be labeled as model-independent as, for the case of ΛCDM, the respective expression does not involve direct observables and the expected values vary for distinct theories. As previously mentioned in Chapter 3, the current available measurements prefer a lower value of matter density than the one obtained by the Planck collaboration. The discrepancy seen for instance in the KiDS data [138] can be another perspective on the $f\sigma_8$ tension as both quantities rely on the same datasets but also hint of new physics. We conclude it might help understanding gravity when future missions such as Euclid [178] measure it with high accuracy.

5. Measuring the flow of cosmological information

Alongside with the idea of testing gravity in a model-independent way is the question of how observations are interpreted. Cosmology mainly relies on Bayesian statistics in data analysis since most of the data requires a lot of time and expensive experimental setups. The Bayesian statistics allows easy computations of probability distributions using previous information, and also model comparison tools which are very important to narrow down the number of possible models. A possible alternative approach is to use information theory to assess the amount of information available from probabity distributions either from forecasts or from datasets. It is possible to follow a Bayesian approach within this framework and to compare different uncertainty measurements is the aim of this chapter. Furthermore, current cosmological data such as the CMB, supernovae type Ia, weak lensing, and galaxy clustering data enable the study of the properties of gravity with high precision, even with the growing complexity of systematical effects and data analysis. As a consequence from the particular parameter sensitivities of cosmological probes, the combination of these datasets has become essential to narrow down the allowed cosmological parameter space. We explore the combination of these cosmological probes taking into account all possible cross-correlations, which at the moment is only possible in a forecast setup. If the probes are consistent, any deviation or discrepancy between the best-fit parameters is a hint for hidden systematics or possible new physics. Thus, it is important to understand not just the possible assumptions done in statistical inference but also to find a good way to measure uncertainties and relate them with parameter degeneracies or the significance of tension in the data.

In a statistical inference problem, the information about how good does a model represent the data is encapsulated in the likelihood function. This probability distribution has typically a Gaussian shape for relatively simple models which is a fair assumption when well-constrained by data [179, 180]. Nonlinearities in the model can be approximated by linear relationships leading to an ideal scenario of a Gaussian distribution and to the possibility of using the Fisher-matrix formalism [181–184], which is ubiquitous in modern cosmology [185–188]. This Gaussian assumption yields a way to estimate the expected parameter covariance from the second derivatives of the logarithmic likelihood. The Fisher-matrix formalism is a crucial tool for forecasting applications assuming that the true model is known and it is usually referred to as the *fiducial* cosmology. Inference from cosmological data is often not limited by statistics but rather by systematics. For this reason, extensions to the Fisher formalism have been proposed that allow the forecasting of systematical errors, i.e. the shift of the best-fit point of a Gaussian likelihood if an unknown systematic is not removed or properly modelled [189–194].

There is no obvious way to quantify the total statistical error of a cosmological probe or the significance of tensions between likelihood obtained with different cosmological probes, even in the case of Gaussian likelihoods. Such a measure of total error would be convenient in quantifying the information content of a particular cosmological probe, or

its parameter degeneracy breaking power, or in applications of experimental design, where one optimises a survey to yield the smallest possible errors [195, 196]. In a larger context, Bayesian evidences [197–199] used in model selection are measures of the consistency between likelihood and prior [200], and are only on possible choice among many others, for instance the Akaike information or the Bayes-information, for preferring a particular model, incorporating a tradeoff between the goodness-of-fit and model complexity [201–204].

In this chapter, we apply the information theory to cosmology, in a similar way to previous work [205, 206]. We compute the absolute entropy through the likelihoods of cosmological probes and relative entropy for the case of different probe combination. These are compared with traditional uncertaity measurements, which are typically simple properties of the Fisher matrix with respect to a specific probe and model. This comparison intends to provide an interpretation for the cosmological likelihoods, suggesting how the information theory can be useful. A possible useful aplication is writting the biases between likelihoods as relative entropies, which can later be used as a sort of scale. This is the case for the dark energy figure of merit that was designed to evaluate the performance of cosmological probes to measure deviations from the ΛCDM model, which is actually a measure of information entropy. We focus on Gaussian distributions where all calculations have analytical solutions. The concept of information entropy is generally applicable to asymmetric or even multimodal distributions. This concept has the advantadge of providing natural generalisations to quantities that are intuitive for Gaussian distributions. This applies in particular to Bayesian evidences or evidence ratios, which can in fact be related to information entropy differences. In the case of information entropy differences, these are interpretable directly without resorting to the rather arbitrarily defined Jeffreys scale and are likewise quantified in units of nats.

Unless otherwise specified, we focus on spatially flat ΛCDM and wCDM cosmologies with adiabatic initial conditions. The fiducial parameter choices are the values of the [83] (TT, TE, EE, low$z$ + lensing), meaning $\Omega_m = 0.3153$, $\sigma_8 = 0.8111$, $h = 0.6736$, $n_s = 0.9649$ and $\Omega_b = 0.0493$. The recombination redshift is chosen as $z_{re} = 11.357$ and galaxy bias as $b = 0.68$ [207]. The dark energy fluid is described by an equation of state parameter with a linear variation in time, commonly known as the CPL parameterisation [208, 209],

$$w(a) = w_0 + (1 - a)w_a, \tag{5.1}$$

where $w_0 = -1$ and $w_a = 0$ recovers the case of a cosmological constant Λ. In this analysis, we assume the characterics of the Euclid survey [210] for the large-scale structure and of Planck for the CMB and CMB-lensing. Sec. 5.1 describes the fundamental concepts of Bayesian statistics, including the Fisher matrix formalism, and how these are related to the essential notions of information theory. Then, we compare absolute entropies to properties of the Fisher matrix in Sec. 5.3 and how probe combination increases the relative entropies in Sec. 5.4. In Sec. 5.2 we summarize the important details of the large scale structure probes that are considered in our analysis. Furthermore, we delve into the famous data tensions: on the H_0 value, the S_8 value and the importance of intrinsic alignments in weak lensing data in a wCDM scenario in Sec. 5.5.

5.1. Relation between statistics and information theory

The likelihood $\mathcal{L}(x_\mu|M,D)$ of a model M for the data D is embedded into Bayes' theorem. This expresses the state of knowledge after carrying out an experiment, the posterior $p(x_\mu|M,D)$, as proportional to the likelihood provided by the experiment times the prior distribution $\pi(x_\mu|M)$,

$$p(x_\mu|M,D) = \frac{\mathcal{L}(D|x_\mu,M)\ \pi(x_\mu|M)}{p(D|M)}, \tag{5.2}$$

where $p(D|M)$ is the evidence defined by

$$p(D|M) = \int \mathrm{d}^n x\ \mathcal{L}(D|x_\mu,M)\ \pi(x_\mu|M). \tag{5.3}$$

The assumption of Gaussian probability distributions is very common in cosmology and it can be easily expressed in terms of the covariance matrix. The model parameters are well constrained when the likelihoods are peaked and the full parameter covariances are small. This study of the statistical errors of cosmological parameters frequently serves to check if future data can help validating new models for dark energy or structure formation. When it comes to distinguish between models, the Occam's razor is the guiding principle at it prefers the simpler models with fewer parameters. For this purpose, the evidence can be useful to quantify the complexity of the model, namely, via the Bayes ratio. The Bayes factor was designed to select the most likely model, independently of the choice or number of parameters. Motivated by the Neyman-Pearson lemma, one compares competing models by constructing their evidence ratio as one would do when comparing likelihoods,

$$B_{ij} = \frac{p(D|M_i)}{p(D|M_j)} \tag{5.4}$$

such that positive values of B_{ij} prefer the model M_i over M_j and vice versa. The Jeffrey's scale is used to assess the strength of evidence for preferring one model over another [211]. Further criteria have been developed and used such as the Akaike, or Bayesian information criteria which additionally account for the number of parameters in the model, each in a different way [201–204].

Information theory has entropy as its main quantity, inspired by the thermodynamics definition. It quantifies the amount of randomness in a distribution. The Shannon entropy S [212] is defined as

$$S = -\int \mathrm{d}^n x\ p(x_\mu) \ln p(x_\mu), \tag{5.5}$$

where $p(x_\mu)$ can be, for instance, the likelihood function. In that case, the Shannon entropy takes small values for a peaked likelihood, as it follows from $S = \ln[2\pi\sigma^2 \exp(1)]/2$ for a Gaussian distribution. The most general definition of entropy is the Rényi entropy S_α [213, 214], parameterised by $\alpha > 0$ and $\alpha \neq 1$,

$$S_\alpha = -\frac{1}{\alpha - 1} \ln \int \mathrm{d}^n x\ p(x_\mu) p^{\alpha-1}(x_\mu), \tag{5.6}$$

where the Shannon-entropy is recovered in the limit $\alpha \to 1$ by application of de l'Hôpital's rule. The Rényi-entropies also increase with the variance for positive values of α as

$S_\alpha = \ln(2\pi\sigma^2\alpha^{\frac{1}{\alpha-1}})/2$. This implies that entropies provide a way of quantifying how constraining data is as they quantify the size of the allowed parameter space.

Interestingly, the Gaussian distribution is very peculiar within the Shannon entropy. The reason is that the Gaussian distribution maximises the Shannon entropy, among all distributions with a fixed variance. This statement can be illustrated using a Gram-Charlier-parameterised distribution $p(x)\mathrm{d}x$ with weak non-Gaussianities [215] described by the cumulants κ_3 and κ_4, both of which are much smaller than one,

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \left[1 + \frac{\kappa_3}{3!\sigma^3} H_3\left(\frac{x}{\sigma}\right) + \frac{\kappa_4}{4!\sigma^4} H_4\left(\frac{x}{\sigma}\right)\right], \tag{5.7}$$

with the Hermite-polynomials $H_n(x)$ of order n. Substituting this series into the definition in Eq. (5.5) and approximating $\ln(1+\epsilon) \simeq \epsilon$ for $|\epsilon| \ll 1$, one obtains at second order the result

$$S = \frac{1}{2}\ln\left[2\pi\sigma^2\exp(1)\right] - \frac{1}{3!}\frac{\kappa_3^2}{\sigma^6} - \frac{1}{4!}\frac{\kappa_4^2}{\sigma^8}, \tag{5.8}$$

by using the orthogonality relation of the Hermite-polynomials,

$$\int \mathrm{d}x \, \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) H_m\left(\frac{x}{\sigma}\right) H_n\left(\frac{x}{\sigma}\right) = n!\delta_{mn}. \tag{5.9}$$

Eq. (5.8) shows that the entropy of the Gaussian distribution is always diminished by non-Gaussianities, because κ_3^2 and κ_4^2 are as squares necessarily positive. Because of this result, we would like to point out that the information entropies that we compute are upper bounds, and that realistic non-Gaussian likelihoods would have lower values for their information entropies than their Gaussian counterparts. It is remarkable that the orthogonality relation in Eq. (5.9) cancels the influence of non-Gaussianities on S to first order, which can be shown by substituting $1 = H_0(x)$ and $x^2 = H_2(x) + H_0(x)$. Sadly, there is no analogous result to Eq. (5.9) for the Rényi-entropy, but Eq. (5.9) can be generalised in principle to hold for non-Gaussianities of arbitrary order κ_n. A multivariate generalisation of Eq. (5.8) can serve as a way to estimate the Shannon-entropy from MCMC-samples of the likelihood without the need of a density estimate as a way to compute $\ln p(x)$. Instead, one would estimate multivariate cumulants from the samples directly and correct the Gaussian result for the information entropy. It can be expected that a similar relationship exists for DALI-approximated likelihoods [216, 217].

The basis of the Fisher matrix formalism is to approximate likelihoods with a multivariate Gaussian distribution. This allows the correspondent covariance matrix to be computed from the averaged gradients of the logarithmic likelihood. Fixing the fiducial model, the Fisher-matrix $F_{\mu\nu}$ [185] is defined as

$$F_{\mu\nu} = -\left\langle \frac{\partial^2 \ln\mathcal{L}}{\partial x_\mu \partial x_\nu} \right\rangle, \tag{5.10}$$

yielding for a measurement of multipole moments $A_{\ell m}$ and $B_{\ell m}$ of Gaussian random fields $A(\theta,\varphi)$ and $B(\theta,\varphi)$ that are described by angular spectra $C_{AA}(\ell)$, $C_{BB}(\ell)$ and $C_{AB}(\ell)$ as

$$F_{\mu\nu} = \sum_\ell \frac{2\ell+1}{2} \mathrm{tr}\left(\frac{\partial}{\partial x_\mu}\ln C \, \frac{\partial}{\partial x_\nu}\ln C\right), \tag{5.11}$$

where the spectra are combined into a common data covariance C.

Quoting the logarithmic curvature $F_{\mu\nu}$ of the likelihood surface in parameter space, the second moments $C_{\mu\nu} = \langle x_\mu x_\nu \rangle$ or confidence intervals is equivalent for a Gaussian distribution. Since the Fisher matrix can correspond to the inverse parameter covariance $C = F^{-1}$, one can write down a multivariate Gaussian distribution as

$$p(x_\mu) = \sqrt{\frac{\det(F)}{(2\pi)^n}} \exp\left(-\frac{1}{2} x_\mu F_{\mu\nu} x_\nu\right), \tag{5.12}$$

using coordinates relative to the best-fit point. Measures of total uncertainty can be derived from the Fisher-matrix in a straightforward way as, for instance, the invariant trace $\mathrm{tr}(F)$, the Frobenius-norm $\mathrm{tr}(F^2)$ or the determinant $\det(F)$ of which we will take the logarithm $\ln\det(F)$ to make the connection to information entropies clearer. Generalisations to the trace and the Forbenius norm of the type $\mathrm{tr}(F^p)$ with $p > 2$ would be restricted in the values that they can assume by the Hölder inequality,

$$\frac{1}{n}\,\mathrm{tr}\,(F) \leq \left[\frac{1}{n}\mathrm{tr}(F^p)\right]^{\frac{1}{p}} \tag{5.13}$$

for arbitrary powers p of Fisher matrices in n dimensions, where the traces can be generalised to arbitrary real valued powers p by using $\mathrm{tr}(F^p) = \mathrm{tr}\exp[p\ln(F)]$.

On the other side, the generalised inequality of the arithmetic and geometric mean implies

$$\frac{1}{n}\mathrm{tr}(F) \geq \det(F)^{\frac{1}{n}}, \tag{5.14}$$

such that the information entropies are bounded by traces of the Fisher matrix, as shown in the next paragraph. Specifically, while $\mathrm{tr}(F) = \sum_\mu \sigma_\mu^{-2}$ is a measure of the total uncertainty of the likelihood, it does not differentiate between correlated and uncorrelated distributions. This is taken care of by $\mathrm{tr}(F^2)$, as the expression contains information from the off-diagonal elements in addition to performing a different weighting of the errors. While all trace relations for arbitrary p are measures of total error, only the determinant provides a geometric interpretation as the volume of parameter space. The dark energy figure of merit is defined as the volume of the w_0-w_a subspace of parameter space bounded by the 1σ-contour. Of all these measures, however, only $\mathrm{tr}(F)$ is additive for statistically independent measurements. Given the inequalities in Eq. (5.13) and Eq. (5.14), we state all results in a scaled way, i.e. $[\mathrm{tr}(F^p)/n]^{1/p}$ and $\det(F)^{1/n}$. The usage of these scaled traces is motivated by the fact that for a diagonal Fisher-matrix with identical entries $1/\sigma^2$, they all return the same value of $1/\sigma^2$ irrespective of n or p.

Analytical expressions for the entropies S and S_α can be derived for a multivariate Gaussian in terms of the determinant of the covariance matrix $C_{\mu\nu}$ as the inverse Fisher-matrix. Specifically, integration by substitution yields directly

$$S = \frac{1}{2}\ln\Big[(2\pi)^n \det(C)\,\exp(n)\Big], \tag{5.15}$$

for the Shannon-entropy, and

$$S_\alpha = \frac{1}{2}\ln\Big[(2\pi)^n \det(C)\,\alpha^{\frac{n}{\alpha-1}}\Big], \tag{5.16}$$

for the Rényi-entropy, such that the univariate case is recovered for $n = 1$ and $\det(C) = \sigma^2$. The two definitions are consistent as in the limit $\alpha \to 1$, the expression $\alpha^{\frac{n}{\alpha-1}}$ converges to

$\exp(n)$. The difference between Shannon and Rényi entropies for Gaussian distributions with identical covariances is given by an additive term,

$$\Delta_n(\alpha) = S_\alpha - S = \frac{1}{2}\left[\ln\left(\alpha^{\frac{n}{\alpha-1}}\right) - n\right]. \tag{5.17}$$

Considering in particular the case of Bhattacharyya entropies [218] defined as S_α with $\alpha = 1/2$, this figure shows that since $\alpha = 1/2 < 1$, it will always be larger than Shannon entropies. This trend becomes stronger with increasing number of random variables n. Typical numbers in cosmology with a ΛCDM or wCDM model with 7 or 8 parameters would then be $\Delta_7(1/2) \simeq 1.35$ and $\Delta_8(1/2) \simeq 1.54$.

Remarkably enough, both entropies are measures of the logarithmic volume of the parameter space bounded by the 1σ contour, implying that the dark energy figure of merit is in fact an inverse information entropy. Interestingly, the entropies are defined as $\det(C)$, which is always strictly positive for the covariance matrix $C_{\mu\nu} = \langle x_\mu x_\nu \rangle$. This comes as a consequence of Gram's inequality, while $\det(x_\mu x_\nu)$ without averaging would be exactly zero. By the choice of the natural logarithm, the unit of entropy is nat.

The information entropies S and S_α are inversely proportional to $\ln\det(F)$ and thus, one should expect similar relations with the measures $\mathrm{tr}(F) = F_{\mu\mu}$ and $\mathrm{tr}(F^2) = F_{\mu\nu}F_{\mu\nu}$ too. As explained before, the Hölder-inequality on Eq. (5.13) and the inequality of the geometric and arithmetic mean on Eq. (5.14) provide bounds on the information entropy S for both the Shannon and Rényi definition in terms of trace invariants of the Fisher matrix. Additivity in the case of statistical independence is a defining property of information entropies that makes them useful for describing the information content. They share this property with Fisher matrices for the case of statistically independent probes, i.e. $F_{\mu\nu} = F^{(1)}_{\mu\nu} + F^{(2)}_{\mu\nu}$ implies $S_\alpha = S^{(1)}_\alpha + S^{(2)}_\alpha$ as a consequence of $p(x_\mu) = p^{(1)}(x_\mu)\, p^{(2)}(x_\mu)$. For the case with statistical non-independence, additivity of the entropies does not hold and therefore, one defines relative entropies between two distributions, also referred to as divergences. For the Shannon entropy, this corresponds to the Kullback-Leibler divergence ΔS [219],

$$\Delta S = D_{\mathrm{KL}} = \int \mathrm{d}^n x\, p(x_\mu) \ln \frac{p(x_\mu)}{q(x_\mu)}, \tag{5.18}$$

and a more general class of α-divergences ΔS_α for Rényi entropies,

$$\Delta S_\alpha = \frac{1}{\alpha - 1} \ln \int \mathrm{d}^n x\, p(x_\mu) \left(\frac{p(x_\mu)}{q(x_\mu)}\right)^{\alpha-1}, \tag{5.19}$$

between two multivariate distributions $p(x_\mu)$ and $q(x_\mu)$.

Relative entropies would be invariant under transformation of the random variables, whereas absolute entropies would not. In fact, they do depend on the choice of parameterisation and even on the choice of units for the parameters, which in particular is less relevant in cosmology as almost all parameters are defined in a dimensionless way, with H_0 or $\chi_H = c/H_0$ being notable exceptions. Under an invertible reparameterisation with a nonzero Jacobian determinant $\det(\partial y_\nu/\partial x_\mu)$, both the Shannon entropy S and, surprisingly, the Rényi entropy S_α too acquire the identical additive term $\ln\det(\partial y_\nu/\partial x_\mu)$. This is valid if the transformation is affine, $y_\nu = A^\mu{}_\nu x_\mu + b_\nu$ with a constant $A^\mu{}_\nu$ and b_ν corresponding to a change in units and a shift of the mean.

In this application, we would like to compute the entropy difference between the posterior $p(x_\mu) = p(x_\mu|M, D) \propto \mathcal{L}(D|x_\mu, M)\pi(x_\mu|M)$ which includes the information provided

by measurement and the prior $q(x_\mu) = \pi(x_\mu|M)$, which reflects the state of knowledge before the data has been taken. We would like to point out that the entropy divergences ΔS and ΔS_α are not symmetric in interchanging prior and posterior. Also, the definition of relative entropy does not admit transitivity when combining multiple independent data sets $D_1, \ldots, D_n$, i.e. in cases where $p(x_\mu|M, D) = \mathcal{L}(D_1|x_\mu, M) \cdots \mathcal{L}(D_n|x_\mu, M)\pi(x_\mu|M)$. As an example, if $\Delta S(1)$ is the entropy divergence between the posterior $\mathcal{L}(D_1|x_\mu, M)\,\pi(x_\mu|M) = \mathcal{L}_1\pi$ and the prior $\pi(x_\mu|M)$,

$$\Delta S(1) = \int \mathrm{d}^n x\, \mathcal{L}_1\pi \ln \frac{\mathcal{L}_1\pi}{\pi} = \int \mathrm{d}^n x\, \mathcal{L}_1\pi \ln \mathcal{L}_1 \tag{5.20}$$

and $S(2)$ the corresponding difference between $\mathcal{L}_1\mathcal{L}_2\pi$ and π,

$$\Delta S(2) = \int \mathrm{d}^n x\, \mathcal{L}_1\mathcal{L}_2\pi \ln \frac{\mathcal{L}_1\mathcal{L}_2\pi}{\pi} = \int \mathrm{d}^n x\, \mathcal{L}_1\mathcal{L}_2\pi \ln \mathcal{L}_1\mathcal{L}_2 \tag{5.21}$$

one can define the entropy decrease $\Delta S(12)$ gained by including the data set D_2 and adding the likelihood $\mathcal{L}_2$ to the state of knowledge $\mathcal{L}_1\pi$ obtained from the data set D_1,

$$\Delta S(12) = \int \mathrm{d}^n x\, \mathcal{L}_1\mathcal{L}_2\pi \ln \frac{\mathcal{L}_1\mathcal{L}_2\pi}{\mathcal{L}_1\pi} = \int \mathrm{d}^n x\, \mathcal{L}_1\mathcal{L}_2\pi \ln \mathcal{L}_2. \tag{5.22}$$

With these definitions, one sees that $\Delta S(2) \neq \Delta S(1) + \Delta S(12)$. Because of this and due to statistical non-independence of cosmological probes, we compute all entropies from the effective Fisher-matrix combining all probes into a single Gaussian likelihood. The same issue appears in the case of Rényi entropies ΔS_α in an identical way.

If now we inverse, i.e. setting $p(x_\mu) = \pi(x_\mu|M)$ and $q(x_\mu) \propto \mathcal{L}(D|x_\mu, M)\pi(x_\mu|M)$, i.e. computing the entropy divergence of the prior relative to the posterior yields an interesting result. The entropy divergence quantifies by how much the entropy will decrease by acquiring new data, i.e. by how much the entropy of the posterior will be different relative to that of the prior. At first sight, one might think that entropies are then additive for statistically independent likelihoods, $\mathcal{L} = \prod_i \mathcal{L}_i$,

$$\Delta S = \int \mathrm{d}^n x\, \pi \ln \frac{\pi}{\mathcal{L}\pi} = -\int \mathrm{d}^n x\, \pi \ln \mathcal{L} = -\sum_i \int \mathrm{d}^n x\, \pi \ln \mathcal{L}_i, \tag{5.23}$$

but the evidence-term $\int \mathrm{d}^n x\, \mathcal{L}(x_\mu)\pi(x_\mu)$ needed for a correctly normalised posterior in fact breaks additivity, as

$$\Delta S(1) = -\int \mathrm{d}^n x\, \pi \ln \mathcal{L}_1 + \ln \int \mathrm{d}^n x\, \mathcal{L}_1\pi \tag{5.24}$$

with the renormalised posterior,

$$q(x_\mu) = \frac{\mathcal{L}(x_\mu)\pi(x_\mu)}{\int \mathrm{d}^n x\, \mathcal{L}(x_\mu)\pi(x_\mu)} \tag{5.25}$$

is not contained in the expression

$$\Delta S(2) = -\int \mathrm{d}^n x\, \pi \ln \mathcal{L}_1 - \int \mathrm{d}^n x\, \pi \ln \mathcal{L}_2 + \ln \int \mathrm{d}^n x\, \mathcal{L}_1\mathcal{L}_2\pi. \tag{5.26}$$

If there are no tensions between the likelihoods and if they are of Gaussian shape, there are analytic relations for the relative Shannon entropy ΔS,

$$\Delta S = \frac{1}{2}\left[\ln\frac{\det(F)}{\det(G)} - n + F^{-1}_{\mu\nu}G_{\mu\nu}\right] \tag{5.27}$$

now expressed in terms of the Fisher matrices $F_{\mu\nu}$ and $G_{\mu\nu}$ of the posterior and the prior, respectively, as well as for the relative Rényi entropy ΔS_α,

$$\Delta S_\alpha = \frac{1}{2}\frac{1}{\alpha-1}\ln\left[\frac{\det^\alpha(F)}{\det^{\alpha-1}(G)\det(A)}\right]. \tag{5.28}$$

Both relationships yield $\Delta S = \Delta S_\alpha = 0$ if $F_{\mu\nu} = G_{\mu\nu}$. It is quite illustrative to substitute $\ln\det(F) = \ln\mathrm{tr}(F)$, yielding

$$\Delta S = \frac{1}{2}\left[\left(\ln(F)_{\mu\mu} - F^{-1}_{\mu\nu}F_{\mu\nu}\right) - \left(\ln(G)_{\mu\mu} - F^{-1}_{\mu\nu}G_{\mu\nu}\right)\right] \tag{5.29}$$

such that ΔS becomes $\langle\Delta\chi^2\rangle/2$ for Gaussian likelihoods as $\mathcal{L} \propto \exp(-\chi^2/2)$, where we substituted $F^{-1}_{\mu\nu}F_{\mu\nu} = n$ for symmetry. The analogous relation for the relative Rényi entropy ΔS_α is

$$\Delta S_\alpha = \frac{1}{2}\frac{1}{\alpha-1}\left[\alpha\ln(F)_{\mu\mu} + (1-\alpha)\ln(G)_{\mu\mu} - \ln(A)_{\mu\mu}\right] \tag{5.30}$$

with

$$A_{\mu\nu} = \alpha F_{\mu\nu} + (1-\alpha)G_{\mu\nu}, \tag{5.31}$$

where one recovers the convexity condition for the matrix-valued logarithm. Again, application of de l'Hôpital's rule for evaluating the limit $\alpha \to 1$ recovers ΔS from ΔS_α. It is not straightforward to find general interpretations of Eq. (5.30) for arbitrary α. In the Shannon case, one finds for S the ratio between the logarithmic volumes of the two likelihoods and the asymmetry of the relative entropy is ensured by the fact that $F^{-1}_{\mu\nu}G_{\mu\nu} \neq F_{\mu\nu}G^{-1}_{\mu\nu}$. One would find a symmetric expression for the Rényi-entropy ΔS_α if $\alpha = 1/2$, namely, the Bhattacharyya-entropy, assuming in this particular case,

$$\Delta S_\alpha = 2\ln\int \mathrm{d}^n x\,\sqrt{p(x_\mu)q(x_\mu)}\,. \tag{5.32}$$

This is then symmetric, with equal prefactors for $F_{\mu\nu}$, $G_{\mu\nu}$ and $A_{\mu\nu} = (F_{\mu\nu} + G_{\mu\nu})/2$. We will come back to this in Sec. 5.6, when discussing the relationship between Bayesian evidence and information entropy.

5.2. Large scale structure probes

As discussed previously, we will assume the data to be given as a collection of spherical harmonic modes. Under the assumption of Gaussian fields, their power spectra entirely determine the statistical properties. In this section, we will briefly describe the probes considered already presented in Sec. 2.2 and how the corresponding spectra are evaluated. we refer to [220] for further details, where is demonstrated the construction of Fisher matrices $F_{\mu\nu}$ from the cosmological probes including all non-vanishing cross-correlations that

would arise [221–223]. This section particularly serves to emphasize that cross-correlations have a dual influence on the inference process by making the data statistically dependent which would decrease the constraining power. On the contrary, they introduce unique handles on investigating structure formation, for instance, through the sensitivity of the integrate Sachs-Wolfe effect the CMB-LSS-correlations to dark energy. We approximate the covariance through a Gaussian with additional power on small scales due to the modelling of nonlinear structure formation [224–226]. Due to the assumption of a true fiducial model, we do not need to worry about covariance matrix variations [227–229]. Further, by the Gaussian assumption, there are no complications arising in relation to covariance matrix estimation [230–233].

5.2.1. Cosmic Microwave Background

The spectra of cleaned, full-sky CMB maps are given by a

$$\langle a^{P*}_{\ell m} a^{P'}_{\ell' m'} \rangle \equiv \hat{C}^{PP'}(\ell) = \Big(C^{PP'}(\ell) + N^{P}(\ell) \Big) \, \delta_{\ell\ell'} \delta_{mm'} \, , \tag{5.33}$$

where $P = T, E, B$ stands for temperature or the two polarization modes respectively, while $C^{TB}(\ell) = C^{EB}(\ell) = 0$. The noise covariance is given by [234]

$$N^{P}(\ell) \equiv \langle n^{P*}_{\ell m} n^{P'}_{\ell m} \rangle = \theta^2_{\rm beam} \sigma^2_P \exp\left(\ell(\ell+1) \frac{\theta^2_{\rm beam}}{8 \ln 2} \right) \delta_{PP'} \, . \tag{5.34}$$

with root mean square σ^2_P and a Gaussian beam with width $\theta_{\rm beam}$. Stage IV CMB experiments [235] will have a very small instrumental noise allowing for measurements up to $\ell \sim 5000$, especially for the polarisation maps. The spectra of the different components are calculated using the `hi-CLASS` code [49].

5.2.2. Large scale structure

The modes of any large scale structure probe can be calculated, to first order, as a weighted line-of-sight integral of the modes of the density field

$$A_{\ell m} = \int \mathrm{d}\chi \, W_A(\chi) \delta_{\ell m}(\chi) \, , \tag{5.35}$$

where χ is the comoving distance and a suitable weighting function $W_A(\chi)$. Corresponding spectra involve integration over Bessel functions due to the spherical basis. However, in the flat sky and Limber approximation, the calculation is simplified greatly and any angular power spectrum is given by [236]

$$C_{AB}(\ell) = \int \frac{\mathrm{d}\chi}{\chi^2} W_A(\chi) W_B(\chi) P_{\delta\delta} \left(\frac{\ell + 0.5}{\chi}, \chi \right) \, . \tag{5.36}$$

Note that the comoving wave vector of a mode k is related to the multipole ℓ via $k = (\ell + 0.5)/\chi$ in the Limber projection. We will continue by listing the weight functions of all probes used:

1. Cosmic shear [237, 238]:

$$W(\chi) = \frac{3\Omega_{\rm m} \chi^2_H}{2a\chi} \int_{\min(\chi,\chi_i)}^{\chi_{i+1}} \mathrm{d}\chi' p(\chi') \frac{\mathrm{d}z}{\mathrm{d}\chi'} \left(1 - \frac{\chi}{\chi'} \right) \, , \tag{5.37}$$

with the Hubble radius $\chi_H = c/H$, i the tomographic bin index and the Jacobi determinant $\mathrm{d}z/\mathrm{d}\chi' = H(\chi')/c$ due to the transformation of the redshift distribution $p(z)\mathrm{d}z$ of background galaxies in redshift z, which is given by [210]

$$p(z)\,\mathrm{d}z \propto z^2 \exp\left[-\left(\frac{z}{z_0}\right)^\beta\right] . \tag{5.38}$$

Typical parameters for stage IV experiments are $z_0 \approx 1$ and $\beta = 3/2$.

2. Galaxy clustering [239–241]

$$W(k,\chi) = \frac{H(\chi)}{c}\, b(k,\chi)\, p(\chi) \text{ if } \chi \in [\chi_i, \chi_{i+1}) , \tag{5.39}$$

where b is the galaxy bias [242] for which we assume [207]:

$$b(\chi) = b_0\,[1 + z(\chi)] , \tag{5.40}$$

with a free positive parameter b_0.

3. Lensing of the CMB [243, 244]:

$$W(\chi) = \frac{\chi_* - \chi}{\chi_*\chi}\frac{H(\chi)}{ca} , \tag{5.41}$$

with the comoving distance to the last scattering surface χ_*.

4. Integrated Sachs-Wolfe effect [245]:

$$W(k,a) = \frac{3}{2\chi_H^3} a^2 E(a)\, F'(k,a) , \tag{5.42}$$

where the prime denotes a derivative with respect to a and

$$F(k,a) = 2\frac{D_+(k,a)}{a} , \tag{5.43}$$

which is measured in cross-correlation with galaxy clustering and weak lensing.

The noise covariance of cosmic shear and galaxy clustering is given by

$$N_{\mathrm{LSS}}(\ell) = \sigma^2 \frac{n_{\mathrm{tomo}}}{\bar{n}_{\mathrm{gal}}}\, \delta_{\ell\ell'}, \tag{5.44}$$

with $\sigma = 0.3$ and $\sigma = 1$ for lensing and galaxy clustering, respectively, describing the intrinsic ellipticity of galaxies and the Poissonian fluctuation of galaxy numbers in each bin, $\bar{n}_{\mathrm{gal}}/n_{\mathrm{tomo}}$. Note that the tomographic bins are chosen such that the same amount of galaxies, i.e. $\bar{n}_{\mathrm{gal}}/n_{\mathrm{tomo}}$, lie in each bin. For CMB-lensing, we assume the noise to be given by the quadratic estimator described in Hu and Okamoto [246], Okamoto and Hu [247] using all five non-vanishing estimators involving T, E and B.

In our analysis, we combine the currently most powerful cosmological probes into a joint likelihood function. Specifically, we start out with spectra of the temperature and polarisation anisotropies in the CMB (labeled as CMB primary), and successively add CMB-lensing, tomographic galaxy clustering (GC) and tomographic weak gravitational

shear (WL), while taking account of all possible cross-correlations. As the reference cosmology, we use the Planck collaboration [83] result. The Fisher matrices used here were computed with the code of [248] where all cross-correlations are taken into account. At the moment, to apply this approach to available data would not be fully correct since it is not available the correlation between probes that have overlap areas of scan as it is the case for some surveys of galaxy clustering and weak lensing. Therefore, we use these specific Fisher matrices as a proof of concept, that is, to assess relations between the Bayesian statistics methods and information theory measures.

5.3. Comparison of uncertainty measures

Firstly, we would like to see how the statistical uncertainty in a ΛCDM or wCDM cosmology is reduced by combining cosmological probes. Starting from constraints from the temperature and polarisation anisotropies of the cosmic microwave background, we add successively gravitational lensing of the CMB, galaxy clustering and weak gravitational lensing, i.e. the large-scale structure probes ordered by decreasing redshift. In doing that, we are considering all nonzero cross-correlations in the data covariance, most notably the integrated Sachs-Wolfe effect between the CMB-temperature and any low-redshift tracer of the large-scale structure, plus the nonzero cross-correlations between the galaxy density and weak lensing.

At some point, the distribution will become very narrow such that their entropies will become negative, irrespective of the Shannon or Rényi definition. This can be explained straightforwardly by considering a one-dimensional Gaussian distribution with variance σ^2, where the relevant term in both entropy definitions is $\ln(\sigma^2)$ which tends towards $-\infty$ as $\sigma^2 \to 0$. The Dirac distribution δ_D with perfect knowledge of the parameters has infinite negative entropy and not zero as a consequence of the continuum limit.

We start by quantifying the absolute entropies S and S_α for the four cosmological data sets separately and put them into relation with other measures of total error that can be directly derived from the Fisher-matrix such as $\mathrm{tr}(F)$ and $\mathrm{tr}(F^2)$, where the inequality

$$\ln\det(F) = \mathrm{tr}\ln F \geq \ln\mathrm{tr}(F) \tag{5.45}$$

is obeyed as it should be expected from a positive definite and symmetric Fisher matrix $F_{\mu\nu}$. As such, the entropies are in fact not only scaling with Fisher invariants but are bounded by them as well, keeping in mind that $2S = n[\ln(2\pi) + 1] - \ln\det(F)$.

The absolute Shannon entropies S and Rényi entropies S_α for $\alpha = 1/2$ are listed in Tab. 5.1 for a ΛCDM cosmology and in Tab. 5.2 for a wCDM cosmology. It is clear for both cosmological models, that the cosmic microwave background is the primary source of information, followed by galaxy clustering and weak lensing, and with CMB-lensing adding the smallest amount of information. Comparing the two cosmological models, the entropies in ΛCDM are smaller, reflecting the reduced parameter space in comparison to wCDM, leading to tighter constraints, smaller entries in the parameter covariance matrix, and consequently, of the information entropies. The Shannon and Rényi entropies are related for the Gaussian distributions by a fixed factor, for which we have chosen to compute the case for $\alpha = 1/2$.

In opposition to the absolute entropies, the relative entropies ΔS and ΔS_α are independent under transformations of the random variable, and for this reason the choice of units does not matter. However, in particular for Cosmology, most of the cosmological

Table 5.1.: Absolute Shannon S and Bhattacharyya entropies S_α with $\alpha = 1/2$, in units of nats, for the likelihood of a ΛCDM model, computed from the Fisher matrices.

Probe	Shannon entropy S	Bhattacharyya entropy S_α
CMB	-29.38	-28.03
CMB lensing	1.03	2.38
galaxy clustering	-13.09	-11.75
weak lensing	-8.21	-6.85

Table 5.2.: Absolute Shannon S and Bhattacharyya entropies S_α with $\alpha = 1/2$, in units of nats, characterising the likelihood of a wCDM model.

Probe	Shannon entropy S	Bhattacharyya entropy S_α
CMB	-44.31	-42.76
CMB lensing	-4.62	-3.00
galaxy clustering	-24.40	-22.85
weak lensing	-59.21	-57.67

parameters are defined in a dimensionless way such that it is sensible to compare absolute entropies directly. Fig. 5.1 refers to the total entropy of all cosmological probes individually, and show their scaling with $\mathrm{tr}(F)/n$, $(\mathrm{tr}(F)/n)^{1/p}$ and $\ln\det(F)/n$. It does not come as a surprise that information entropies visibly scale with trace invariants of the Fisher matrix. As stated before, the primary CMB has the highest information content for a ΛCDM, followed by galaxy clustering, weak lensing and CMB-lensing, in that particular order. Additonally, the inequalities of Eq. (5.13) and Eq. (5.14) are fulfilled as expected. Likewise, in Fig. 5.2 shows for the same cosmological probes their respective information content for a wCDM cosmology, where the increased parameter space typically comes along with higher uncertainties and thus, higher entropies. In this case, the gain of information on cosmology is highest for weak lensing, as the CMB is not as sensitive to details of the dark energy equation of state.

5.4. Information update in relative entropies

The total statistical uncertainty can be measured by both the information entropies and invariants of the Fisher-matrix. As such, one should expect a relation between S (or generally S_α) with tr(F), $\mathrm{tr}(F^2)$ and $\det(F)$ at every stage of combining cosmological probes.

Although it is clear that for an uncorrelated multivariate Gaussian distribution with covariance $C_{\mu\nu} \propto \delta_{\mu\nu}$, the entropies of the individual distributions add, $S = \sum_\mu S_\mu$, the same does not hold if correlations are present. In fact, the total entropy is bounded by the conditional and marginal variances, respectively. For both Shannon and Rényi entropies, the conditional error results from the corresponding inverse entry of the Fisher matrix, $\sigma^2_{\mu,c} = (F_{\mu\mu})^{-1}$, such that with $S^{(c)}_\mu \propto \sigma^2_\mu$. Ignoring non-relevant prefactors, one obtains $\exp(-S^{(c)}_\mu) = F_{\mu\mu}$. Using the Hadamard-inequality, one then finds $\exp(-S) = \det(F) \leq \prod_\mu F_{\mu\mu} = \prod_\mu \sigma^2_{\mu,c} = \prod_\mu \exp(-S^{(c)}_\mu) = \exp\left(\sum_\mu S^{(c)}_\mu\right)$, such that $S \geq \sum_\mu S^{(c)}_\mu$ for conditional variances. Conversely, the marginalised variance is computed from the inverse Fisher-matrix, $\sigma^2_{\mu,c} = (F^{-1})_{\mu\mu}$, as well as $\det(F^{-1}) = 1/\det(F)$, implying $\exp(S) =$

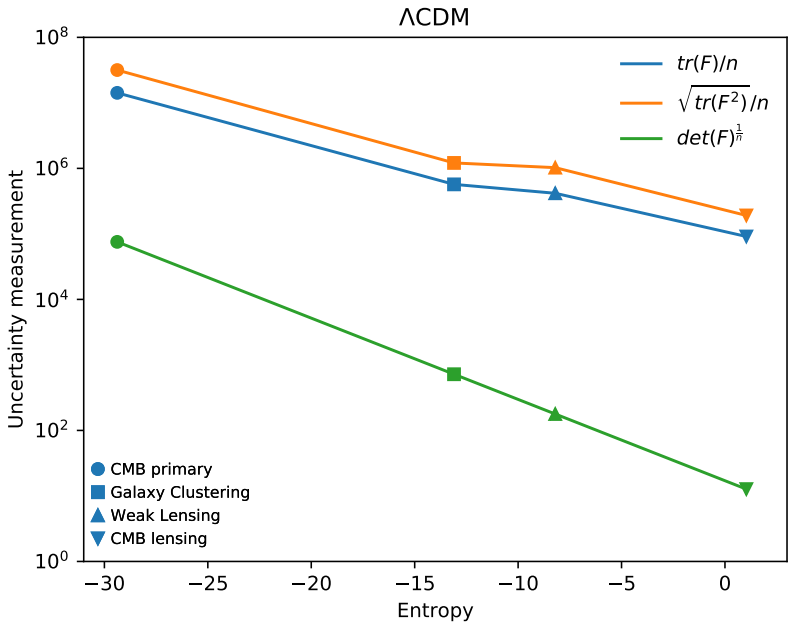

Figure 5.1.: Absolute Shannon entropy S in units of nats for the likelihood of a ΛCDM cosmology, constrained through primary CMB fluctuations, CMB lensing, galaxy clustering and weak lensing individually, plotted against $\mathrm{tr}(F)/n$, $[\mathrm{tr}(F^p)/n]^{1/p}$, $p = 2$, and $\det(F)^{1/n}$.

$\det(F^{-1})$. Then, $\exp(S) = \det(F^{-1}) \leq \prod_\mu (F^{-1})_{\mu\mu} = \prod_\mu \sigma^2_{\mu,m} = \prod_\mu \exp(-S^{(m)}_\mu) = \exp\left(\sum_\mu S^{(m)}_\mu\right)$, and from that $S \leq \sum_\mu S^{(m)}_\mu$. In summary, $\sum_\mu S^c_\mu \leq S \leq \sum_\mu S^{(m)}_\mu$, where equality is given for the uncorrelated case.

The Cramér-Rao inequality asserts that the estimated variance of a distribution is bounded by the Fisher matrix from below, where equality between the variances σ^2 and F^{-1} is only given for a Gaussian distribution. If one estimates the covariance matrix from a non-Gaussian distribution, the resulting variance would be larger than that of a Gaussian distribution for the same Fisher-matrix. Naturally, this assigns an entropy to that covariance through $S \propto \ln\det(C)$ which yield a larger result than $-\ln\det(F)$. This statement is not in contradiction with the property of the Gaussian distribution to maximise S for a given covariance, because the actual value of S depends on the shape of the distribution. The actual value of S has to be either computed from the functional shape or be estimated from data, through $S = -\int \mathrm{d}^n x\, p \ln p$ or $S_\alpha = -\int \mathrm{d}^n x\, p p^{\alpha-1}/(\alpha - 1)$.

It is a standard derivation to show by functional extremisation $\delta S = 0$ of $S = -\int \mathrm{d}x p(x) \ln p(x)$ while incorporating the boundary conditions $\int \mathrm{d}x\, p(x) = 1$ and $\int \mathrm{d}x\, p(x) x^2 = \sigma^2$ with Lagrange-multipliers, that the Gaussian distribution is in fact the one with the largest possible entropy for fixed variance. We would like to point out that the Gaussian distribution is likewise the solution if one fixes the Fisher matrix $F = \langle(\partial \ln p)^2\rangle = \int \mathrm{d}x\, p(\partial \ln p)^2$. Formulating the entropy functional as the averaged logarithmic curvature,

$$S = -\int \mathrm{d}x\, p \ln p + \lambda\left[\int \mathrm{d}x\, p - 1\right] + \mu\left[\int \mathrm{d}x\, p(\partial \ln p)^2 - F\right] \tag{5.46}$$

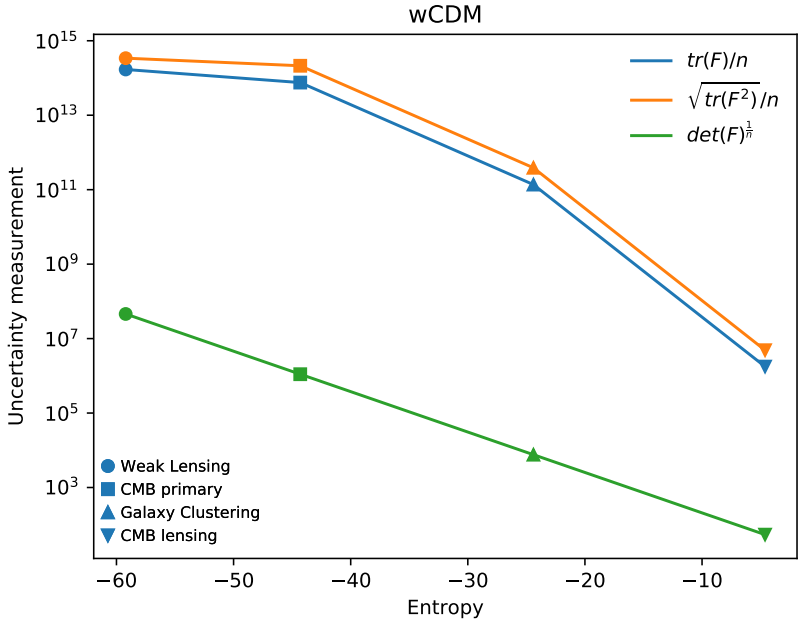

Figure 5.2.: Absolute Shannon entropy S in units of nats for the likelihood of a wCDM cosmology, constrained through primary CMB fluctuations, CMB lensing, galaxy clustering and weak lensing individually, plotted against $\mathrm{tr}(F)/n$, $[\mathrm{tr}(F^p)/n]^{1/p}$, $p = 2$, and $\det(F)^{1/n}$.

yields as a solution to $\delta S = 0$ the differential equation

$$\ln p(x) + 1 + \lambda + \mu \frac{\partial^2 p}{p} = 0 \tag{5.47}$$

using $\partial^2 p/p = -(\partial \ln p)^2$. While identifying F with σ^{-2}, this is solved by the Gaussian distribution $p(x) = \exp(-x^2/(2\sigma^2))/\sqrt{2\pi\sigma^2}$.

We compute information entropy differences for combinations of cosmological data sets and again add successively, in order of decreasing redshift, CMB-lensing, tomographic galaxy clustering and tomographic weak gravitational lensing to the primary CMB-fluctuations. In this computation, we use the resulting combined Fisher matrix including all cross-correlations in order to quantify the gain of information relative to the previous combination of probes. The results are shown in Fig. 5.3 for both the Kullback-Leibler divergence ΔS as well as the α-divergence ΔS_α for $\alpha = 1/2$ for a ΛCDM cosmology in black solid and dashed lines. The analysis is repeated for a marginalised likelihood of each individual parameter Ω_m, σ_8, h, n_s, w_0 and w_a, where the above discussed inequalities for the sum of the marginalised entropies in comparison to the total entropy become relevant. These results are represented in Fig. 5.4. For the ΛCDM case, the σ_8 parameter is the one that gains more information from the combination of all probes compared to the starting point of CMB. On the case of wCDM, the parameter with the highest gain is the matter density parameter Ω_m.

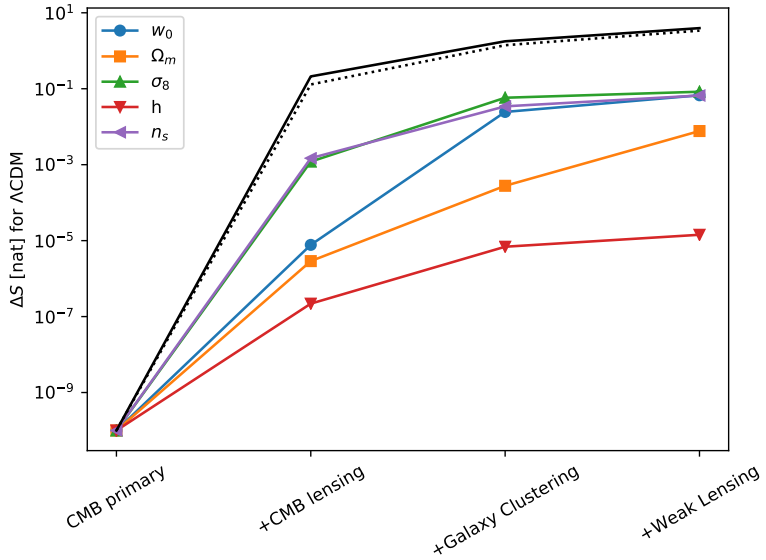

Figure 5.3.: Relative entropies ΔS from the marginalised likelihoods for each of the parameters in a ΛCDM cosmology in units of nats, i.e. for the marginal Gaussian distributions for Ω_m, σ_8, h, n_s, and w_0, for combining successively the four probes from high to low redshift: primary CMB, CMB lensing, galaxy clustering and weak lensing. The black solid and dashed line correspond respectively to the relative entropies ΔS and ΔS_α with $\alpha = 1/2$ for all the model parameters considered.

5.5. Entropy increase through systematics

The assumption of unbiased measurements has been always included throughout this analysis. As such, the averaged likelihoods should be peaked at the fiducial cosmology because the data was on average equal to the theoretical prediction. We will next relax this assumption by considering shifts in the likelihood functions of different cosmological probes due to systematical errors. As introduced by [249], the figure of bias Q from the Fisher-matrix $F_{\mu\nu}$ and the shifts δ_μ of the best-fit point can take these shifts into account. It is obtained through the quadratic form $Q^2 = \sum_{\mu\nu} F_{\mu\nu}\delta_\mu\delta_\nu$ and it is connected to the Kullback-Leibler-divergence $\Delta S = Q^2/2$, if the covariance is unaffected by the systematic. The interpretation of Q is straightforwardly associated to the systematic error in units of the statistical error. It is likewise obvious that there is an effect of systematic errors on relative entropies. Indeed, the explicit relationship for the Kullback-Leibler divergence ΔS between two Gaussian distributions with Fisher matrices $F_{\mu\nu}$ and $G_{\mu\nu}$ has a term involving δ_μ,

$$\Delta S = \frac{1}{2}\left(\delta_\mu G_{\mu\nu}\delta_\nu + \ln\frac{\det(F)}{\det(G)} - n + F^{-1}_{\mu\nu}G_{\mu\nu}\right) \tag{5.48}$$

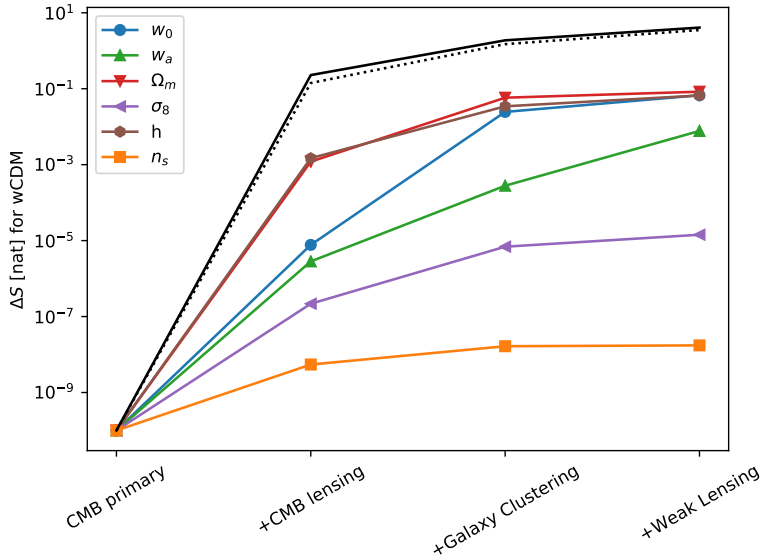

Figure 5.4.: Relative entropies ΔS and ΔS_α for $\alpha = 1/2$ from the marginalised likelihoods for each of the parameters in a wCDM cosmology in units of nats, i.e. for the marginal Gaussian distributions for Ω_m, σ_8, h, n_s, w_0, and w_a, for a successive combination of the four probes primary CMB, CMB lensing, galaxy clustering and weak lensing. The black solid and dashed line correspond respectively to the relative entropies ΔS and ΔS_α with $\alpha = 1/2$ for all the model parameters considered.

which reverts to Eq. (5.27) in the case of vanishing tension, $\delta_\mu = 0$. The analogous relationship for the relative Rényi entropy ΔS_α can be derived to be

$$\Delta S_\alpha = \frac{1}{2}\frac{1}{\alpha - 1} \ln \left(\frac{\det^\alpha(F)}{\det^{\alpha-1}(G)\det(A)} \times \exp\left[-\frac{\alpha}{2} \delta_\alpha \left(F_{\alpha\beta} - \alpha F_{\alpha\mu} A^{-1}_{\mu\nu} F_{\nu\beta} \right) \delta_\beta \right] \right), \tag{5.49}$$

again with

$$A_{\mu\nu} = \alpha F_{\mu\nu} + (1 - \alpha) G_{\mu\nu}, \tag{5.50}$$

where the previous relation for the Rényi entropy ΔS_α is recovered for $\delta_\mu = 0$, as the exponential becomes equal to one.

For the application of the above concepts, we consider three well-known examples of tensions between likelihoods, assmuning Gaussian distributions. These examples are the tension in the value of the Hubble-parameter H_0 between the CMB and Cepheids, the tension in the (Ω_m, σ_8)-plane between the CMB and weak lensing, and intrinsic alignments as a contaminant in weak lensing data as a theoretical forecast. Regarding the interpretation of ΔS and ΔS_α, the Kullback-Leibler divergence for two identical Gaussian distributions shifted by δ is given by $(\delta/(\sqrt{2}\sigma))^2$, such that the square root measures the number of standard deviations by which the Gaussian distributions are displaced relative to each other. This immediately suggests the interpretation of the integrated probability

to obtain values larger than the actual bias δ as

$$P(x \geq \delta) = \frac{1}{2}\mathrm{erfc}\left(\frac{\delta}{\sqrt{2}\sigma}\right), \tag{5.51}$$

where one could substitute the square root of the entropy difference ΔS inside the complementary error function.

5.5.1. Hubble parameter H_0 from Cepheids and the CMB

The Hubble parameter H_0 quantifies the current rate of expansion of the Universe, but values from the CMB [83] and the local value of H_0 [250] are in disagreement by 4.4σ of statistical significance. The CMB temperature fluctuations provides cosmological parameter values with very good statistical precision, assuming a ΛCDM cosmology. More recently, the improvements on the distance ladder using near-infrared Cepheids variables in host galaxies with recent type Ia supernovae reduced the uncertainty on H_0 to 2%. Typically this measurement is model-independent as it follows directly from the Hubble-Lemaître law and does not dependent on a specific cosmological model. Up to this point, all the methods were in agreement but the recent better control over systematics revealed two distinct sides for this H_0 quantity. The origin is yet unknown and most of the formulated solutions involve hidden systematics or new physics. This tension clearly shows up as a nonzero Kullback-Leibler divergence ΔS. With CMB as a reference to be updated by Cepheids, this gives a value of $\Delta S = 44$ nats, and has a value of $Q^2 = 88$ for the figure of bias.

5.5.2. (Ω_m, σ_8)-plane from the CMB and weak lensing

There is an older unsolved disparity between determinations of Ω_m and σ_8 from the CMB and from weak gravitational lensing. This appers as the lensing data preferring smaller values for both parameters relative to the CMB. However, it is worth to recall the well-known degeneracy between these parameters as a lensing essentially determines the product of both parameters and not the parameters individually. For this reason and since the best fit value for (Ω_m, σ_8) does not have a Gaussian uncertainty, we opt to use the parameter $S_8 = \sigma_8\sqrt{\Omega_m/0.3}$. This estimator encapsulates the information of both parameters and has a Gaussian uncertainty. Otherwise, it would not have been possible to use these data as one needs a full correlation matrix between the parameters to compute correctly this shift between likelihoods. One obtains an entropy difference $\Delta S = 2$ nats for the Kullback-Leibler divergence, between Planck's CMB observation and KiDS's weak lensing data set [251], with CMB as the reference value. As for the figure of bias, we obtain $Q = 1$. Again, we would like to emphasise that in contrast to quantities like Q, one does not assume Gaussianity of the distributions.

5.5.3. wCDM and lensing with intrinsic alignments

Lastly, we quantify the effect of intrinsic alignments in weak lensing data on parameter estimation and the bias that they cause. From a physical point of view, intrinsic alignments are mechanisms related to tidal interaction of galaxies with the large-scale structure, which causes them to have correlated intrinsic shapes. This then changes the fundamental assumption that lensing is the only mechanism to generate shape correlations. One can

derive estimation biases for the wCDM parameter set including the dark energy equation of state parameters w_0 and w_a. This is possible only when deriving intrinsic ellipticity spectra using tidal shearing for elliptical and tidal torquing for spiral galaxies including the cross correlation that exists between gravitational lensing and the intrinsic shapes of elliptical galaxies. Expressed in terms of the ratio δ/σ, those are in fact significant for a weak lensing survey like Euclid. In this particular case, we work with the approximation that the intrinsic alignments only give rise to a nonzero bias δ_μ while keeping the covariance, or equivalently, the Fisher matrix $F_{\mu\nu}$ fixed. The numerical value of the Kullback-Leibler divergence is computed to be $\Delta S = 549573$ nats, for an analysis of Euclid's weak lensing data in the framework of a wCDM cosmology.

5.6. Bayesian evidence and information entropy

Bayesian evidence as a criterion for model selection provides a trade-off between the size of the statistical errors and the model complexity. It is straightforward to show that, for a Gaussian likelihood $\mathcal{L}(D|x_\mu, M)$ with a Fisher matrix $F_{\mu\nu}$ and a Gaussian prior $\pi(x_\mu|M)$ with the inverse covariance $P_{\mu\nu}$, the evidence $p(D|M)$ is given by

$$p(D|M) = \sqrt{\frac{\det(F)\det(P)}{(2\pi)^n \det(F+P)}}, \tag{5.52}$$

which implies a scaling $\propto \pi^{-n}/2$ disfavouring models with high complexity.

The expression for the evidence $p(D|M)$ changes to

$$p(D|M) = \sqrt{\frac{\det(F)\det(P)}{(2\pi)^n \det(F+P)}} \exp\left\{-\frac{1}{2}\delta_\mu \left[P_{\mu\nu} - F_{\mu\alpha}(F+P)^{-1}_{\alpha\beta} F_{\beta\nu}\right] \delta_\nu\right\}, \tag{5.53}$$

if likelihood and prior are displaced by δ_μ relative to each other. Comparing this expression with Eq. (5.49), shows that the two expressions are related to each other if $\alpha = 1 - \alpha$, i.e. if $\alpha = 1/2$, which is the Bhattacharyya entropy. For this particular case, the Rényi entropy would weigh both likelihood and prior equally,

$$\Delta S_\alpha = 2 \ln \int \mathrm{d}^n x \sqrt{\mathcal{L}(D|x_\mu, M)\pi(x_\mu|M)}, \tag{5.54}$$

with the natural bound $\sqrt{\mathcal{L}(D|x_\mu, M)\pi(x_\mu|M)} \leq [\mathcal{L}(D|x_\mu, M) + \pi(x_\mu|M)]/2$. This is given by the inequality of the geometric and arithmetic mean such that $\Delta S_\alpha \geq 0$ as both $\mathcal{L}(D|x_\mu, M)$ and $\pi(x_\mu|M)$ are normalised. Because the logarithm is a concave function, one can use Jensen's inequality to write $\ln \int \mathrm{d}^n x \sqrt{\mathcal{L}(D|x_\mu, M)\pi(x_\mu|M)} \geq \int \mathrm{d}^n x \ln \sqrt{\mathcal{L}\pi} = \int \mathrm{d}^n x \, (\ln \mathcal{L} + \ln \pi)/2$, such that $\Delta S_\alpha \leq \int \mathrm{d}^n x \, (\ln \mathcal{L} + \ln \pi)$, bounding the evidence from above, although in most cases this particular bound is diverging.

Finally, expressing the evidence $p(D|M)$ for Gaussian distributions in terms of the Rényi entropy yields

$$p(D|M) = \exp(-\Delta S_\alpha) \sqrt{\det(F+P)}\, \pi^{-n/2}. \tag{5.55}$$

This relation shows that the evidence is made up from three contributions. On the one hand, it decreases $\propto \pi^{-n/2}$ if the dimensionality of the parameter space, i.e. the model complexity is increased. Also, the determinant of the Fisher matrix generates a scaling

$\propto \prod_\mu 1/\sigma_\mu$, such that models with large errors are assigned low evidences as well as being a measure of the dissimilarity of likelihood and prior. Taking the logarithm shows that

$$\ln p(D|M) = -\Delta S_\alpha + \frac{1}{2}\mathrm{tr}\ln(F+P) - \frac{n}{2}\ln\pi, \tag{5.56}$$

meaning that the evidence reflects the inverse volume of the permissible parameter space, allowed by combining likelihood and prior.

Comparing Bayes evidences and information entropy differences shows a clear mathematical relationship between the two, implying perhaps that one could in principle state the relative entropy instead of the Bayesian evidence ratio. The advantage in doing that would be to avoid the empirical Jeffreys scale. Instead, relative entropies would be stated in units of nats. In addition, the usage of the evidence ratio appears to be motivated by the Neyman-Pearson lemma known from hypothesis testing, where it ensures that two hypotheses are tested against each other with the most efficient test statistic. While it seems to be unclear whether the Neyman-Pearson lemma applies to evidences as well in the sense if evidence ratios constitute the most efficient statistical test to decide between two models, these difficulties are avoided by relative entropies. They are axiomatically defined and unambiguous, and the evidence difference, in units of nats, can be computed for every likelihood, with the only complication originating from having to estimate $\ln p(x_\mu)$ from samples in the case of non-Gaussian likelihoods or priors.

Using the current Fisher generator code, it is not possible to directly do this comparison since the prior is fixed to be an unbounded uniform distribution. In future work, we will explore this with the usage of the Monte Carlo Markov chains upgrade of this code, where this comparison can be more extensively explored.

5.7. Summary

The current framework in Cosmology is data driven and employing Bayesian statistics. Given the complexity of systematical effects that blur the cosmological data, a better understanding of statistics is required, or perhaps just a different interpretation. In this work, we have applied concepts from information theory and made an analogy with the commonly used quantities of statistical inference.

Using the same key ingridient of likelihood functions, we compared the associated Fisher matrix and information entropy following the Shannon and the Bhattacharyya definitions. We show how these scale with measures of total error derived from the Fisher matrix, in particular, for the case of $\mathrm{tr}(F)$, $\mathrm{tr}(F^2)$ or $\ln\det(F)$. This comparison was done using specifically the likelihoods for the spectra of the cosmic microwave background temperature and polarisation anisotropies, CMB lensing, galaxy clustering and weak gravitational lensing by the cosmic large-scale structure. Assuming a Gaussian distribution for the cosmological parameters, this computation comes out as purely analytical given the parameter covariance matrix, or equivalently, the inverse of the Fisher matrix.

We compared the conventional ΛCDM and wCDM cosmologies with the aim of quantifying the information content of these cosmological probes. From the absolute entropies for each probe individually, we can conclude that the CMB is the primary source of information for the ΛCDM model and weak lensing for the wCDM scenario. This arises naturally as constraints on the dark energy equation of state stem from lower redshift probes such as weak lensing. Next, we computed the entropy differences for combination

of cosmological data sets, in decreasing order of redshift, having as reference the CMB. This update of the likelihood function as more information is taken into account resembles the usual Bayesian approach. For the individual and all parameters considered of ΛCDM and wCDM cosmologies, we find that the information about the σ_8 is significantly different from the CMB with the additional information for the ΛCDM scenario, where as for the wCDM scenario that is achieved by the Ω_m.

The data tensions can be expressed as a shift between likelihoods which one wants to minimize. We computed the relative entropy and the figure of bias for the three known data discrepancies: on the H_0 parameter from Cepheid variable stars and the cosmic microwave background, the tension in (Ω_m, σ_8) between weak lensing and the CMB, and the bias caused by intrinsic alignments in weak gravitational lensing in wCDM models. Also, we showed how the Bayesian evidence and information entropy are related, suggesting that the information entropy can be used in a model comparison analysis. There is a relation between the entropy difference ΔS_α and the evidence $p(D|M)$ if α is chosen to be $1/2$. This comparison is beyond the Fisher code used in this work and an MCMC approach will enable this work in the future.

The main goal of this work is to show the wide range of applications of the information theory concepts. The connection between Bayesian statistics and information entropy measures is straightforward which enables another perspective on distinguishing uncertainties components, systematical errors or tensions between datasets and evidences. Even further, it can provide with a viable alternative to the Jeffreys scale which thresholds are sometimes not conclusive or even reliable for model comparison. The expressions here derived are assuming a Gaussian likehood which is a common and generally good assumption in Cosmology but information entropies are not dependent of this assumption. Further exploration of these concepts is one of the next goals of future work. Information entropies are suited for any probability distribution which acts as an advantadge with respect to a Fisher matrix formalism. On the other hand, the additivity and other Fisher matrix properties are extremly useful, surpassing the Kullback-Leibler divergence with its lack of triangule property. The progress in cosmology envolves not just new theoretical possibilities but also a new take on statistical inference, which hopefully can clear out the data tensions and develop our knowledge on the history and dynamics of the Universe.

6. Summary and outlook

In this dissertation, we covered different approaches to assess deviations to the standard model of cosmology, the ΛCDM model. This includes estimators built to test gravity and a new framework to measure information in statistical inference.

In Chapter 2, we have explained the foundations of General Relativity and the ΛCDM model. Regarding alternative theories, we described the difference between modified gravity and dark energy models, detailing the Horndeski theory. We introduced common parametrizations of modified gravity and details about the structure formation. On the observational side, we characterized the most important cosmological probes: the cosmic microwave background, weak lensing, galaxy clustering and Hubble parameter.

Next, in Chapter 3, we explored how the anisotropic stress parameter can be computed in a model-independent way. The recent gravitational-wave events have ruled out several models by setting the speed of gravity to be nearly the speed of light. This constrained, for example, a few sectors of the Horndeski theories. Since most of the cosmological data points towards a ΛCDM description of the Universe, we need to test gravity in a model-independent way. The anisotropic stress parameter, η, is defined by the ratio of the gravitational potentials Φ and Ψ. In General Relativity, $\eta = 1$ for all scales and redshifts. If η is determined to be different from one, it means that the perfect fluid approach is not valid. This allows us to discriminate and rule out models that have this assumption.

We estimate the anisotropic stress parameter, η, in a model-independent way. For this, we use the set of observable quantities that do not depend on the assumptions of the initial conditions of the Universe or about the primordial power spectrum. Also in these quantities, it is not assumed a particular description of dark matter or an expression for the galaxy bias. With the ratios of these observable quantities, we can build model-independent parameters [65, 145] that correspond to linear relations of cosmological data. If one substitutes the model-independent parameters in the linearized Poisson equations in Fourier space, we obtain an expression for $\eta_{\rm obs}$ written in terms of the data observables.

These data are compilations of the Hubble parameter, the $f\sigma_8$, and the E_G statistics. Each of these datasets has measurements taken at different redshifts. Additionally, the model-independent estimator for η includes derivatives of the direct data. For these reasons, we need data reconstruction methods to obtain a continuous function of redshift. Firstly, we use the binning method which consists of a weighted average for a specific redshift interval. This simple method can be useful in large datasets as it has the least amount of assumptions. However, it does not capture higher mode behaviors and it strongly limits the number of data points available after applied. As a second method, we use the Gaussian Processes method. This is a generalization of the Gaussian distribution to function space where it is assumed that the data are Gaussian distributed. It is not exactly a non-parametric method because a kernel function is chosen as the covariance matrix. Finally, we do a polynomial regression where the exponents and coefficients are determined by each dataset. This method provides an easy computation for the derivatives and a guaranteed smooth function of redshift. It requires a lower inconsistency in the dataset which means there should not a big scatter of values for a small redshift interval.

The binning method reduces the redshift range to three points at which we compare each method. Gaussian Process is the method that delivers the most stringent error bars. For the first bin, the results from the different methods are compatible with each other. The binning method is 1.5σ away from the Gaussian Processes, in the second bin. The last and farthest bin, the errors are much larger and the comparison is not significant. Conclusions are dependent on the redshift and method. For instance, in the case of the second bin of the binning, it is two or even three sigma away from the standard gravity value $\eta = 1$. In general, the results are in every bin compatible with at least two of the three methods. The polynomial regression appears to be the most conservative error bars. We quote that as our method and we find $\eta_{\rm obs} = 0.44 \pm 0.92$ at $z = 0.294$, $\eta_{\rm obs} = 0.42 \pm 0.89$ at $z = 0.58$, and $\eta_{\rm obs} = -0.14 \pm 3.01$ at $z = 0.86$.

The future large scale surveys will soon obtain large amounts of data. For example, Euclid combines galaxy clustering and lensing with unprecedented quality. The forecasts of [62] show that a constant $\eta_{\rm obs}$ could be measured up to a few percent. To revisit this estimator when the Euclid [178] data are available is the next step. We would like to emphasize that to performing null-tests is as important as testing specific models.

We explored another gravity test called the E_G statistics in Chapter 4. This estimator distinguishes modified gravity models through the comparison of weak lensing and galaxy clustering signals. The theoretical definition simplifies to a ratio of the matter density over the growth function for the ΛCDM model. However, the observational definition of this estimator takes the galaxy-lensing and galaxy-galaxy two-point correlation functions, and only recovers the simplified expression under certain conditions. We have explored this definition deeper intending to understand how far the correspondence between the theoretical and observational definitions is valid.

It has been shown before by [146] that there are a few choices possible in the E_G statistics. First, it depends on the redshift distribution of sources $P_s(z)$, which is usually defined by the experimental setup. The cutoff-scale R_0 is chosen and this is defined as the scale below which information is discarded. The projection length δ_ℓ at the galaxy-galaxy cross-correlation is another parameter that is chosen. Finally, it also depends on the bias function $b(z, k)$ which is usually taken as linear but in reality this an assumption. It is important to confirm that this definition should allow for alternative theories of gravity. This should be done by carefully account for the geometry of the sources and lensed galaxies. After deriving the equations and implement them numerically, we identified the relevant parameter space. This is work in progress but it is present in this dissertation as another example of an estimator build to test gravity.

Unlike the model-independent estimation for η, the E_G statistics has been measured by a few lensing and galaxy clustering surveys. Given the current data tensions which also include the estimation of the matter density, the E_G statistics needs to be taken with a grain of salt. We point out that an estimator truly is model-independent can help assess properties of gravity. A correct calculation of the E_G statistics could be interesting to measure with future missions such as Euclid [178].

Finally, we applied information theory concepts to the cosmological framework in Chapter 5. The precision era of Cosmology includes plenty of observables and data which are often combined in a single analysis. At the moment each analysis is also more complex where systematical effects can easily undermine its accuracy or a discovery. The majority of the currently available data prefers a ΛCDM description of the Universe while still allowing for alternative theories. The motivation for a new approach in statistical inference

stems from the lack of distinction between theories and it is also suggested by the recent data tensions.

Typically in statistical inference, the main quantity is the likelihood function which encompasses both the model and the data. If the likelihood function is a Gaussian distribution, it is defined by a mean vector and a covariance matrix. For a fixed model, we can compute the Fisher matrix from which the inverse is the covariance matrix. When forecasting for future surveys, the Fisher matrix is the primary tool used to estimate the parameter uncertainties. Through geometrical matrix properties, such as the trace, it is how often the information from a Fisher matrix is summarized. As a comparison, we have used the information theory concept of entropy to measure the amount of information in a dataset. We computed the Shannon and the Bhattacharyya entropy definitions and we have shown how these scale with measures of total error derived from the Fisher matrix. In particular, we compare the entropy with the Fisher matrix properties $\mathrm{tr}(F)$, $\mathrm{tr}(F^2)$ or $\ln\det(F)$. Specifically, this comparison was done using the likelihoods for the spectra of the cosmic microwave background temperature and polarization anisotropies, CMB lensing, galaxy clustering and weak gravitational lensing by the cosmic large-scale structure. This computation comes out as purely analytical since we assume a Gaussian distribution defined by the cosmological parameters covariance matrix, or equivalently, the inverse of the Fisher matrix.

We compared the ΛCDM and wCDM cosmologies to quantify how much information can each cosmological probe provide. The computed absolute entropies show that the CMB is the primary source of information for the ΛCDM model and weak lensing for the wCDM scenario. This may reflect from the fact that weak lensing probes the late Universe and therefore naturally constraints on the dark energy equation of state. The combination of cosmological datasets is also a choice that can help break parameter degeneracies or hide them. For this reason, we set the CMB as the reference and updated the corresponding likelihood with more information from the remaining probes, in decreasing order of redshift. In this Bayesian-like approach, we computed the entropy differences for probe combination, for all and each parameter considered in a ΛCDM and wCDM cosmology. We find that the information about the σ_8 is significantly different from the CMB with the additional information for the ΛCDM scenario, whereas for the wCDM scenario that is done by the Ω_m.

The degree of dissimilarity between likelihoods is what is important to measure when understanding the data tensions. The relative entropies and the figure of bias are two ways of evaluating this. We have computed those quantities for the H_0 tension between Cepheid variable stars and the cosmic microwave background, the tension in (Ω_m, σ_8) parameter plane between weak lensing and the CMB, and the bias caused by intrinsic alignments in weak gravitational lensing in wCDM models. Finally, we derived how the Bayesian evidence and information entropy are related. This means that information entropies can be useful to compare models. A particularly interesting quantity is the relative entropy ΔS_α if α is chosen to be $1/2$ when compared with the evidence $p(D|M)$. However, this comparison requires a full Monte-Carlo Markov chain approach to fully compute the posterior distribution which will be done in future work. Also, the same sampling code for the posterior distribution could enable to take the comparison further to a non-Gaussian distribution since the information theory concepts can be applied in any probability distribution.

The goal of this dissertation was to understand what are the next steps towards finding a

complete description of the Universe. This is especially important when most observational data prefer the ΛCDM model but at the same time, several questions remain unsolved. We consider of fundamental value a model-independent approach in cosmology, or at least, to be fully aware of every assumption. We propose two ways to address this issue: measuring estimators built for testing deviations of general relativity or ΛCDM, and an alternative point of view on how statistics is used. The future of cosmology requires not just new theories but also new perspectives on statistical inference. These steps hopefully contribute to clear out the data tensions and deepen our understanding of the Universe.

A. Details of the Polynomial Regression Method for the reconstruction of η_{obs}

We want to estimate the value of the functions $y^{(i)}$, for $i = 1, 2, 3, 4$ at a number of arbitrary points, labeled by subscripts $A, B, C \ldots$, which we call the interpolated points and assume that there is a domain $\mathcal{D}$ common to all datasets, in which all the interpolated points are contained. We further assume that the three initial datasets, $y^{(j)}$, for $j = 0, 1, 2$ are independent of each other. Now we use for all initial datasets, a polynomial of the form

$$g^{(j)} = \sum_{\alpha=0}^{N_j} (1+x)^\alpha \,, \tag{A.1}$$

with N_j the maximum order of the polynomial, which depends on the characteristics of each dataset $y^{(j)}$, and will be explained further below. Then the function f in Eq. (3.24) will have $N_j + 1$ coefficients, ranging from $\bar{A}_0$ to $\bar{A}_{N_j}$. If we now take the derivative of this function, we obtain

$$f^{(j)\prime} = \sum_{\mu=1}^{N_j} -\mu \bar{A}_\mu^{(j)} g_\mu^{(j)} = -\sum_{\mu=1}^{N_j} \bar{B}_\mu^{(j)} g_\mu^{(j)} \,, \tag{A.2}$$

where $g_\alpha^{(j)}$ is the α-th term in the sum $g^{(j)}$. For notational simplicity we define the indices α, β to always run from 0 to N_j, while the indices μ, ν will run from 1 to N_j. As we can see, the derivative functions $f^{(j)\prime}$ have one coefficient less, because there is no A_0 coefficient. The relation between the old and new coefficients is

$$\bar{B}_\mu = \mu \bar{A}_\mu \,. \tag{A.3}$$

This means that the covariance matrix $(F^j)^{-1}$ of the coefficients $\bar{A}^j$ has to be modified with a Jacobian of the form

$$J^j_{\mu\alpha} = \frac{\partial \bar{B}_\mu}{\partial \bar{A}_\alpha} = \mu \delta_{\alpha\mu} = \text{diag}(0, 1, 2, ..., N_j) \,, \tag{A.4}$$

to obtain the covariance matrix $\tilde{F}$ of the new coefficients

$$(\tilde{F}^j)^{-1}_{\mu\nu} = J^j_{\mu\alpha} (F^j_{\alpha\beta})^{-1} J_{\beta\nu} \,. \tag{A.5}$$

Since $\alpha, \beta = 0, \ldots, N_j$ and $\mu, \nu = 1, \ldots, N_j$ the Jacobian is a rectangular matrix of dimensions $(N_j - 1) \times N_j$, therefore the $\tilde{F}$ matrices have a dimension equal to the original F minus unity.

Summarizing, we have the following four functions at the wanted points A

$$\bar{f}_A^{(a)} = \bar{B}^{(a)}_{\{\alpha,\mu\}} p^{(a)}_{A\{\alpha,\mu\}} \,. \tag{A.6}$$

Where due to the derivative, we have the following basis functions,

$$p_\alpha^{(1)} = g_\alpha^{(1)}\,, \qquad p_\alpha^{(2)} = g_\alpha^{(2)}\,, \qquad p_\mu^{(3)} = -g_\mu^{(0)}\,, \qquad p_\mu^{(4)} = -g_\mu^{(1)}\,, \tag{A.7}$$

for $\alpha = 0, \ldots, N_j$ and $\mu = 1, \ldots, N_j$. Which in turn leads to a change in the vector of coefficients, such that they read now

$$\bar{B}_\alpha^{(1)} = \bar{A}_\alpha^{(1)}\,, \qquad \bar{B}_\alpha^{(2)} = \bar{A}_\alpha^{(2)}\,, \qquad \bar{B}_\mu^{(3)} = \mu\bar{A}_\mu^{(0)}\,, \qquad \bar{B}_\mu^{(4)} = \mu\bar{A}_\mu^{(1)}\,. \tag{A.8}$$

The Fisher matrices for $\bar{B}^{(1)}$ and $\bar{B}^{(2)}$, are $F^{(1)}$ and $F^{(2)}$, respectively. For $\bar{B}^{(3)}$ the Fisher matrix is $\tilde{F}^{(3)}$, while for $\bar{B}^{(4)}$ it is $\tilde{F}^{(4)}$. The $\tilde{F}$ matrices have a dimension smaller by one unit than the original F,

$$C_{\alpha\beta}^{(1,1)} = \mathrm{Var}(\bar{B}_\alpha^{(1)}\bar{B}_\beta^{(1)}) = \left(F^{(1)}\right)^{-1}_{\alpha\beta}\,, \tag{A.9}$$

$$C_{\alpha\beta}^{(2,2)} = \mathrm{Var}(\bar{B}_\alpha^{(2)}\bar{B}_\beta^{(2)}) = \left(F^{(2)}\right)^{-1}_{\alpha\beta}\,, \tag{A.10}$$

$$C_{\mu\nu}^{(3,3)} = \mathrm{Var}(\bar{B}_\mu^{(3)}\bar{B}_\nu^{(3)}) = \left(\tilde{F}^{(3)}\right)^{-1}_{\mu\nu}\,, \tag{A.11}$$

$$C_{\mu\nu}^{(4,4)} = \mathrm{Var}(\bar{B}_\mu^{(4)}\bar{B}_\nu^{(4)}) = \left(\tilde{F}^{(4)}\right)^{-1}_{\mu\nu}\,, \tag{A.12}$$

$$C_{\alpha\beta}^{(1,4)} = \mathrm{Var}(\bar{B}_\alpha^{(1)}\bar{B}_\beta^{(4)}) = \mathrm{Var}(\bar{A}_\alpha^{(1)}\beta\bar{A}_\beta^{(1)}) = \left(F^{(1)}\right)^{-1}_{\alpha\gamma} J_{\gamma\beta}\,. \tag{A.13}$$

The full matrix $\mathcal{C}_{ab,AB}$ is our final result: the covariance matrix at any two different points $x_A,\ x_B$ for any pairs of datasets $f^{(a)}, f^{(b)}$

$$\mathcal{C}_{ab,AB} = C_{\alpha\beta}^{(a,b)} p_{A\alpha}^{(a)} p_{B\beta}^{(b)}\,. \tag{A.14}$$

Acknowledgements

Acknowledgements

I would like to start by thanking my supervisor Luca Amendola for giving me this opportunity and guiding me along on this journey. It was not an easy time for me but I have learned a lot in the process and not just about Physics and the Universe. Grazie Mille for your patience. Another big thank you or gracias (it should be Danke!) to Björn Malte Schäfer for his fascination for Physics and enthusiasm for the small details of life. Also, I would like to thank Eduard Thommes for taking care of so many administrative details and sharing the same enthusiasm for music.

I am grateful to Matthias Bartelmann for being my second referee to this PhD dissertation.

Throughout this learning process, I realized I needed some help. I would like to thank Eunice Casanova, Olga Degtyareva, and Frank-Hagen Hofmann for the precious guidance into insights, finding my way into being productive and deeply understanding myself. It was life-changing and empowering.

Next, I would like to thank all the crackpots that handle my crazy times and my depressive times. Sharing an office can be a very interesting experience and so many memories from the past years were amazing exactly because of these guys. I thank all my office mates, in particular, the last ones the ones that saw this dissertation coming into shape: Maya, David, Giorgio, Arvid, Leonardo, and Raquel. Thank you Manuel, the only constant office mate, for all the time we shared discussing Physics, organizational stuff or just daily life. Thank you, Julius, for making the office so dynamic. And finally, to the true crackpots, thank you, Henrik, for all the crazy hours and kind words. To Santiago, thank you for all the teachings from Mathematica to dancing salsa. Thank you, Frank, for your the neverending energy and good times. To Kevin, a huge thank you for all the music we did together as well as everything you taught me. To Lorenzo for being the craziest adorable friend I could have, you honestly changed my life with all the music.

My time in Heidelberg would have not been the same if it wasn't for a few more friends. I would like to thank Oscar for all the BroZeit and coffee expertise. To Mary, thank you for your uplifting smile and friendship. To Javi for all the support in many areas of my life, thank you. To Veronika thank you for the lovely dinners and time together. And finally, thank you, Patrick, Daniel and all the Portuguese people that made me feel like home every time we met. Speaking of which, part of me remains in Portugal where Mariana, Martins, Catarina, and Patrick are my keystones, being there to support the next adventure, as it has been for a while. And even if I went to the dark side, my astrofamily João, Filipe, Raquel, Elsa, and Catarina remains connected. To all of you, a big thank you for remembering me of myself.

Even the distance, my family is one of the most important things in my life. I would regularly come back and always be sure that we would all share a nice meal as if the

distance was never there. I'm glad nowadays it's easy to get a new picture of my baby cousins or quickly have a video call where I can see all my family, sing "Happy Birthday" on every anniversary that I am away and ask to ship some cake. I would like to thank in particular to my grandma for always caring about me.

Finally, I would like to thank to the most important people in my life. À minha mãe e ao seu hábito adorável de estar presente no meu dia-a-dia, todos os dias, chamada ou mensagem, pelas visitas frequentes para que não estejamos mais do que dois meses sem nos encontrarmos e partilhar um chá e o sofá (ou equivalente na minha casa pequenina). Ao meu pai e às suas perspetivas, apoio e boa disposição. Que não falte uma boa música do BB King e um delicioso jantar. Não existem palavras para agradecer o quão sortuda sou em termos como o meu núcleo de família. And last, but not least, Manuel. So much of my strength and perseverance were inspired and sourced by you being in my life. It has been an amazing journey together to which the distance barely did any harm. Thank you so much for all your kind words, the stupid jokes and every moment you were there, physically or virtually, for it would not have been possible to achieve this without you.

Bibliography

[1] A. G. Riess, S. A. Rodney, D. M. Scolnic, *et al.*, "Type Ia Supernova Distances at z > 1.5 from the Hubble Space Telescope Multi-Cycle Treasury Programs: The Early Expansion Rate," (2017), arXiv:1710.00844.

[2] S. Alam, M. Ata, S. Bailey, *et al.*, "The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample," Monthly Notices of the Royal Astronomical Society (2016), 10.1093/mnras/stx721, arXiv:1607.03155.

[3] C. Blake, S. Brough, M. Colless, *et al.*, "The WiggleZ Dark Energy Survey: joint measurements of the expansion and growth history at z < 1," Monthly Notices of the Royal Astronomical Society 425, 405 (2012).

[4] A. M. Pinho, S. Casas, and L. Amendola, "Model-independent reconstruction of the linear anisotropic stress η," JCAP 1811, 027 (2018), arXiv:1805.00027 [astro-ph.CO].

[5] L. Amendola, A. M. Pinho, and S. Casas, "Model-independent measures of gravity at large scales," International Journal of Modern Physics A 33, 1844022 (2018), https://doi.org/10.1142/S0217751X18440220.

[6] L. Amendola, D. Bettoni, A. M. Pinho, and S. Casas, "Measuring gravity at cosmological scales," Universe 6 (2020), 10.3390/universe6020020.

[7] A. M. Pinho, R. Reiscke, and B. M. Schäfer, "Information entropy in cosmological inference problems," (2020), in preparation.

[8] L. Amendola and A. M. Pinho, "The $E_G$statistics, revisited," (2020), in preparation.

[9] B. Bertotti, L. Iess, and P. Tortora, "A test of general relativity using radio links with the Cassini spacecraft," Nature 425, 374 (2003).

[10] M. Kramer, I. H. Stairs, R. N. Manchester, M. A. McLaughlin, A. G. Lyne, R. D. Ferdman, M. Burgay, D. R. Lorimer, A. Possenti, N. D'Amico, and et al., "Tests of general relativity from timing the double pulsar," Science 314, 97–102 (2006).

[11] B. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, and et al., "Observation of gravitational waves from a binary black hole merger," Physical Review Letters 116 (2016), 10.1103/physrevlett.116.061102.

[12] S. Perlmutter *et al.* (Supernova Cosmology Project), "Measurements of Ω and Λ from 42 high redshift supernovae," Astrophys. J. 517, 565 (1999), arXiv:astro-ph/9812133 [astro-ph].

[13] A. G. Riess *et al.* (Supernova Search Team), "Observational evidence from supernovae for an accelerating universe and a cosmological constant," Astron. J. 116, 1009 (1998), arXiv:astro-ph/9805201 [astro-ph].

[14] K. C. Freeman, "On the disks of spiral and s0 galaxies," Astron. J. 160, 811 (1970).

[15] V. C. Rubin and J. Ford, W. Kent, "Rotation of the andromeda nebula from a spectroscopic survey of emission regions," Astron. J. 159, 379 (1970).

[16] V. C. Rubin, J. Ford, W. K., and N. Thonnard, "Rotational properties of 21 sc galaxies with a large range of luminosities and radii, from ngc 4605 (r=4kpc) to ugc 2885 (r=122kpc)." Astron. J. 238, 471 (1980).

[17] A. Einstein, "The Field Equations of Gravitation," Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.) 1915, 844 (1915).

[18] A. Einstein, "Cosmological Considerations in the General Theory of Relativity," Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.) 1917, 142 (1917).

[19] W. de Sitter, "Einstein's theory of gravitation and its astronomical consequences. Third paper," Monthly Notices of the Royal Astronomical Society 78, 3 (1917).

[20] J. Martin, "Everything you always wanted to know about the cosmological constant problem (but were afraid to ask)," Comptes Rendus Physique 13, 566–665 (2012).

[21] C. P. Burgess, "Quantum gravity in everyday life: General relativity as an effective field theory," Living Reviews in Relativity 7 (2004), 10.12942/lrr-2004-5.

[22] A. Ashtekar, M. Reuter, and C. Rovelli, "From General Relativity to Quantum Gravity," arXiv e-prints , arXiv:1408.4336 (2014), arXiv:1408.4336 [gr-qc].

[23] I. Zlatev, L. Wang, and P. J. Steinhardt, "Quintessence, cosmic coincidence, and the cosmological constant," Physical Review Letters 82, 896–899 (1999).

[24] L. Amendola and S. Tsujikawa, *Dark Energy: Theory and Observations* (2010).

[25] A. Joyce, L. Lombriser, and F. Schmidt, "Dark energy versus modified gravity," Annual Review of Nuclear and Particle Science 66, 95–122 (2016).

[26] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, "Modified gravity and cosmology," Physics Reports 513, 1–189 (2012).

[27] L. Amendola, "Coupled quintessence," Physics Review D 62, 043511 (2000), arXiv:astro-ph/9908023 [astro-ph].

[28] L. Amendola, S. Appleby, D. Bacon, T. Baker, M. Baldi, N. Bartolo, A. Blanchard, C. Bonvin, S. Borgani, E. Branchini, C. Burrage, S. Camera, C. Carbone, L. Casarini, M. Cropper, C. de Rham, C. Di Porto, A. Ealet, P. G. Ferreira, F. Finelli, J. García-Bellido, T. Giannantonio, L. Guzzo, A. Heavens, L. Heisenberg, C. Heymans, H. Hoekstra, L. Hollenstein, R. Holmes, O. Horst, K. Jahnke, T. D. Kitching, T. Koivisto, M. Kunz, G. La Vacca, M. March, E. Majerotto, K. Markovic, D. Marsh, F. Marulli, R. Massey, Y. Mellier, D. F. Mota, N. J. Nunes, W. Percival, V. Pettorino, C. Porciani, C. Quercellini, J. Read,

M. Rinaldi, D. Sapone, R. Scaramella, C. Skordis, F. Simpson, A. Taylor, S. Thomas, R. Trotta, L. Verde, F. Vernizzi, A. Vollmer, Y. Wang, J. Weller, and T. Zlosnik, "Cosmology and Fundamental Physics with the Euclid Satellite," Living Reviews in Relativity 16, 6 (2013), arXiv:1206.1225 [astro-ph.CO].

[29] I. D. Saltas, I. Sawicki, L. Amendola, and M. Kunz, "Anisotropic stress as a signature of nonstandard propagation of gravitational waves," Phys. Rev. Lett. 113, 191101 (2014).

[30] C. Skordis, "Consistent cosmological modifications to the einstein equations," Physical Review D 79 (2009), 10.1103/physrevd.79.123527.

[31] A. Pourtsidou, C. Skordis, and E. J. Copeland, "Models of dark matter coupled to dark energy," Physical Review D 88 (2013), 10.1103/physrevd.88.083505.

[32] G. W. Horndeski, "Second-Order Scalar-Tensor Field Equations in a Four-Dimensional Space," International Journal of Theoretical Physics 10, 363 (1974).

[33] A. De Felice and S. Tsujikawa, "Conditions for the cosmological viability of the most general scalar-tensor theories and their applications to extended Galileon dark energy models," Journal of Cosmology and Astroparticle Physics 2012, 007 (2012), arXiv:1110.3878 [gr-qc].

[34] C. Deffayet and D. A. Steer, "A formal introduction to Horndeski and Galileon theories and their generalizations," Classical and Quantum Gravity 30, 214006 (2013), arXiv:1307.2450 [hep-th].

[35] G. Gubitosi, F. Piazza, and F. Vernizzi, "The effective field theory of dark energy," Journal of Cosmology and Astroparticle Physics 2013, 032–032 (2013).

[36] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, "Essential building blocks of dark energy," Journal of Cosmology and Astroparticle Physics 2013, 025–025 (2013).

[37] L. Amendola, "Linear and nonlinear perturbations in dark energy models," Physics Review D 69, 103524 (2004), arXiv:astro-ph/0311175 [astro-ph].

[38] V. Pettorino and C. Baccigalupi, "Coupled and extended quintessence: Theoretical differences and structure formation," Phys. Rev. D 77, 103003 (2008).

[39] L. Amendola, M. Baldi, and C. Wetterich, "Quintessence cosmologies with a growing matter component," Physical Review D 78 (2008), 10.1103/physrevd.78.023015.

[40] C. Wetterich, "Growing neutrinos and cosmological selection," Physics Letters B 655, 201–208 (2007).

[41] C. Deffayet, G. Dvali, and G. Gabadadze, "Accelerated universe from gravity leaking to extra dimensions," Physics Review D 65, 044023 (2002), arXiv:astro-ph/0105068 [astro-ph].

[42] D. Blas, O. Pujolàs, and S. Sibiryakov, "Models of non-relativistic quantum gravity: the good, the bad and the healthy," Journal of High Energy Physics 2011, 18 (2011), arXiv:1007.3503 [hep-th].

[43] Y. Akrami, S. F. Hassan, F. Könnig, A. Schmidt-May, and A. R. Solomon, "Bimetric gravity is cosmologically viable," Physics Letters B 748, 37 (2015), arXiv:1503.07521 [gr-qc].

[44] M. Wittner, F. Könnig, N. Khosravi, and L. Amendola, "Beyond dRGT - a new massive gravity theory?" arXiv e-prints , arXiv:1801.07643 (2018), arXiv:1801.07643 [gr-qc].

[45] F. Koennig, Y. Akrami, L. Amendola, M. Motta, and A. R. Solomon, "Stable and unstable cosmological models in bimetric massive gravity," Physics Review D 90, 124014 (2014), arXiv:1407.4331 [astro-ph.CO].

[46] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, "Exploring gravitational theories beyond Horndeski," Journal of Cosmology and Astroparticle Physics 2015, 018 (2015), arXiv:1408.1952 [astro-ph.CO].

[47] B. Hu, M. Raveri, N. Frusciante, and A. Silvestri, "Effective field theory of cosmic acceleration: An implementation in camb," Phys. Rev. D 89, 103530 (2014).

[48] E. Bellini and I. Sawicki, "Maximal freedom at minimum cost: linear large-scale structure in general modifications of gravity," Journal of Cosmology and Astroparticle Physics 2014, 050–050 (2014).

[49] M. Zumalacárregui, E. Bellini, I. Sawicki, J. Lesgourgues, and P. G. Ferreira, "hi_class: Horndeski in the Cosmic Linear Anisotropy Solving System," JCAP 2017, 019 (2017), arXiv:1605.06102 [astro-ph.CO].

[50] J. Lesgourgues, "The cosmic linear anisotropy solving system (class) i: Overview," (2011), arXiv:1104.2932 [astro-ph.IM].

[51] B. P. Abbott *et al.*, "Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A," The Astrophysical Journal 848, L13 (2017).

[52] B. P. Abbott *et al.*, "Search for Post-merger Gravitational Waves from the Remnant of the Binary Neutron Star Merger GW170817," The Astrophysical Journal 851, L16 (2017).

[53] A. Albert *et al.*, "Search for High-energy Neutrinos from Binary Neutron Star Merger GW170817 with ANTARES, IceCube, and the Pierre Auger Observatory," The Astrophysical Journal 850, L35 (2017).

[54] B. P. Abbott, others (LIGO Scientific Collaboration, and Virgo), "Estimating the Contribution of Dynamical Ejecta in the Kilonova Associated with GW170817," The Astrophysical Journal 850, L39 (2017).

[55] B. P. Abbott *et al.*, "Multi-messenger Observations of a Binary Neutron Star Merger*," The Astrophysical Journal Letters 848, L12 (2017).

[56] L. Lombriser and A. Taylor, "Breaking a dark degeneracy with gravitational waves," Journal of Cosmology and Astroparticle Physics 3, 031 (2016), arXiv:1509.08458.

[57] L. Lombriser and N. A. Lima, "Challenges to self-acceleration in modified gravity from gravitational waves and large-scale structure," Physics Letters B 765, 382 (2017), arXiv:1602.07670.

[58] J. M. Ezquiaga and M. Zumalacárregui, "Dark Energy After GW170817: Dead Ends and the Road Ahead," Physical Review Letters 119, 251304 (2017).

[59] J. Sakstein and B. Jain, "Implications of the Neutron Star Merger GW170817 for Cosmological Scalar-Tensor Theories," Physical Review Letters 119, 251303 (2017).

[60] P. Creminelli and F. Vernizzi, "Dark Energy after GW170817 and GRB170817A," Physical Review Letters 119, 251302 (2017).

[61] T. Baker *et al.*, "Strong Constraints on Cosmological Gravity from GW170817 and GRB 170817A," Physical Review Letters 119, 251301 (2017).

[62] L. Amendola, S. Fogli, A. Guarnizo, M. Kunz, and A. Vollmer, "Model-independent constraints on the cosmological anisotropic stress," Phys. Rev. D89, 063538 (2014), arXiv:1311.4765 [astro-ph.CO].

[63] V. Mukhanov, *Physical Foundations of Cosmology* (Cambridge University Press, 2005).

[64] M. Kunz, "The phenomenological approach to modeling the dark energy," Comptes Rendus Physique 13, 539–565 (2012).

[65] L. Amendola *et al.*, "Observables and unobservables in dark energy cosmologies," Physics Review D (2012), 10.1103/PhysRevD.87.023501, arXiv:1210.0439.

[66] I. Sawicki, I. D. Saltas, M. Motta, L. Amendola, and M. Kunz, "Nonstandard gravitational waves imply gravitational slip: On the difficulty of partially hiding new gravitational degrees of freedom," Physical Review D 95 (2017), 10.1103/physrevd.95.083520.

[67] H. Nersisyan, N. A. Lima, and L. Amendola, "Gravitational wave speed: Implications for models without a mass scale," (2018), arXiv:1801.06683 [astro-ph.CO].

[68] A. H. Guth, "Inflationary universe: A possible solution to the horizon and flatness problems," Phys. Rev. D 23, 347 (1981).

[69] X. Chen, "Primordial non-gaussianities from inflation models," Advances in Astronomy 2010, 1–43 (2010).

[70] A. Lewis, A. Challinor, and A. Lasenby, "Efficient computation of cosmic microwave background anisotropies in closed friedmann-robertson-walker models," The Astrophysical Journal 538, 473–476 (2000).

[71] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, "Large-scale structure of the universe and cosmological perturbation theory," Physics Reports 367, 1–248 (2002).

[72] G. Gamow, "The origin of elements and the separation of galaxies," Phys. Rev. 74, 505 (1948).

[73] A. A. Penzias and R. W. Wilson, "A Measurement of Excess Antenna Temperature at 4080 Mc/s." The Astrophysics Journal 142, 419 (1965).

[74] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, and E. L. Wright, "The cosmic microwave background spectrum from the fullcobefiras data set," The Astrophysical Journal 473, 576–587 (1996).

[75] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson, "Structure in the COBE Differential Microwave Radiometer First-Year Maps," The Astrophysical Journal Letters 396, L1 (1992).

[76] E. L. Wright, S. S. Meyer, C. L. Bennett, N. W. Boggess, E. S. Cheng, M. G. Hauser, A. Kogut, C. Lineweaver, J. C. Mather, G. F. Smoot, R. Weiss, S. Gulkis, G. Hinshaw, M. Janssen, T. Kelsall, P. M. Lubin, J. Moseley, S. H., T. L. Murdock, R. A. Shafer, R. F. Silverberg, and D. T. Wilkinson, "Interpretation of the Cosmic Microwave Background Radiation Anisotropy Detected by the COBE Differential Microwave Radiometer," The Astrophysical Journal Letters 396, L13 (1992).

[77] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, and et al., "Nine-year wilkinson microwave anisotropy probe (wmap) observations: Final maps and results," The Astrophysical Journal Supplement Series 208, 20 (2013).

[78] Planck Collaboration, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, J. Carron, B. Casaponsa, A. Challinor, H. C. Chiang, L. P. L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, J. Delabrouille, J. M. Delouis, F. X. Désert, E. Di Valentino, C. Dickinson, J. M. Diego, S. Donzelli, O. Doré, M. Douspis, A. Ducout, X. Dupac, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, E. Falgarone, Y. Fantaye, J. Fergusson, R. Fernandez-Cobos, F. Finelli, F. Forastieri, M. Frailis, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gruppuso, J. E. Gudmundsson, J. Hamann, W. Hand ley, F. K. Hansen, G. Helou, D. Herranz, E. Hivon, Z. Huang, A. H. Jaffe, W. C. Jones, A. Karakci, E. Keihänen, R. Keskitalo, K. Kiiveri, J. Kim, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, J. M. Lamarre,

M. Langer, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, J. P. Leahy, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Lilley, V. Lindholm, M. López-Caniego, P. M. Lubin, Y. Z. Ma, J. F. Macías-Pérez, G. Maggio, D. Maino, N. Mand olesi, A. Mangilli, A. Marcos-Caballero, M. Maris, P. G. Martin, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. D. Meerburg, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M. A. Miville-Deschênes, D. Molinari, A. Moneti, L. Montier, G. Morgante, A. Moss, S. Mottet, M. Münchmeyer, P. Natoli, H. U. Nørgaard-Nielsen, C. A. Oxborrow, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, T. J. Pearson, M. Peel, H. V. Peiris, F. Perrotta, V. Pettorino, F. Piacentini, L. Polastri, G. Polenta, J. L. Puget, J. P. Rachen, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, C. Rosset, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, E. P. S. Shellard, M. Shiraishi, C. Sirignano, G. Sirri, L. D. Spencer, R. Sunyaev, A. S. Suur-Uski, J. A. Tauber, D. Tavagnacco, M. Tenti, L. Terenzi, L. Toffolatti, M. Tomasi, T. Trombetti, J. Valiviita, B. Van Tent, L. Vibert, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K. Wehus, M. White, S. D. M. White, A. Zacchei, and A. Zonca, "Planck 2018 results. I. Overview and the cosmological legacy of Planck," arXiv e-prints , arXiv:1807.06205 (2018), arXiv:1807.06205 [astro-ph.CO].

[79] R. K. Sachs and A. M. Wolfe, "Perturbations of a Cosmological Model and Angular Variations of the Microwave Background," The Astrophysical Journal 147, 73 (1967).

[80] D. J. Eisenstein, "Dark energy and cosmic sound [review article]," New Astronomy Reviews 49, 360 (2005).

[81] J. Silk, "Fluctuations in the Primordial Fireball," Nature 215, 1155 (1967).

[82] J. M. Kovac, E. M. Leitch, C. Pryke, J. E. Carlstrom, N. W. Halverson, and W. L. Holzapfel, "Detection of polarization in the cosmic microwave background using dasi," Nature 420, 772–787 (2002).

[83] N. Planck Collaboration, Aghanim *et al.*, "Planck 2018 results. VI. Cosmological parameters," ArXiv e-prints (2018), arXiv:1807.06209.

[84] M. Bartelmann and P. Schneider, "Weak gravitational lensing," Physics Reports 340, 291–472 (2001).

[85] W. Hu and M. Tegmark, "Weak lensing: Prospects for measuring cosmological parameters," The Astrophysical Journal 514, L65–L68 (1999).

[86] D. N. Limber, "The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field." The Astrophysical Journal 117, 134 (1953).

[87] S. Cole and G. Efstathiou, "Gravitational lensing of fluctuations in the microwave background radiation," Monthly Notices of the Royal Astronomical Society 239, 195 (1989).

[88] U. Seljak, "Gravitational lensing effect on cosmic microwave background anisotropies: A power spectrum approach," The Astrophysical Journal 463, 1 (1996).

[89] W. Hu, "Weak lensing of the cmb: A harmonic approach," Physical Review D 62 (2000), 10.1103/physrevd.62.043007.

[90] A. LEWIS and A. CHALLINOR, "Weak gravitational lensing of the cmb," Physics Reports 429, 1–65 (2006).

[91] P. J. E. Peebles, "Statistical Analysis of Catalogs of Extragalactic Objects. I. Theory," The Astrophysical Journal 185, 413 (1973).

[92] E. J. Groth and P. J. E. Peebles, "Statistical analysis of catalogs of extragalactic objects. VII. Two- and three-point correlation functions for the high-resolution Shane-Wirtanen catalog of galaxies." The Astrophysical Journal 217, 385 (1977).

[93] D. J. Baumgart and J. N. Fry, "Fourier Spectra of Three-dimensional Data," The Astrophysical Journal 375, 25 (1991).

[94] A. F. Heavens and A. N. Taylor, "A spherical harmonic analysis of redshift space," Monthly Notices of the Royal Astronomical Society 275, 483–497 (1995).

[95] A. Raccanelli, F. Montanari, D. Bertacca, O. Doré, and R. Durrer, "Cosmological measurements with general relativistic galaxy correlations," Journal of Cosmology and Astroparticle Physics 2016, 009–009 (2016).

[96] H. Seo and D. J. Eisenstein, "Improved forecasts for the baryon acoustic oscillations and cosmological distance scale," The Astrophysical Journal 665, 14–24 (2007).

[97] S. Casas, M. Kunz, M. Martinelli, and V. Pettorino, "Linear and non-linear Modified Gravity forecasts with future surveys," Phys. Dark Univ. 18, 73 (2017), arXiv:1703.01271 [astro-ph.CO].

[98] A. J. Ross, W. J. Percival, and M. Manera, "The Information Content of Anisotropic Baryon Acoustic Oscillation Scale Measurements," Mon. Not. Roy. Astron. Soc. 451, 1331 (2015), arXiv:1501.05571 [astro-ph.CO].

[99] V. Desjacques, D. Jeong, and F. Schmidt, "Large-scale galaxy bias," Physics Reports 733, 1–193 (2018).

[100] N. Kaiser, "Clustering in real space and in redshift space," Monthly Notices of the Royal Astronomical Society 227, 1 (1987).

[101] C. Alcock and B. Paczynski, "An evolution free test for non-zero cosmological constant," Nature 281, 358 (1979).

[102] W. E. Ballinger, J. A. Peacock, and A. F. Heavens, "Measuring the cosmological constant with redshift surveys," Monthly Notices of the Royal Astronomical Society 282, 877 (1996), arXiv:astro-ph/9605017 [astro-ph].

[103] T. Delubac *et al.* (BOSS), "Baryon acoustic oscillations in the Lyα forest of BOSS DR11 quasars," Astron. Astrophys. 574, A59 (2015), arXiv:1404.1801 [astro-ph.CO].

[104] A. Font-Ribera *et al.* (BOSS), "Quasar-Lyman α Forest Cross-Correlation from BOSS DR11 : Baryon Acoustic Oscillations," JCAP 1405, 027 (2014), arXiv:1311.1767 [astro-ph.CO].

[105] E. Hubble, "A relation between distance and radial velocity among extra-galactic nebulae," Proceedings of the National Academy of Sciences 15, 168 (1929), https://www.pnas.org/content/15/3/168.full.pdf.

[106] D. Camarena and V. Marra, "A new method to build the (inverse) distance ladder," (2019), arXiv:1910.14125 [astro-ph.CO].

[107] J. Guy, P. Astier, S. Nobili, N. Regnault, and R. Pain, "SALT: a spectral adaptive light curve template for type Ia supernovae," Astronomy and Astrophysics 443, 781 (2005), arXiv:astro-ph/0506583 [astro-ph].

[108] L. Wang, G. Goldhaber, G. Aldering, and S. Perlmutter, "Multicolor Light Curves of Type Ia Supernovae on the Color-Magnitude Diagram: A Novel Step toward More Precise Distance and Extinction Estimates," The Astrophysical Journal 590, 944 (2003), arXiv:astro-ph/0302341 [astro-ph].

[109] A. G. Riess, W. H. Press, and R. P. Kirshner, "A Precise Distance Indicator: Type IA Supernova Multicolor Light-Curve Shapes," The Astrophysical Journal 473, 88 (1996), arXiv:astro-ph/9604143 [astro-ph].

[110] M. M. Phillips, "The Absolute Magnitudes of Type IA Supernovae," The Astrophysical Journal Letters 413, L105 (1993).

[111] A. G. Riess, S. Casertano, W. Yuan, L. M. Macri, and D. Scolnic, "Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond ΛCDM," The Astrophysical Journal 876, 85 (2019), arXiv:1903.07603 [astro-ph.CO].

[112] G. Pietrzyński, D. Graczyk, A. Gallenne, W. Gieren, I. B. Thompson, B. Pilecki, P. Karczmarek, M. Górski, K. Suchomska, M. Taormina, and et al., "A distance to the large magellanic cloud that is precise to one per cent," Nature 567, 200–203 (2019).

[113] M. J. Reid, D. W. Pesce, and A. G. Riess, "An improved distance to ngc 4258 and its implications for the hubble constant," The Astrophysical Journal 886, L27 (2019).

[114] R. Jimenez and A. Loeb, "Constraining cosmological parameters based on relative galaxy ages," The Astrophysical Journal 573, 37–42 (2002).

[115] M. Moresco *et al.*, "Improved constraints on the expansion rate of the Universe up to z 1.1 from the spectroscopic evolution of cosmic chronometers," JCAP 1208, 006 (2012), arXiv:1201.3609 [astro-ph.CO].

[116] M. Moresco, "Raising the bar: new constraints on the Hubble parameter with cosmic chronometers at $z \sim 2$," Mon. Not. Roy. Astron. Soc. 450, L16 (2015), arXiv:1503.01116 [astro-ph.CO].

[117] M. Moresco, L. Pozzetti, A. Cimatti, R. Jimenez, C. Maraston, L. Verde, D. Thomas, A. Citro, R. Tojeiro, and D. Wilkinson, "A 6% measurement of the Hubble parameter at z˜0.45: direct evidence of the epoch of cosmic re-acceleration," Journal of Cosmology and Astroparticle Physics 5, 014 (2016), arXiv:1601.01701.

[118] D. Stern, R. Jimenez, L. Verde, *et al.*, "Cosmic Chronometers: Constraining the Equation of State of Dark Energy. II. A Spectroscopic Catalog of Red Galaxies in Galaxy Clusters," (2009), 10.1088/0067-0049/188/1/280, arXiv:0907.3152.

[119] J. Simon, L. Verde, and R. Jimenez, "Constraints on the redshift dependence of the dark energy potential," Physics Review D 71 (2004), 10.1103/PhysRevD.71.123001, arXiv:0412269 [astro-ph].

[120] C. Zhang, H. Zhang, S. Yuan, S. Liu, T.-J. Zhang, and Y.-C. Sun, "Four New Observational $H(z)$ Data From Luminous Red Galaxies of Sloan Digital Sky Survey Data Release Seven," (2012), 10.1088/1674–4527/14/10/002, arXiv:1207.4541.

[121] H. Yu, B. Ratra, and F.-Y. Wang, "Hubble Parameter and Baryon Acoustic Oscillation Measurement Constraints on the Hubble Constant, the Deviation from the Spatially-Flat Lambda cdm Model, The Deceleration-Acceleration Transition Redshift, and Spatial Curvature," (2017), arXiv:1711.03437.

[122] A. Gómez-Valent and L. Amendola, "H0from cosmic chronometers and type ia supernovae, with gaussian processes and the novel weighted polynomial regression method," Journal of Cosmology and Astroparticle Physics 2018, 051–051 (2018).

[123] W. L. Freedman, B. F. Madore, T. Hoyt, I. S. Jang, R. Beaton, M. G. Lee, A. Monson, J. Neeley, and J. Rich, "Calibration of the tip of the red giant branch (trgb)," (2020), arXiv:2002.01550 [astro-ph.GA].

[124] W. Yuan, A. G. Riess, L. M. Macri, S. Casertano, and D. M. Scolnic, "Consistent calibration of the tip of the red giant branch in the large magellanic cloud on the hubble space telescope photometric system and a redetermination of the hubble constant," The Astrophysical Journal 886, 61 (2019).

[125] W. L. Freedman, B. F. Madore, D. Hatt, T. J. Hoyt, I. S. Jang, R. L. Beaton, C. R. Burns, M. G. Lee, A. J. Monson, J. R. Neeley, and et al., "The carnegie-chicago hubble program. viii. an independent determination of the hubble constant based on the tip of the red giant branch," The Astrophysical Journal 882, 34 (2019).

[126] A. G. Riess, "The expansion of the universe is faster than expected," Nature Reviews Physics 2, 10–12 (2019).

[127] H.-Y. Wu and D. Huterer, "Sample variance in the local measurements of the hubble constant," Monthly Notices of the Royal Astronomical Society 471, 4946–4955 (2017).

[128] B. L. Hoscheit and A. J. Barger, "Large Local Void, Supernovae Type Ia, and the Kinematic Sunyaev-Zel'dovich Effect in a Lambda-LTB Model," in *American*

Astronomical Society Meeting Abstracts #230, American Astronomical Society Meeting Abstracts, Vol. 230 (2017) p. 314.05.

[129] W. D. Kenworthy, D. Scolnic, and A. Riess, "The local perspective on the hubble tension: Local structure does not impact measurement of the hubble constant," The Astrophysical Journal 875, 145 (2019).

[130] M. Smith, D. J. Bacon, R. C. Nichol, H. Campbell, C. Clarkson, R. Maartens, C. B. D'Andrea, B. A. Bassett, D. Cinabro, D. A. Finley, and et al., "The effect of weak lensing on distance estimates from supernovae," The Astrophysical Journal 780, 24 (2013).

[131] N. Canac, G. Aslanyan, K. N. Abazajian, R. Easther, and L. C. Price, "Testing for new physics: neutrinos and the primordial power spectrum," Journal of Cosmology and Astroparticle Physics 2016, 022–022 (2016).

[132] J. L. Bernal, L. Verde, and A. G. Riess, "The trouble withh0," Journal of Cosmology and Astroparticle Physics 2016, 019–019 (2016).

[133] E. Mörtsell and S. Dhawan, "Does the hubble constant tension call for new physics?" Journal of Cosmology and Astroparticle Physics 2018, 025–025 (2018).

[134] C. Heymans, E. Grocutt, A. Heavens, M. Kilbinger, T. D. Kitching, F. Simpson, J. Benjamin, T. Erben, H. Hildebrandt, H. Hoekstra, Y. Mellier, L. Miller, L. Van Waerbeke, M. L. Brown, J. Coupon, L. Fu, J. Harnois-Déraps, M. J. Hudson, K. Kuijken, B. Rowe, T. Schrabback, E. Semboloni, S. Vafaei, and M. Velander, "CFHTLenS tomographic weak lensing cosmological parameter constraints: Mitigating the impact of intrinsic galaxy alignments," Monthly Notices of the Royal Astronomical Society 432, 2433 (2013), https://academic.oup.com/mnras/article-pdf/432/3/2433/12627095/stt601.pdf.

[135] P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al., "Planck2013 results. xx. cosmology from sunyaev–zeldovich cluster counts," Astronomy and Astrophysics 571, A20 (2014).

[136] F. Köhlinger, M. Viola, B. Joachimi, H. Hoekstra, E. van Uitert, H. Hildebrandt, A. Choi, T. Erben, C. Heymans, S. Joudaki, and et al., "Kids-450: the tomographic weak lensing power spectrum and constraints on cosmological parameters," Monthly Notices of the Royal Astronomical Society 471, 4412–4435 (2017).

[137] E. van Uitert *et al.*, "KiDS+GAMA: Cosmology constraints from a joint analysis of cosmic shear, galaxy-galaxy lensing and angular clustering," (2017), arXiv:1706.05004.

[138] A. Amon, C. Blake, C. Heymans, C. D. Leonard, M. Asgari, M. Bilicki, A. Choi, T. Erben, K. Glazebrook, J. Harnois-Déraps, and et al., "Kids+2dflens+gama: testing the cosmological model with the eg statistic," Monthly Notices of the Royal Astronomical Society 479, 3422–3437 (2018).

[139] T. Abbott, F. Abdalla, A. Alarcon, J. Aleksić, S. Allam, S. Allen, A. Amara, J. Annis, J. Asorey, S. Avila, and et al., "Dark energy survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing," Physical Review D 98 (2018), 10.1103/physrevd.98.043526.

[140] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, *et al.*, "Planck 2015 results. XIII. Cosmological parameters," Astronomy and Astrophysics 594, A13 (2016), arXiv:1502.01589.

[141] S. Joudaki, H. Hildebrandt, D. Traykova, N. E. Chisari, C. Heymans, A. Kannawadi, K. Kuijken, A. H. Wright, M. Asgari, T. Erben, H. Hoekstra, B. Joachimi, L. Miller, T. Tröster, and J. L. van den Busch, "KiDS+VIKING-450 and DES-Y1 combined: Cosmology with cosmic shear," arXiv e-prints , arXiv:1906.09262 (2019), arXiv:1906.09262 [astro-ph.CO].

[142] A. De Felice, T. Kobayashi, and S. Tsujikawa, "Effective gravitational couplings for cosmological perturbations in the most general scalar–tensor theories with second-order field equations," Physics Letters B 706, 123 (2011).

[143] A. D. Felice and S. Tsujikawa, "Conditions for the cosmological viability of the most general scalar-tensor theories and their applications to extended Galileon dark energy models," Journal of Cosmology and Astroparticle Physics 2012, 007 (2012).

[144] F. Könnig *et al.*, "Stable and unstable cosmological models in bimetric massive gravity," Physical Review D 90, 124014 (2014), arXiv:1407.4331.

[145] M. Motta *et al.*, "Probing dark energy through scale dependence," Physical Review D 88, 124035 (2013).

[146] C. D. Leonard, P. G. Ferreira, and C. Heymans, "Testing gravity with E_G: mapping theory onto observations," Journal of Cosmology and Astroparticle Physics (2015), 10.1088/1475-7516/2015/12/051, arXiv:1510.04287.

[147] S. de la Torre *et al.*, "The VIMOS Public Extragalactic Redshift Survey (VIPERS). Gravity test from the combination of redshift-space distortions and galaxy-galaxy lensing at $0.5 < z < 1.2$," Astron. Astrophys. 608, A44 (2017), arXiv:1612.05647 [astro-ph.CO].

[148] P. Zhang, M. Liguori, R. Bean, and S. Dodelson, "Probing Gravity at Cosmological Scales by Measurements which Test the Relationship between Gravitational Lensing and Matter Overdensity," Physical Review Letters 99, 141302 (2007).

[149] A. G. Riess, S. Casertano, W. Yuan, L. Macri, J. Anderson, J. W. MacKenty, J. B. Bowers, K. I. Clubb, A. V. Filippenko, D. O. Jones, and B. E. Tucker, "New Parallaxes of Galactic Cepheids from Spatially Scanning the Hubble Space Telescope: Implications for the Hubble Constant," The Astrophysical Journal 855, 136 (2018), arXiv:1801.01120 [astro-ph.SR].

[150] A. Gómez-Valent and L. Amendola, "H0 from cosmic chronometers and Type Ia supernovae, with Gaussian Processes and the novel Weighted Polynomial Regression method," (2018), arXiv:1802.01505.

[151] C. Blake, S. Joudaki, C. Heymans, *et al.*, "RCSLenS: Testing gravitational physics through the cross-correlation of weak lensing and large-scale structure," Monthly Notices of the Royal Astronomical Society (2016), 10.1093/mnras/stv2875, arXiv:1507.03086.

[152] F. Beutler, M. Blake, C.and Colless, *et al.*, "The 6dF Galaxy Survey: z=0 measurements of the growth rate and σ8," Monthly Notices of the Royal Astronomical Society 423, 3430 (2012).

[153] T. Okumura, C. Hikage, T. Totani, *et al.*, "The Subaru FMOS galaxy redshift survey (FastSound). IV. New constraint on gravity theory from redshift space distortions at $z = 1.4$," Publications of the Astronomical Society of Japan, Volume 68, Issue 3, id.38 24 pp. 68 (2015), 10.1093/pasj/psw029, arXiv:1511.08083.

[154] Y.-S. Song and W. J. Percival, "Reconstructing the history of structure formation using redshift distortions," Journal of Cosmology and Astroparticle Physics (2008), 10.1088/1475-7516/2009/10/004, arXiv:0807.0810.

[155] A. J. Hawken, B. R. Granett, A. Iovino, *et al.*, "The VIMOS Public Extragalactic Redshift Survey: Measuring the growth rate of structure around cosmic voids," Astronomy & Astrophysics 607 (2016), 10.1051/0004-6361/201629678, arXiv:1611.07046.

[156] S. de la Torre, L. Guzzo, J. A. Peacock, *et al.*, "The VIMOS Public Extragalactic Redshift Survey (VIPERS). Galaxy clustering and redshift-space distortions at z=0.8 in the first data release," Astronomy & Astrophysics, Volume 557, id.A54, 19 pp. 557 (2013), 10.1051/0004-6361/201321463, arXiv:1303.2622.

[157] F. G. Mohammad, B. R. Granett, L. Guzzo, *et al.*, "The VIMOS Public Extragalactic Redshift Survey (VIPERS): An unbiased estimate of the growth rate of structure at z = 0.85 using the clustering of luminous blue galaxies," (2017), arXiv:1708.00026.

[158] C. Howlett, A. J. Ross, L. Samushia, W. J. Percival, and M. Manera, "The clustering of the SDSS main galaxy sample – II. Mock galaxy catalogues and a measurement of the growth of structure from redshift space distortions at z = 0.15," Monthly Notices of the Royal Astronomical Society 449, 848 (2015).

[159] L. Samushia, W. J. Percival, and A. Raccanelli, "Interpreting large-scale redshift-space distortion measurements," Monthly Notices of the Royal Astronomical Society 420, 2102 (2012).

[160] R. Tojeiro *et al.*, "The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: measuring structure growth using passive galaxies," Mon. Not. Roy. Astron. Soc. 424, 2339 (2012), arXiv:1203.6565 [astro-ph.CO].

[161] C.-H. Chuang and Y. Wang, "Modelling the anisotropic two-point galaxy correlation function on small scales and single-probe measurements of H(z), $D_A(z)$ and $f(z)\sigma_8(z)$ from the Sloan Digital Sky Survey DR7 luminous red galaxies," Monthly Notices of the Royal Astronomical Society 435, 255 (2013), arXiv:1209.0210.

[162] H. Gil-Marín *et al.*, "The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: RSD measurement from the LOS-dependent power spectrum of DR12 BOSS galaxies," Mon. Not. Roy. Astron. Soc. 460, 4188 (2016), arXiv:1509.06386 [astro-ph.CO].

[163] H. Gil-Marín, W. Percival, A. Cuesta, *et al.*, "The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: BAO measurement from the LOS-dependent power spectrum of DR12 BOSS galaxies," Monthly Notices of the Royal Astronomical Society 460, 4210 (2016).

[164] C. Chuang, F. Prada, M. Pellejero-Ibanez, *et al.*, "The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: single-probe measurements from CMASS anisotropic galaxy clustering," Monthly Notices of the Royal Astronomical Society 461 (2016).

[165] A. Cabré and E. Gaztañaga, "Clustering of luminous red galaxies - I. Large-scale redshift-space distortions," Monthly Notices of the Royal Astronomical Society 393, 1183 (2009).

[166] L. Guzzo, M. Pierleoni, B. Meneux, *et al.*, "A test of the nature of cosmic acceleration using galaxy redshift distortions," Nature 451, 541 (2008).

[167] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning* (MIT Press, 2006).

[168] M. Seikel, C. Clarkson, and M. Smith, "Reconstruction of dark energy and expansion dynamics using Gaussian processes," Journal of Cosmology and Astroparticle Physics (2012), 10.1088/1475-7516/2012/06/036, arXiv:1204.2832.

[169] F. Melia and M. K. Yennapureddy, "Model Selection Using Cosmic Chronometers with Gaussian Processes," (2018), arXiv:1802.02255.

[170] T. Holsclaw, U. Alam, B. Sanso, H. Lee, K. Heitmann, S. Habib, and D. Higdon, "Nonparametric Reconstruction of the Dark Energy Equation of State from Diverse Data Sets," Physical Review D 84, 083501 (2011), arXiv:1104.2041.

[171] C. Mignone and M. Bartelmann, "Model-independent determination of the cosmic expansion rate," Astronomy and Astrophysics 481, 295–303 (2008).

[172] S. Benitez-Herrera, F. Röpke, W. Hillebrandt, C. Mignone, M. Bartelmann, and J. Weller, "Model-independent reconstruction of the expansion history of the universe from type ia supernovae," Monthly Notices of the Royal Astronomical Society 419, 513–521 (2011).

[173] S. Benitez-Herrera, E. E. O. Ishida, M. Maturi, W. Hillebrandt, M. Bartelmann, and F. Röpke, "Cosmological parameter estimation from sn ia data: a model-independent approach," Monthly Notices of the Royal Astronomical Society 436, 854–858 (2013).

[174] E. Jullo *et al.*, "Testing gravity with galaxy-galaxy lensing and redshift-space distortions using CFHT-Stripe 82, CFHTLenS and BOSS CMASS datasets," Astron. Astrophys. 627, A137 (2019), arXiv:1903.07160 [astro-ph.CO].

[175] M. Kunz, "Degeneracy between the dark components resulting from the fact that gravity only measures the total energy-momentum tensor," Physical Review D 80, 123001 (2009).

[176] R. Reyes, R. Mandelbaum, U. Seljak, T. Baldauf, J. E. Gunn, L. Lombriser, and R. E. Smith, "Confirmation of general relativity on large scales from weak lensing and galaxy velocities," Nature 464, 256 (2010), arXiv:1003.2185 [astro-ph.CO].

[177] L. Amendola, E. Menegoni, C. Di Porto, M. Corsi, and E. Branchini, "Constraints on a scale-dependent bias from galaxy clustering," Phys. Rev. D95, 023505 (2017), arXiv:1502.03994 [astro-ph.CO].

[178] L. e. a. Amendola, "Cosmology and Fundamental Physics with the Euclid Satellite," ArXiv e-prints (2016), arXiv:1606.00180.

[179] R. A. Fisher, "The logic of inductive inference," JRSS 98, 39 (1935).

[180] R. Trotta, "Bayesian methods in cosmology," Lecture notes for the 44th Saas Fee Advanced Course on Astronomy and Astrophysics, "Cosmology with wide-field surveys" (2017), 1701.01467.

[181] L. Wolz, M. Kilbinger, J. Weller, and T. Giannantonio, "On the validity of cosmological fisher matrix forecasts," JCAP 09, 9 (2012).

[182] R. G. Crittenden, G.-B. Zhao, L. Pogosian, L. Samushia, and X. Zhang, "Fables of reconstruction: controlling bias in the dark energy equation of state," JCAP 2012, 048 (2012), 1112.1693.

[183] F. Elsner and B. D. Wandelt, "Fast calculation of the fisher matrix for cosmic microwave background experiments," A+A 540, L6 (2012), 1202.4898.

[184] S. Khedekar and S. Majumdar, "Cosmology with the largest galaxy cluster surveys: going beyond fisher matrix forecasts," JCAP 2, 30 (2013).

[185] M. Tegmark, A. Taylor, and A. Heavens, "Karhunen-loeve eigenvalue problems in cosmology: how should we tackle large data sets?" ApJ 480, 22 (1997), astro-ph/9603021.

[186] D. Coe, "Fisher matrices and confidence ellipses: a quick-start guide and software," ArXiv 0906.4123 (2009).

[187] B. A. Bassett, Y. Fantaye, R. Hlozek, and J. Kotze, "Fisher4cast users' manual," ArXiv 0906.0993 (2009).

[188] B. M. Schäfer and R. Reischke, "Describing variations of the fisher-matrix across parameter space," MNRAS 460, 3398 (2016), 1603.03626.

[189] M. Loverde, L. Hui, and E. Gaztañaga, "Magnification-temperature correlation: The dark side of integrated sachs-wolfe measurements," PRD 75, 043519 (2007).

[190] N. Taburet, N. Aghanim, M. Douspis, and M. Langer, "Biases on the cosmological parameters and thermal Sunyaev-Zel'dovich residuals," MNRAS 392, 1153 (2009).

[191] A. Amara and A. Réfrégier, "Systematic bias in cosmic shear: extending the fisher matrix," MNRAS 391, 228 (2008).

[192] B. M. Schaefer, M. Douspis, and N. Aghanim, "Implications of bias evolution on measurements of the integrated sachs-wolfe effect: errors and biases in parameter estimation," MNRAS 397, 925 (2009).

[193] B. M. Schäfer, A. F. Kalovidouris, and L. Heisenberg, "Parameter estimation biases due to contributions from the rees-sciama effect to the integrated sachs-wolfe spectrum," MNRAS 416, 1302 (2011).

[194] D. Kirk, I. Laszlo, S. Bridle, and R. Bean, "Optimising cosmic shear surveys to measure modifications to gravity on cosmic scales," MNRAS 430, 197 (2013), 1109.4536.

[195] C. R. Jenkins and J. A. Peacock, "The power of bayesian evidence in astronomy," MNRAS 413, 2895 (2011), 1101.4822.

[196] M. Kerscher and J. Weller, "On model selection in cosmology," SciPost Physics Lecture Notes 9 (2019), 1901.07726.

[197] R. Trotta, "Applications of bayesian model selection to cosmological parameters," MNRAS 378, 72 (2007).

[198] R. Trotta, "Bayes in the sky: Bayesian inference and model selection in cosmology," Contemporary Physics 49, 71 (2008), 0803.4089.

[199] B. Santos, N. C. Devi, and J. S. Alcaniz, "Bayesian comparison of non-standard cosmologies using type ia supernovae and BAO data," PRD 95, 123514 (2017), 1603.06563.

[200] A. Liddle, P. Mukherjee, and D. Parkinson, "Model selection in cosmology," Astronomy and Geophysics 47, 040000 (2006).

[201] A. R. Liddle, P. Mukherjee, D. Parkinson, and Y. Wang, "Present and future evidence for evolving dark energy," PRD 74, 123506 (2006), astro-ph/0610126.

[202] P. Mukherjee, D. Parkinson, P. S. Corasaniti, A. R. Liddle, and M. Kunz, "Model selection as a science driver for dark energy surveys: Model selection and dark energy surveys," MNRAS 369, 1725 (2006).

[203] A. F. Heavens, T. D. Kitching, and L. Verde, "On model selection forecasting, dark energy and modified gravity," MNRAS 380, 1029 (2007).

[204] K. H. Knuth, M. Habeck, N. K. Malakar, A. M. Mubeen, and B. Placek, "Bayesian evidence and model selection," Digital Signal Processing Special Issue in Honour of William J. (Bill) Fitzgerald, 47, 50 (2015).

[205] J. Carron, A. Amara, and S. Lilly, "Probe combination in large galaxy surveys : Application of fisher information and shannon entropy to weak lensing," MNRAS 417, 1938 (2011).

[206] S. Grandis, S. Seehars, A. Refregier, A. Amara, and A. Nicola, "Information gains from cosmological probes," JCAP 2016, 034 (2016), 1510.06422.

[207] S. Ferraro, B. D. Sherwin, and D. N. Spergel, "WISE measurement of the integrated Sachs-Wolfe effect," PRD 91, 083533 (2015).

[208] M. Chevallier and D. Polarski, "Accelerating Universes with Scaling Dark Matter," International Journal of Modern Physics D 10, 213 (2001), arXiv:gr-qc/0009008 [gr-qc].

[209] E. V. Linder, "Exploring the Expansion History of the Universe," PRL 90, 091301 (2003), arXiv:astro-ph/0208512 [astro-ph].

[210] R. Laureijs, J. Amiaux, S. Arduini, J. . Augueres, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet, and et al., "Euclid Definition Study Report," ArXiv e-prints 1110.3193 (2011), arXiv:1110.3193 [astro-ph.CO].

[211] S. Nesseris and J. Garcia-Bellido, "Is the jeffreys' scale a reliable tool for bayesian model comparison in cosmology?" JCAP 2013, 036 (2013), 1210.7652.

[212] C. E. Shannon, "A mathematical theory of communication," The Bell System Technical Journal 27, 379 (1948).

[213] A. Rényi, "On measures of information and entropy," Proceedings of the fourth Berkeley Symposium o Mathematics, Statistics and Probability , 547 (1960).

[214] L. Golshani and E. Pasha, "Rényi entropy rate for gaussian processes," Information Sciences 180, 1486 (2010).

[215] D. L. Wallace, "Asymptotic approximations to distributions," Annals of Mathematical Statistics 29, 635 (1958).

[216] E. Sellentin, M. Quartin, and L. Amendola, "Breaking the spell of gaussianity: forecasting with higher order fisher matrices," Monthly Notices of the Royal Astronomical Society 441, 1831 (2014).

[217] E. Sellentin, "A fast, always positive definite and normalizable approximation of non-gaussian likelihoods," Monthly Notices of the Royal Astronomical Society 453, 893 (2015).

[218] A. K. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," Bull. Calcutta Math. Soc. 35, 99 (1943).

[219] S. Kullback and R. A. Leibler, "On information and sufficiency," Ann. Math. Statist. 22, 79 (1951).

[220] R. Reischke, A. S. Mancini, B. M. Schäfer, and P. M. Merkel, "Investigating scalar-tensor gravity with statistics of the cosmic large-scale structure," MNRAS 482, 3274 (2019), arXiv:1804.02441 [astro-ph.CO].

[221] T. D. Kitching, A. F. Heavens, J. Alsing, T. Erben, C. Heymans, H. Hildebrandt, H. Hoekstra, A. Jaffe, A. Kiessling, Y. Mellier, and others, "3d cosmic shear: cosmology from CFHTLenS," Monthly Notices of the Royal Astronomical Society 442, 1326 (2014).

[222] A. Nicola, A. Refregier, and A. Amara, "Integrated approach to cosmology: Combining CMB, large-scale structure and weak lensing," PRD 94, 083517 (2016), 1607.01014.

[223] P. M. Merkel and B. M. Schaefer, "Parameter constraints from weak lensing tomography of galaxy shapes and cosmic microwave background fluctuations," MNRAS 469, 2760 (2017), 1707.08153.

[224] S. Hilbert, J. Hartlap, and P. Schneider, "Cosmic shear covariance: The log-normal approximation," A+A 536, A85 (2011), 1105.3980.

[225] I. Kayo, M. Takada, and B. Jain, "Information content of weak lensing power spectrum and bispectrum: including the non-gaussian error covariance matrix," MNRAS 429, 344 (2013).

[226] E. Krause and T. Eifler, "CosmoLike - cosmological likelihood analyses for photometric galaxy surveys," MNRAS 470, 2100 (2017), 1601.05779.

[227] T. Tao and V. Vu, "Random covariance matrices: Universality of local statistics of eigenvalues," The Annals of Probability 40, 1285 (2012).

[228] D. J. Paz and A. G. Sanchez, "Improving the precision matrix for precision cosmology," MNRAS 454, 4326 (2015), 1508.03162.

[229] R. Reischke, A. Kiessling, and B. M. Schäfer, "Variations of cosmic large-scale structure covariance matrices across parameter space," MNRAS 465, 4016 (2016), 1607.03136.

[230] A. Taylor and B. Joachimi, "Estimating cosmological parameter covariance," MNRAS 442, 2728 (2014).

[231] E. Sellentin and A. F. Heavens, "Parameter inference with estimated covariance matrices," MNRAS 456, L132 (2016), 1511.05969.

[232] E. Sellentin and A. F. Heavens, "Quantifying lost information due to covariance matrix estimation in parameter inference," MNRAS 464, 4658 (2016), 1609.00504.

[233] E. Sellentin and A. F. Heavens, "On the insufficiency of arbitrarily precise covariance matrices," MNRAS 473, 2355 (2018), 1707.04488.

[234] L. Knox, "Determination of inflationary observables by cosmic microwave background anisotropy experiments," PRD 52, 4307 (1995).

[235] R. J. Thornton, P. A. R. Ade, S. Aiola, F. E. Angilè, M. Amiri, J. A. Beall, D. T. Becker, H.-M. Cho, S. K. Choi, P. Corlies, K. P. Coughlin, R. Datta, M. J. Devlin, S. R. Dicker, R. Dünner, J. W. Fowler, A. E. Fox, P. A. Gallardo, J. Gao, E. Grace, M. Halpern, M. Hasselfield, S. W. Henderson, G. C. Hilton, A. D. Hincks, S. P. Ho, J. Hubmayr, K. D. Irwin, J. Klein, B. Koopman, D. Li, T. Louis, M. Lungu, L. Maurin, J. McMahon, C. D. Munson, S. Naess, F. Nati, L. Newburgh, J. Nibarger, M. D. Niemack, P. Niraula, M. R. Nolta, L. A. Page, C. G. Pappas, A. Schillaci, B. L. Schmitt, N. Sehgal, J. L. Sievers, S. M. Simon, S. T. Staggs, C. Tucker, M. Uehara, J. van Lanen, J. T. Ward, and E. J. Wollack, "The Atacama Cosmology Telescope: The Polarization-sensitive ACTPol Instrument," ApJS 227, 21 (2016).

[236] D. N. Limber, "The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field. II." ApJ 119, 655 (1954).

[237] M. Bartelmann and P. Schneider, "Weak gravitational lensing," Physics Reports 340, 291 (2001).

[238] H. Hoekstra and B. Jain, "Weak Gravitational Lensing and Its Cosmological Applications," Annual Review of Nuclear and Particle Science 58, 99 (2008).

[239] D. J. Baumgart and J. N. Fry, "Fourier spectra of three-dimensional data," ApJ 375, 25 (1991).

[240] H. A. Feldman, N. Kaiser, and J. A. Peacock, "Power-spectrum analysis of three-dimensional redshift surveys," ApJ 426, 23 (1994).

[241] A. F. Heavens and A. N. Taylor, "A spherical harmonic analysis of redshift space," MNRAS 275, 483 (1995).

[242] V. Desjacques, D. Jeong, and F. Schmidt, "Large-scale galaxy bias," Physics Reports 733, 1 (2018).

[243] C. M. Hirata and U. Seljak, "Reconstruction of lensing from the cosmic microwave background polarization," PRD 68, 083002 (2003).

[244] A. Lewis and A. Challinor, "Weak gravitational lensing of the CMB," Physics Reports 429, 1 (2006).

[245] R. K. Sachs and A. M. Wolfe, "Perturbations of a Cosmological Model and Angular Variations of the Microwave Background," ApJ 147, 73 (1967).

[246] W. Hu and T. Okamoto, "Mass Reconstruction with Cosmic Microwave Background Polarization," ApJ 574, 566 (2002).

[247] T. Okamoto and W. Hu, "Cosmic microwave background lensing reconstruction on the full sky," Phys. Rev. D 67, 083002 (2003).

[248] R. Reischke and B. M. Schäfer, "Realistic systematic biases induced by residual intrinsic alignments in cosmic shear surveys," ArXiv 1910.05994 , arXiv:1910.05994 (2019), arXiv:1910.05994 [astro-ph.CO].

[249] B. M. Schäfer and P. M. Merkel, "Angular spectra of the intrinsic galaxy ellipticity field, their observability and their impact on lensing in tomographic surveys," arXiv e-prints , arXiv:1506.07366 (2015), arXiv:1506.07366 [astro-ph.CO].

[250] A. G. Riess, S. Casertano, W. Yuan, L. M. Macri, and D. Scolnic, "Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond ΛCDM," ApJ 876, 85 (2019), arXiv:1903.07603 [astro-ph.CO].

[251] H. Hildebrandt, M. Viola, C. Heymans, S. Joudaki, K. Kuijken, C. Blake, T. Erben, B. Joachimi, D. Klaes, L. Miller, C. B. Morrison, R. Nakajima, G. Verdoes Kleijn, A. Amon, A. Choi, G. Covone, J. T. A. de Jong, A. Dvornik, I. Fenech Conti, A. Grado, J. Harnois-Déraps, R. Herbonnet, H. Hoekstra, F. Köhlinger, J. McFarland, A. Mead, J. Merten, N. Napolitano, J. A. Peacock, M. Radovich, P. Schneider, P. Simon, E. A. Valentijn, J. L. van den Busch, E. van Uitert, and L. Van Waerbeke, "Kids-450: cosmological parameter constraints from tomographic weak gravitational lensing," MNRAS 465, 1454 (2017), arXiv:1606.05338 [astro-ph.CO].

[252] I. Olkin and F. Pukelsheim, "The distance between two random vectors with given dispersion matrices," Linear Algebra and its Applications 48, 257 (1982).

[253] D. C. Dowson and B. V. Landau, "The fréchet distance between multivariate normal distributions," Journal of Multivariate Analysis 12, 450 (1982).