UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE FÍSICA

BAYESIAN STATISTICAL METHODS ON LARGE SCALE STRUCTURE
COSMOLOGY

TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIAS, MENCIÓN FÍSICA

BRYAN LESTER SAGREDO BRIONES

PROFESOR GUÍA:
DOMENICO SAPONE

MIEMBROS DE LA COMISIÓN:
ROLANDO DÜNNER PLANELLA
NELSON PADILLA
GONZALO PALMA QUILODRÁN

SANTIAGO DE CHILE
2018

Resumen

En esta tesis se introduce el formalismo de la estadística Bayesiana aplicado a cosmología ΛCDM y otros modelos de energía oscura, enfocado en los observables de agrupamiento de galaxias y de crecimiento cósmico. Se explora un gran rango de aplicaciones del marco de trabajo Bayesiano.

Primero, se exploran las posibilidades de predicción de la estadística Bayesiana con nuestro primer proyecto, el cual consiste en la aplicación del método de Aproximación de Verosimilitudes con Derivadas (DALI por sus siglas en inglés) para mejorar las predicciones de Matriz de Fisher de un experimento de agrupamiento de galaxias tipo LSST. El método contiene una expansión de Taylor hasta el tercer orden a partir del punto de parámetros de confianza, capturando formas de las regiones de confianza que van más allá de las usuales elipses de Fisher en la bibliografía. Además se compara con muestras de Cadenas de Markov Monte Carlo para mostrar la efectividad del método.

Luego, realizamos un proyecto acerca de la aplicación del formalismo de la Robustez Interna a una compilación de datos de crecimiento cósmico, el cual es un método Bayesiano que potencialmente puede detectar outliers (datos aislados), errores sistemáticos o nuevas leyes físicas en los datos, considerando la posibilidad de que subconjuntos de los datos sigan diferentes parámetros o modelos (incluyendo de esta manera el aspecto de comparación de modelos básico de la estadística Bayesiana). No se encuentran errores sistemáticos ni outliers en el set de datos, así asegurando su robustez interna.

Finalmente, tomamos por completo el campo de comparación de modelos Bayasiana, y lo hacemos via un estudio acerca de diferentes métodos de comparación de modelos cosmológicos. Se comparan varios modelos de energía oscura usando datos crecimiento cósmico y expansión cósmica, y esto se hace utilizando cuatro criterios de comparación: Comparación de evidencias, Criterio de Información Bayesiano, Criterio de Información de Akaike y un método reciente de Figura de Mérito. Luego, se discute acerca de la efectividad y conveniencia de cada uno de ellos.

Abstract

In this thesis, we introduce the Bayesian statistics formalism applied to ΛCDM cosmology and other dark energy models, focused on galaxy clustering and cosmic growth observables. We explore a wide range of applications of the Bayesian framework.

First, we explore the forecasting possibilities of Bayesian statistics with our first project, which consists in the application of the Derivative Approximation of LIkelihoods (DALI) method to improve Fisher Matrix forecasts of an LSST-like galaxy clustering survey. The method features a Taylor expansion up to third order from the fiducial parameter point, capturing shapes for the confidence regions that go beyond the usual Fisher ellipses in the bibliography. We also compare with Markov Chain Monte Carlo samples to show the effectiveness of the method.

Next, we take on a project about the application of the Internal Robustness formalism to a compilation of cosmic growth data, which is a Bayesian method that can detect outliers, systematics or potential new physics in data considering the possibility that subsets of the data follow different parameters or models (including in this manner the model comparison basics of Bayesian statistics). We found no systematics nor outliers in the dataset, thus ensuring its internal robustness.

Finally, we fully take on the Bayesian model comparison field, and we perform this via a study about contrasting different comparison methods of cosmological models. We rank several dark energy models using cosmic growth and cosmic expansion data, and we do this by utilizing four different comparison criteria: Evidence comparison, Bayesian Information Criterion, Akaike Information Criteria and a recent Figure of Merit method. We also develop a fifth method that builds on the Figure of Merit method. We then discuss the effectiveness and convenience of each of them.

That might be an action that's been repeated endlessly in any era, in any world. Humans always pursue things they can't reach. Doesn't it single-heartedly symbolize such yearnings?

The legend ends, and history begins.

Agradecimientos

Este trabajo no habría sido posible sin la llegada de Domenico, mi profesor guía, a Chile, donde ofreció un tema de investigación en donde me pude desempeñar con mucho entusiasmo y quedé totalmente prendado con la investigación. Desde aquel momento siempre he contado con su cercanía, ligereza, experiencia y, de vez en cuando, con su disciplina. Muchas gracias por permitirme conocer el mundo mientras aprendía cosmología y ciencia de datos.

Agradezco profundamente el apoyo y enorme expertiz de Savvas Nesseris en mi estadía en la Universidad Autónoma de Madrid en donde se desarrolló el capítulo 6 de esta tesis.

Es importante mencionar y agradecer el fundamental trabajo de Javier Silva al sumarse al proyecto del capítulo 7, el cual ayudó a sacar adelante con su trabajo duro.

Un gran saludo a mis amigos de la facultad, como los institutanos, los de la sección 7 y todos con quien he compartido en mi paso por el lugar. Saludos a la Salita de Astro y Física y toda su comunidad.

Agradezco mucho haber participado de los magnos eventos de las CosmoPizza y CosmoBeer. Mención honrosa a mis amigos del tercer piso del edificio de física. Será difícil la hora del café después de almuerzo sin ustedes.

Muchas gracias a las academias de física y astronomía del Instituto Nacional, que me dieron el puntapié para querer estudiar algo tan desquiciado como es la física.

No es posible no agradecer a mi familia: mamá, papá, Ronald, Alexander, Daniela por apoyarme siempre en mis proyectos de vida tan extraños. También quiero agradecer a mi segunda familia: Celia y Andrés, quienes me han inspirado mucho estos años con su cercanía. Sin el apoyo de todos ustedes no podría estar hoy escribiendo esto.

Es fundamental agradecer de todo corazón a mi compañera Bárbara, mi mayor pilar en los años que llevo en la facultad, quien siempre me ha apoyado en los momentos difíciles y acompañado a celebrar en los menos difíciles (en los difíciles también). En particular le agradezco por la última etapa de la confección de esta tesis, en donde su apoyo fue crucial para seguir adelante cada día.

Finalmente, muchas gracias a ti, querido lector, por invertir tu tiempo en aprender de mi trabajo *(Finally, thank you very much, reader, for investing your time on learning about my work)*.

Contents

List of Tables

List of Figures

Introduction

Albert Einstein published his theory of gravity, General Relativity, in 1915 [4], changing completely the way we understood the Universe. This led to a better understanding of relativistic astrophysical objects like black holes, neutron stars, active galactic nuclei, phenomena like the Doppler effect, gravitational waves, and gravitational lensing, just to name a few. Especially, gravity shapes the large-scale structure of the Universe through the gravitational interaction of objects like galaxies, clustering up and forming patterns known as the *cosmic web*.

Several experimental tests have confirmed General Relativity to be the correct theory of gravity [5]. Modern cosmology then began as one of the first applications of General Relativity[1]. Another important cornerstone of our current cosmovision is the cosmological principle, which states that our Universe is homogeneous and isotropic on large scales. Such a hypothesis has been validated by many experiments, all which are related to the observed distribution of galaxies in our Universe, and it has been crucial at setting the stage to develop a cosmological perturbation theory, which describes the overall growth and dynamics of the components of the Universe. Another important part of our current understanding of the Universe was the discovery of its expansion by Edwin Hubble [7], pointing that objects separate between each other.

The ΛCDM model, also called the concordance model, is nowadays the standard model that it is able to describe the dynamics of the Universe, as confirmed by several observations [8, 9]. It is governed by the laws of General Relativity, the cosmological principle, and cosmological perturbation theory. Before carrying on, let us illustrate the current state of the chronology of the Universe.

According to ΛCDM, our Universe began with a hot *big-bang*, a result that comes from reversing the time in an expanding Universe, where physics as we know it breaks down, due to the immense density and energy levels. After that, the Universe experienced a period of fast expansion called *inflation*, which is parametrized as a field-driven process which is not explained by the standard model of particle physics. After the inflationary epoch, the Universe is dominated by radiation and, as the Universe expanded and cooled down, light particles like Hydrogen and heavier elements were formed (in processes called *recombination* and *nucleosynthesis*, respectively) and the Universe was then composed of a baryon-photon coupled fluid. As the expansion and cooling continued, eventually, photons decoupled from

---

[1]Even recently, Einstein's theory was confirmed again via gravitational wave detections on 2016 by the LIGO/VIRGO consortia [6].

baryons, and electrons followed afterward, thus liberating photons with valuable information of the state of the Universe at that point, constituting the Cosmic Microwave Background (CMB). The CMB is an (approximately) homogeneous radiation that comes from every direction in the sky, evidencing that ancient period of high energy where baryon and photons were coupled. Quantum fluctuations of the inflaton field correspond to the primordial seeds of perturbations, and their imprints are still preserved on the CMB radiation, these seeds have grown during the matter-dominated era forming the structures we see today. Summarizing, the Universe started as an infinitesimal point in spacetime, experienced a quick expansion through inflation, and then started to cool down and form larger and larger bound objects as it expanded: from sub-atomic particles to clusters of galaxies.

Two important parts of the contemporary cosmological picture are the enigmatical *dark matter* and *dark energy*. They are the distinctive features of ΛCDM, making up for its name: dark energy in the form of a cosmological constant (Λ) plus cold dark matter (CDM).

The need of intruding a dark matter component comes from observations where the matter density of the Universe is not able to describe properly the dynamics of galaxies with respect to their rotational curves. Additionally, large-scale observations indicate that the matter density does not match the observed luminous matter content [10]. Another important probe for dark matter is the CMB, as it also captures the behavior of its perturbations. Dark matter can be thought of as a form of matter that does not interact (or it interacts weakly) with light. Since it is very hard to detect it directly, experiments need to map it only through their gravitational interaction with ordinary matter, examples of this are galaxy cluster dynamics and gravitational lensing. Dark matter constitutes around 25% of the energy content of the Universe, while baryonic matter only makes for a 4% [11], it is clear that Dark matter is the main driver on the formation of structure. Dark matter, however, has its issues, represented mostly by *the small scales problem* [12]. There is a generalized mismatch (that can be subdivided in at least 5 specific issues) between observations versus numerical simulations and semi-analytic models of galaxy formation in a ΛCDM Universe. These issues may lead to the discovery of unknown physics in the dark sector, such as the possibility of interactions between baryons and dark matter. The scales of these phenomena far in the non-linear regime (or too small), for which they cannot be properly described by first-order perturbation theory.

Dark energy was introduced into the budget of our Universe in 1998 since the supernovae of type Ia measurements from Refs. [13, 14]. Such experiments captured an unexpected diminishing in the luminosity of the farther supernovae Ia, implying that they were even farther than expected. This showed that the Universe was in a phase of accelerated expansion. Something like this effect could be explained if we add a component that pushes objects away from each other, challenging the dynamics of the Universe known so far. A fluid with a negative pressure can then fuel the acceleration of the Universe, thus giving place to dark energy in the theoretical framework. This peculiar component of the Universe constitutes about 70% percent of the energy content of the Universe. Introducing the cosmological constant gives rise to a sort of philosophical issue, known as *the fine-tuning problem* [15]. In order to relate the cosmological constant Λ to a more fundamental property of the Universe, it could be seen as the energy of the vacuum. However, making this connection is in an enormous discrepancy with particle physics, which can predict the value of the vacuum energy.

This discordance is so extraordinary that it spans more than 120 orders of magnitude[2]. Nonetheless, this might be an issue of the Standard Model of Particle Physics, rather than ΛCDM.

One of the most important probes of dark energy is the large-scale structure, with experiments that study the clustering properties of galaxies. The distribution of galaxies carries precise information about the Universe since their positions can be described (statistically) by linear cosmological perturbation theory up to a great degree. According to this framework, the small perturbations on the matter density, which come as initial conditions from inflation, evolve along with the expansion of the Universe and eventually form stars, galaxies and clusters of galaxies in a *bottom-up* fashion, a process which may defy common sense. Of course, the positioning of galaxies is intimately related to the clustering of mass, which includes dark matter, and this relation can be parametrized by a *bias factor*, which can in principle depend on time and the properties of galaxies. The large-scale structure is nowadays one of the most constraining probes available. Galaxy clustering surveys like the Sloan Digital Sky Survey [16] have already constrained cosmological parameters with high precision, and will continue to improve over next generations, with experiments like the Large Synoptic Survey Telescope [17] and the Euclid Survey [18], which over the next decade will map the positions of billions of galaxies with unprecedented accuracy and range. Our attention to the large-scale structure cosmology lies in that it might be the key to finally obtain answers to the mysteries of the dark sector, and maybe even the true nature about the origin of the Universe, if it was discovered that any model other than General Relativity dictates the dynamics of the Universe (for both dark energy and modified gravity cases).

Throughout the years, several cosmological models have taken the spotlight in the scientific community, depending on the available data at the time. Historically, simplicity and the number of free parameters have been the cornerstones to find the best description of phenomena. In this sense, ΛCDM excels but, at the same time, it leaves many questions yet to answer, which are related to the fundamental theory of the model itself, as we mentioned earlier. Due to these problems the scientific community was forced to develop a phenomenological description of the models in order to describe their behavior and look for alternatives to ΛCDM, such as models that maintain dark energy, but not involving fluids, such as Quintessence or K-essence. There are also models that modify the type of dark energy fluid, like the Chaplygin gas dark energy or coupled dark energy. We can mention theories that modify the type of spacetime while retaining most of the other properties of ΛCDM, like the Lemaître-Tolman-Bondi (LTB) inhomogeneous metric model. There are also Modified Gravity theories that deviate from General Relativity, and explain the dynamics of the Universe without the need of dark energy, like $f(R)$ gravity, or the Dvali-Gabadadze-Parroti brane-world model (DGP) [19].

Over the last few decades, the field of cosmology has been set through a revolution, driven by the technological improvements in astronomical and cosmological surveys: the sensitivity, efficiency, and detection range has been improved by orders of magnitude with every new generation of surveys, allowing us to observe deeper, further in time and in broader ranges.

---

[2]The value of vacuum energy density predicted by quantum physics is $\rho_\Lambda \approx 10^{71}$ GeV, while the vacuum energy density *as it was produced by the cosmological constant* only amounts to $\rho_\Lambda \approx 10^{-47}$ GeV.

Along with this, the increasing computing power, performance, and throughput[3] available for the scientific community has been giving a hard fight to keep up with the even more increasing size of the data products of each new experiment. Often this is called *the era of precision cosmology*, as surveys aim to constrain the cosmological parameters within percent accuracy, something unprecedented a few decades ago.

When having more precise and numerous data, statistical errors drop and, then, the limiting factor on the amount of information that can be extracted from the measurements will be the systematic errors of surveys, which are reduced by refining observational techniques. Another way of maximizing the obtained information is to improve the statistical techniques used to analyze data. This thesis will focus on the latter topic, where Bayesian statistical methods have been crucial when moving into the era of precision cosmology.

Bayesian statistics has been a fundamental tool for the analysis of cosmological data. It is a versatile framework that is naturally scalable with the complexity of the statistical treatment of the data, and this gets more important as the volume and complexity of data increases. Another factor for the flourishing of Bayesian statistics in cosmology is the increase of the computational capabilities, allowing for the application of many numerical techniques for data analysis that were impossible to implement, say, twenty years ago. These techniques have to be formulated properly under a statistical point of view, what is achieved naturally with the Bayesian formalism. Bayesian statistics' point of view is intuitive to understand: it gives a quantitative measure of our belief in a theory, given a set of physical measures. Of course, it is also important to develop tools that minimize numerical cost while maximizing the amount of output information.

The most general questions one wants to answer with Bayesian statistics are: *Which is the best model to describe observations?* and *which are the parameters of a theory that better describe observations?*. The former question addresses the topic of model selection, whereas the latter addresses parameter inference.

Parameter inference is treated as the search for the confidence regions of the parameters of a theory. An important method within the scope of the latter question is the Fisher Matrix. This method consists in approximating the logarithm of a probability distribution to first order in a Taylor expansion, resulting in a Gaussian distribution for the model's parameters. This tool is less expensive numerically in comparison with the more usual Markov Chain Monte Carlo (MCMC) methods for obtaining confidence regions for parameters. One important application of the Fisher Matrix formalism is forecasting the performance of future surveys by applying the method to a simulated Universe with a fiducial model, as it would be observed by the survey's telescope. These forecasts have the goal of pointing what could be improved in order to optimize each survey, thus maximizing their scientific returns.

---

[3]Data throughput refers to the *amount* of data processed. High Throughput Computing (HTC) is different than High Performance Computing (HPC) in the sense that, in an HTC facility, each machine is not necessarily large or fast, but the vast number of machines, which can take the form of a network of desktop computers, could process more data than an HPC facility. On the other hand, an HPC facility is closely and efficiently connected in the same room, allowing for more communication between nodes. So, the degree of parallelization needed for a given massive computation may be better suited to either HPC or HTC treatment.

The main objective of this thesis is to introduce the Bayesian statistics framework applied to cosmology and use it in three independent projects, focused on the careful use of the Bayesian techniques explained, which are based on all of the three pillars of Bayesian analysis.

Outline

On Chapter 1, we underline the major theoretical background regarding cosmology with the ΛCDM model, including cosmic distances and standard candles and rulers.

Continuing, on Chapter 2 we briefly go through the equations needed for computing cosmological quantities related to cosmological perturbations, in the case of non-interacting matter and for the interacting case, we outline the ideas behind the Boltzmann equation.

Building on the foundations from the previous chapters, Chapter 3 delves on the specific topic of this thesis, the large-scale structure of the Universe, including redshift space distortions and baryon acoustic oscillations, and how we actually observe all of it.

Next, we change the switch to the Bayesian statistics formulations on Chapter 4, where we outline how to extract valuable information from observations, where computational considerations are key.

Beginning our projects, on Chapter 5 we implement the Derivative Approximation of Likelihoods, an extension to the Fisher Matrix method, to forecast an LSST-like galaxy clustering survey, where we also compare to a more computationally expensive Markov Chain Monte Carlo method.

Next, we jump into the model comparison part with the project of Chapter 6, which is the implementation of the Internal Robustness method to a set of growth rate observations, which is a novel test that can be applied to almost any dataset (given that it is not too large), giving new insights into the relation of the dataset's subsets.

Our next project, on Chapter 7 is a *comparison of comparison methods* for different dark energy models (equation-of-state models), using growth rate data and direct measures of the Hubble constant with cosmic chronometers. We contrast four ways of comparing the models, while also proposing a new method that improves on one of the former methods.

Finally, we present our conclusions and outlook.

Chapter 1

Fundamentals of the ΛCDM model

In this chapter, we settle the foundations for our cosmological study, based on the theory of General Relativity, along with the current picture of the cosmological landscape. Throughout this thesis, we will use units so that $c = 1$, being c the speed of light. We will state whenever it is important to come back and use other units. Greek letters index numbers starting from 0 and Latin letters start from 1. Dots will represent derivatives with respect to the time coordinate (except on Chapter 2), additionally, covariant derivatives and partial derivatives of tensors are denoted by placing (next to an index) a semicolon or a comma, respectively. Dots represent derivatives with respect to time. The main reference for this part of the thesis is Ref. [11].

1.1 The Friedmann Equations

The ΛCDM scenario is described by a curvature-free Friedman-Lemaître-Robertson-Walker (FLRW) metric [20, 21, 22, 23]. We now write the more general (including curvature) FLRW spacetime, which describes a homogeneous and isotropic Universe, and its line element is given by

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = -\mathrm{d}t^2 + a^2\left[\frac{\mathrm{d}r^2}{1 - Kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2)\right] . \tag{1.1}$$

Here, $g_{\mu\nu}$ is the metric tensor, $a(t)$ is the scale factor dependent on cosmic time t, and we normalize it to 1 at the present epoch. r, θ, ϕ are spherical comoving coordinates. K represents the curvature of the universe, and it can take the values $-1, 0, +1$, corresponding to an open, flat or closed Universe, respectively.

The Einstein's field equations [24] are given by

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} , \tag{1.2}$$

where $G_{\mu\nu}$ is the Einstein's tensor, describing the geometry of the universe, G is the Newtonian constant, and $T_{\mu\nu}$ is the energy-momentum tensor, accounting for the energy and matter present in the Universe.

The Einstein's tensor is defined through the Ricci tensor and, from it, it can be expressed with the Christoffel symbols, connecting to the geometry of the metric. Now we define these quantities:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R \ , \tag{1.3}$$

$$R_{\mu\nu} = \Gamma^{\alpha}_{\mu\nu,\alpha} - \Gamma^{\alpha}_{\mu\alpha,\nu} + \Gamma^{\alpha}_{\beta\alpha}\Gamma^{\beta}_{\mu\nu} - \Gamma^{\alpha}_{\beta\nu}\Gamma^{\beta}_{\alpha\mu} \ , \tag{1.4}$$

$$\Gamma^{\gamma}_{\alpha\beta} = \frac{1}{2}g^{\gamma\eta}(g_{\alpha\eta,\beta} + g_{\beta\eta,\alpha} - g_{\alpha\beta,\eta}) \ , \tag{1.5}$$

where $R_{\mu\nu}$ is the Ricci tensor, $R = g^{\mu\nu}R_{\mu\nu}$ the Ricci scalar, and $\Gamma^{\gamma}_{\alpha\beta}$ the Christoffel symbol, which finally links the Einstein tensor to the metric $g_{\mu\nu}$.

We will express the energy-momentum tensor considering that the contents in the universe are only perfect fluids with density ρ and pressure p.

$$T_{\mu\nu} = (\rho + p)u_{\mu}u_{\nu} + pg_{\mu\nu} \ . \tag{1.6}$$

Here, $u^{\mu} = (-1, 0, 0, 0)$ is the *background* four-velocity of the fluid in comoving coordinates. By background, we refer to quantities that depend on time only, in order to study the the averaged dynamics. In Chapter 2 we will consider deviations from this paradigm.

Combining the Einstein's field equations with the energy-momentum tensor form above gives the Friedmann equations, see Ref. [11]

$$H^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2} \ , \tag{1.7}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) \ . \tag{1.8}$$

On these equations we define the Hubble parameter as $H = \dot{a}/a$. It will be useful to also define the Hubble constant $H_0 = H(a = 1)$ and the adimensional Hubble constant through $100\,h\,\text{km/s/Mpc} = H_0$. An important caveat on this parameter is its tension, as there are recent observational constraints on this parameter that maintain a considerable tension, as CMB measurements point $h = 0.674 \pm 0.005$ [9] while local measurements (of galaxy distances through Cepheids) obtain $h = 0.7324 \pm 0.0174$ [25].

1.1.1 Background evolution of matter species

Each type of fluid ρ_i can be characterized by its equation of state (EoS) parameter w_i, which is the ratio of its pressure to density.

$$w_i = \frac{p_i}{\rho_i} \ , \tag{1.9}$$

This EoS will tell us how the density evolves respect to the scale factor. With that in mind, the conservation of the energy-momentum tensor, or $T^{0\nu}_{;\nu} = 0$ can help us derive

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0 \ . \tag{1.10}$$

Therefore, in the case of a constant w_i, Eq. (1.10) can be solved as

$$\rho_i(a) = \rho_{i,0} a^{-3(1+w_i)} . \tag{1.11}$$

Here, the subscript '0' indicates present time at $a = 1$ (this will hold for the rest of the thesis). We can also treat the case of a time-dependent (via the scale factor a) EoS $w_i(a)$, in which we replace w_i in Eq. (1.9) with the effective equation of state parameter $\hat{w}_i$, given by

$$\hat{w}_i = \frac{1}{\log(a)} \int_a^1 \frac{w_i(x)}{x} \mathrm{d}x . \tag{1.12}$$

The matter species considered are radiation and matter. We now detail more about them, along with other components that can be described similarly:

Radiation: Radiation comprises relativistic particles like photons and massless neutrinos, these behave with a pressure equal to one-third of its density, a result from statistical mechanics. Their EoS is

$$w_r = 1/3 \Rightarrow \rho_r \propto a^{-4} . \tag{1.13}$$

This decrease in density can alternatively be derived more intuitively by considering how the expansion affects an ensemble of photons: a factor of a^{-3} comes from the change in number density of the particles, and another a^{-1} factor from the stretch of the light wave, making it increase its wavelength, hence decreasing its frequency, resulting in a proportional decrease in energy.

Matter: When we speak loosely of matter, we refer to non-relativistic particles, like baryons, massive neutrinos and cold dark matter (*cold* refers to having small speeds compared to light's). Its equation of state is 0 since its pressure is negligible. The EoS for this species is

$$w_m = 0 \Rightarrow \rho_m \propto a^{-3} . \tag{1.14}$$

Baryonic matter and cold dark matter are often considered separately in analyses, to take into account the epoch of the baryon-photon coupled fluid, while dark matter is already starting to form more dense regions. The baryon-photon coupled era lasted until the until the *drag epoch*, this is the time when baryons were released from the drag of photons via Compton scattering. Do not confuse the drag epoch with the recombination time, as the latter is the time when photons were released from the Thomson scattering interactions with electrons, which is the point where the CMB was released, a moment that also defines the *last scattering surface*.

Dark energy: Dark energy can be treated in several ways, as we mentioned in the Introduction. For now, we will consider it as a fluid with an equation of state $w_{\mathrm{de}} = -1$ in the ΛCDM paradigm. This comes from the addition of a cosmological constant Λ to the Einstein's field equation, in order to explain the current acceleration of the Universe

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \quad \rightarrow \quad G_{\mu\nu} + g_{\mu\nu}\Lambda = 8\pi G T_{\mu\nu} . \tag{1.15}$$

This addition modifies the Friedmann equation as follows

$$H^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2} + \frac{\Lambda}{3} . \tag{1.16}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3} . \tag{1.17}$$

Considering this cosmological constant as a fluid, one can return to the original Friedmann Eqs. (1.7),(1.8) by defining a constant density $\rho_{\mathrm{de}} = \Lambda/(8\pi G)$ and, noticing it is independent of the scale factor, we define an effective EoS as

$$w_{\mathrm{de}} = -1 \ . \tag{1.18}$$

The most simple extensions to ΛCDM are modifications to this EoS, namely the wCDM model, that has a free parameter $w = w_{\mathrm{de}}$, and the CPL model [26, 27] also known as $w_0 w_a$CDM, which has a time-dependent dark energy EoS parametrized by $w_{\mathrm{de}}(a) = w_0 + w_a(1 - a)$ (notice this is just a Taylor expansion over $a = 1$ to first order). This time-dependency has to be treated with Eq.1.12 in order to include it in the density function.

An interesting issue to mention here is *the coincidence problem* [28]. This problem addresses why the energy density parameters of dark energy and matter are so similar, as they are of the same order today, which demands that the densities of these species *had to be set* to an incredibly precise value in the early stages of the Universe, considering that the matter density evolves with a power law while the dark energy density is a constant. Intents on how to answer this question often need the anthropic principle, hence converting this problem into a philosophical issue.

Curvature: The curvature term in the Friedmann equation (Eq. (1.7)) can also be treated as a matter species, by defining the density

$$\rho_k = \frac{-3K}{8\pi G a^2} \Rightarrow \rho_k \propto a^{-2} \ . \tag{1.19}$$

Therefore, it can be considered as a fluid with an effective equation of state

$$w_k = -1/3 \ . \tag{1.20}$$

Considering only these species to conform our Universe, then Eq.1.7 can be recast in a straightforward expression,

$$\frac{H^2}{H_0^2} = \Omega_{r,0} a^{-4} + \Omega_{m,0} a^{-3} + \Omega_{k,0} a^{-2} + \Omega_{\mathrm{de},0} \ . \tag{1.21}$$

Here, we introduced the density parameters $\Omega_{i,0}$, given by

$$\Omega_{i,0} = \frac{\rho_{i,0}}{\rho_{\mathrm{crit},0}} = \frac{8\pi G \rho_i}{3 H_0^2} \ . \tag{1.22}$$

With $\rho_{i,0}$ the density of the species at present time, and $\rho_{\mathrm{crit},0} = 3H_0^2/8\pi G$ the critical density of the Universe at present time, which corresponds to the current average density of it. These density parameters are very convenient since they are constants and they also fulfill

$$\Omega_{r,0} + \Omega_{m,0} + \Omega_{k,0} + \Omega_{\mathrm{de},0} = 1 \ . \tag{1.23}$$

Nonetheless, the density parameters can be further generalized to their time-dependant versions, Ω_i, which are expressed by

$$\Omega_i = \frac{\rho_i(a)}{\rho_{crit}(a)} = \frac{H_0^2}{H(a)^2} \Omega_{i,0} a^{-3(1+w_i)} \ , \tag{1.24}$$

where $\rho_{\rm crit} = 3H^2/8\pi G$ the critical density of the Universe at a given scale factor. Thus, these quantities follow, at all times,

$$1 = \Omega_r + \Omega_m + \Omega_k + \Omega_{\rm de} \,. \tag{1.25}$$

The different power laws that the species follow can define domination eras for each of them: The radiation-dominated era (starting after inflation), the matter-dominated era and the dark energy-dominated era, which were mentioned on the Introduction. Currently, we live in the dark energy era, and only recently, since it started at around ~ 10 billion years of Universe age, out of the actual ~ 14 billion years old. Most of the Universe's history has been within the matter domination era, since the Universe was ~ 14000 years old when it started, with the electron-photon decoupling time happening afterward at an age of ~ 378000 years [10].

The current values for the density parameters are, according to [9]:

$$\Omega_{m,0} = 0.3111 \pm 0.0056 \,, \tag{1.26}$$
$$\Omega_{\rm de,0} = 0.6889 \pm 0.0056 \,, \tag{1.27}$$
$$\Omega_{k,0} = 0 \qquad \text{(Assumed)} \,, \tag{1.28}$$
$$\Omega_{k,0} = 0.0007 \pm 0.0019 \qquad \text{(If allowed)} \,. \tag{1.29}$$

As the radiation-dominated era is so far in the history of the Universe, the energy density of radiation has decreased significantly, due to its EoS. Consequently, the density parameter of radiation, which is valued at present time, is very small in comparison to the other parameters, namely $\Omega_{r,0} \approx 8.4 \cdot 10^{-5}$ [10].

1.2 Cosmic distances

Distances are fundamental to portrait how the Universe is configured, therefore in this section, we will illustrate how distances are defined and observed, plus some important related concepts. In this section we mainly follow Ref.[29].

1.2.1 Redshift

The redshift is an important phenomenon of the light that cosmological objects emit, crucial to define distances in the Universe. This comes from the fact that objects in an expanding Universe receed from us, thus experiencing a *cosmological Doppler effect*. In specific, electromagnetic waves stretch during their travel from its emission to us, evidencing the expansion of the space component of an FLRW Universe. The redshift, symbolized by z, is a way of measuring the Doppler effect. It is defined by the wavelengths at emission and reception, namely $\lambda_{\rm e}$ and λ_0, respectively,

$$z = \frac{\lambda_0 - \lambda_{\rm e}}{\lambda_{\rm e}} \,. \tag{1.30}$$

10

If we note that the wavelength is proportional to the scale factor, as any other distance, we arrive at a simple relation for the scale factor and the redshift,

$$\frac{\lambda_e}{\lambda_0} = \frac{a(t_e)}{a(t_0)} \quad \Rightarrow \quad a = \frac{1}{1+z} \tag{1.31}$$

Here, t_e and t_0 represent the times of emission and reception of the wave, respectively. We remind the reader that our convention for the current scale factor is $a(t_0) = 1$.

As the wavelength increases, the effect on visible light is that it gets shifted to the color red, hence the name of the phenomenon. This effect can be measured by spectroscopy: comparing the absorption or emission lines of the spectrum of light with samples of the same chemical composition in the lab. Thus, it is possible to obtain the scale factor by using the redshift of the observed object, opening a window into the past of our Universe.

1.2.2 Comoving distance

The comoving distance d_c is related to the geodesic path of a photon traveling between two points in the same line of sight. To define the comoving distance, we express the FLRW metric (Eq. (1.1)) again, now including the speed of light c,

$$ds^2 = -c^2 dt^2 + a^2 \left[d\chi^2 + r^2(\chi)(d\theta^2 + \sin^2\theta d\phi^2) \right]. \tag{1.32}$$

Here, χ relates to the old coordinate $r(\chi)$ and on the curvature K:

$$r(\chi) = \begin{cases} \sin\chi & (K = +1), \\ \chi & (K = 0), \\ \sinh\chi & (K = -1). \end{cases} \tag{1.33}$$

This expression can also be compressed trough

$$r(\chi) = \frac{1}{\sqrt{-K}} \sinh\left(\chi\sqrt{-K}\right). \tag{1.34}$$

If we consider an object at χ_1, the comoving distance d_c between us, at χ_0, and the object will be determined by the geodesic $ds = 0$, since photons must travel in null geodesics. This equation transforms into $cdt = a(t)d\chi$ when considering a trajectory free of angular components. The definition is as follows:

$$d_c \equiv \chi_1 = \int_0^{\chi_1} d\chi = -c \int_0^{t_1} \frac{dt}{a(t)}. \tag{1.35}$$

Now, if we use $da = H(a)a dt$ and $da = -dz/(1+z)^2$, we get

$$d_c = -c \int_1^{a_1} \frac{da}{H(a)a^2} = c \int_0^{z_1} \frac{dz}{H(z)}. \tag{1.36}$$

Finally, with $E(z) \equiv H(z)/H_0$ we arrive at:

$$d_c(z) = \frac{c}{H_0} \int_0^z \frac{dz'}{E(z')} \, .$$

(1.37)

The quantity c/H_0 is known as the *Hubble distance*, and its scale is in the order of the radius of the Universe. Similarly, $1/H_0$ is the *Hubble time* and is in the scale of the age of the Universe.

1.2.3 Transverse comoving distance

A closely related quantity to the comoving distance is the transverse comoving distance d_M, also known as proper motion distance. If two objects at the same redshift are separated by an angle θ, they are separated by $d_M \cdot \theta$. This is the coordinate $r(\chi)$ from Eq. (1.32). In other words,

$$d_M(z) = \frac{1}{\sqrt{-K}} \sinh\left(d_c(z)\sqrt{-K} \right).$$

(1.38)

Using $\Omega_{k,0} = -Kc^2/H_0^2$ instead of K, this can be rearranged in the more useful form:

$$d_M(z) = \frac{c\, H_0^{-1}}{\sqrt{-\Omega_{k,0}}} \sin\left(\sqrt{-\Omega_{k,0}} \int_0^z \frac{dz'}{E(z')} \right),$$

(1.39)

which, in the case of a flat cosmology, reduces to the comoving distance:

$$d_M(z) = \frac{c}{H_0} \int_0^z \frac{dz'}{E(z')} \, .$$

(1.40)

1.2.4 Luminosity distance and standard candles

Luminosity can be used as a way of measuring distances. We start by considering a source object with an intrinsic luminosity L_e (we use again 'e' for emission and '0' for the present time/arrival), and the flux (energy over surface) we observe is F_0, which defines an observed luminosity L_0 via $F_0 = L_0/A_0$, where A_0 is the area of a sphere at $a = 1$.

The luminosity distance is defined trough

$$d_L \equiv \sqrt{\frac{L_e}{4\pi F_0}} \, .$$

(1.41)

This means that we can obtain the distance of an object if we know its intrinsic luminosity and observe its flux. If we consider the area in a curved spacetime to be related to the transverse comoving distance via $A_0 = 4\pi d_M^2$, we will only need to compute the ratio of luminosities,

$$d_L = d_M \sqrt{\frac{L_e}{L_0}} \, .$$

(1.42)

The luminosity of an object is the energy emitted over time, so the cosmological Doppler effect stretches the wavelength of the light emitted, decreasing the energy emitted per photon. Similar to the redshift case explained earlier, this induces a ratio

$$\frac{E_e}{E_0} = \frac{\lambda_0}{\lambda_e} = \frac{1}{a} ,$$

(1.43)

where E and λ is the energy and wavelength of a photon, and a is the scale factor at the time of emission. This also affects the period Δt of the wave, since the stretch of the wavelength is proportional to it, or

$$\frac{\Delta t_e}{\Delta t_0} = \frac{\lambda_0}{\lambda_e} = a .$$

(1.44)

Finally, writing the luminosity as $L = E/\Delta t$ yields

$$d_L(z) = (1+z)d_M(z) ,$$

(1.45)

where we included the redshift via $a = 1/(1+z)$.

Supernovae Ia (SN Ia) [30] are a subtype of supernovae that is produced in a binary system where a white dwarf is matter-fed by an accreting companion star, until the Chandrasekhar limit of mass is surpassed and an explosion (of nuclear nature) occurs, meaning that the electron-degenerated pressure of the white dwarf was surpassed by the gravity of it. SN Ia are important in cosmology, thanks to their *standard candle* property, meaning that all supernovae Ia have an almost equal luminosity (or it can be standardized using other of their properties like light-curve peak, stretch, and color [31]), which allows them to place strong cosmological constraints, especially on dark energy. In particular, they drove the discovery of the accelerated expansion of the Universe [13, 14].

The luminosity distance is obtained in SN Ia measurements, where the distance modulus μ of the star is the observed quantity, and it is linked to the luminosity distance through

$$\mu = m - M = 5\log_{10}(D_L) + \mu_0$$

(1.46)

with $D_L = (H_0/c)d_L$ the adimensional luminosity distance and μ_0 a nuisance parameter, which we assume equal for all SN Ia. The distance modulus μ is the difference between the magnitude m of the star (a logarithmic measure of the brightness), and the absolute magnitude M, which corresponds to the magnitude of the star if it was located 10 pc away from the observer.

1.2.5   Angular diameter distance and standard rulers

Observing an object that subtends a certain angle can be used as a way of measuring its distance if we know the size of the object. The angular diameter distance is defined as the ratio of the physical size of an object to the angular size that it is observed to subtend, i.e.

$$d_A \equiv \frac{\Delta x}{\Delta \theta} ,$$

(1.47)

with Δx the diameter of the object, orthogonal to the line of sight, and $\Delta\theta$ its subtended angle. It is related to the transverse comoving distance via

$$d_A(z) = \frac{d_M(z)}{(1+z)} \; . \tag{1.48}$$

The explanation comes from the factor $\Delta\theta \sim 1/a = (1+z)$ that arises when moving from the observed angular separation to the actual separation in the current redshift.

The angular diameter distance is of great importance in cosmology, because it is used to measure the Baryon Acoustic Oscillations (BAO) scale r_s, which is the sound horizon at the *baryon drag epoch*, a point in time which is close to the end of recombination. So, it can be defined by

$$r_s = \int\limits_0^{\tau_{\mathrm{drag}}} c_s(\tau)\mathrm{d}\tau \; . \tag{1.49}$$

where τ is the conformal time, defined through $\mathrm{d}\tau = \mathrm{d}t/a$ (τ_{drag} evaluated at the drag epoch) and c_s the sound speed of the baryon-photon fluid.

The drag scale r_s represents the maximum (comoving) scale at which sound waves could propagate before the drag epoch. Before the drag epoch, sound waves propagated back and forth in the fluid, hence the 'oscillations' part of BAO. So, different parts of the baryon-photon fluid could not communicate through sound in distances larger than r_s. If the cosmological principle is correct, this phenomenon must have left a statistical imprint in the CMB and the large-scale structure of galaxies which has been found. The importance of r_s is explained by it being a *standard ruler*, not in the sense of an object of constant size, but rather a 'constant comoving-sized ruler', a ruler which is also obtained via the statistics of galaxies, as we will see later.

Then, by using Eqs. (1.48) and (1.36) while observing transverse and parallel comoving scales $r_\perp$ and $r_\parallel$ through an angle $\Delta\theta$ and a small redshift range Δz, respectively, one can relate these quantities [32]

$$d_A(z) = \frac{r_\perp}{(1+z)\Delta\theta} \; , \tag{1.50}$$

$$H(z) = c\,\frac{\Delta z}{r_\parallel} \; . \tag{1.51}$$

In order to link these quantities to the BAO case, we can simply set both scales to the sound horizon, $r_\perp = r_\parallel = r_s$. This last move may be an oversimplification, since modern BAO surveys like SDSS-IV [33] use more complex models to obtain relations between $r_\perp$, $r_\parallel$ and r_s. These models are based on techniques involving numerical simulations of the 2-point correlation function of galaxies (we treat the 2-point correlation function on Chapter 3). The current constraint on the drag scale, $r_s = 147.18 \pm 0.29$ Mpc, was obtained via CMB observations [9], therefore its standard ruler property is fully exploitable to enhance the quality of BAO constraints.

Chapter 2

A slice of cosmological perturbation theory

Perturbation theory is the treatment of small perturbations (inhomogeneities) in a background homogeneous Universe, and how these primordial seeds evolve following the laws of General Relativity.

In this thesis, we mainly focus on the statistical analysis of observed quantities like the galaxy power spectrum or the growth rate of matter perturbations and, it is for that reason that we have to settle the theoretical foundations for these quantities to be constructed, which are based on perturbation theory.

In order to confront data to theory (or make forecasts out of a theory), we must solve the perturbation differential equations (2.31)-(2.34) plus (2.37)-(2.38) or the Boltzmann equation (2.52), be it analytically, numerically, or using a solver like CLASS [34, 35] or CAMB [36, 37]. That said, those equations are the main results from this chapter. Now, we delve into the mathematical formalism behind those equations.

Linear perturbation theory is a nice way to obtain equations out of a complex theory like General Relativity, and it works very well at large scales, but care must be taken when going to the smaller scales, where first order perturbations alone are insufficient to describe the dynamics of astrophysical objects, and a non-linear treatment or semi-analytical models must be sought.

Before we begin, we state our preference to work in Fourier space and, since we are doing first order perturbations, all the equations are linear and therefore we can drop the $e^{i\mathbf{k}\cdot\mathbf{x}}$ multiplying terms that appear when taking spatial derivatives. It is important to mention that, in this chapter only, the dots represent derivatives with respect to the conformal time τ. Our main references for the chapter are Refs. [11] and [38].

2.1 Choosing the Gauge

There are different ways to define the perturbations in an FLRW metric. This arises from the freedom of choosing the observer for the metric, thanks to the gauge-invariance of General Relativity. We now introduce the two most used gauges.

2.1.1 Synchronous Gauge

We first define the Synchronous gauge through

$$\mathrm{d}s^2 = a^2[-\mathrm{d}\tau^2 + (\delta_{ij} + h_{ij})\mathrm{d}x^i\mathrm{d}x^j] \,, \tag{2.1}$$

with h_{ij} a tensor for spatial perturbations. The name stands for 'equal time', linked to the fact that, by definition, there are no perturbations in time in this gauge. Boltzmann solvers work using this gauge to maximize computational efficiency, due to the arising of *batch-like* calculations on matrices.

2.1.2 Conformal Newtonian Gauge

We now present the perturbed metric, in the conformal Newtonian gauge, also known as the longitudinal gauge or shear-free gauge.

$$\mathrm{d}s^2 = a^2[-(1 + 2\psi)\mathrm{d}\tau^2 + (1 - 2\phi)\mathrm{d}x_i\mathrm{d}x^i] \,, \tag{2.2}$$

where ψ and ϕ are time and space perturbations[1]. This metric has the advantage of having a diagonal metric, which simplifies the calculations immensely. We are now going to utilize this gauge to demonstrate the standard equations needed to evolve its perturbations. There are, of course, transformation rules to go from one gauge to another whenever needed.

2.2 Perturbing General Relativity

In order to perturb the Einstein's equations (1.2), we consider a variation on both the Einstein tensor and the energy-momentum tensor. This is splitting them into a background and a perturbed part according to

$$G^\mu_\nu = G^{\mu(0)}_\nu + \delta G^\mu_\nu \,, \tag{2.3}$$

$$T^\mu_\nu = T^{\mu(0)}_\nu + \delta T^\mu_\nu \,, \tag{2.4}$$

---

[1]This sign convention is often different in the bibliography on the topic, which affects the following equations. We follow the convention of [38].

where the (0) superscript denotes a background part. Then, the background evolution is obtained by solving the Einstein equation at zeroth order, while the perturbed parts follow the first order equation,

$$G^{\mu(0)}_{\nu} = 8\pi G T^{\mu(0)}_{\nu} \,, \tag{2.5}$$

$$\delta G^{\mu}_{\nu} = 8\pi G \, \delta T^{\mu}_{\nu} \,. \tag{2.6}$$

To obtain the expression for the perturbed Einstein tensor, we propagate the perturbations, starting from the metric through the Christoffel symbols, and then to the Ricci tensor and scalar, finally getting to the Einstein tensor.

2.3 Perturbed Einstein tensor

So, we begin by splitting the metric in a background (FLRW) and perturbed part $g_{\mu\nu} = g^{(0)}_{\mu\nu} + \delta g_{\mu\nu}$, where

$$\delta g_{\mu\nu} = a^2 \begin{pmatrix} -2\psi & 0 \\ 0 & 2\phi\delta_{ij} \end{pmatrix} \,. \tag{2.7}$$

Then, we also write the perturbed Christoffel symbol

$$\delta\Gamma^{\mu}_{\nu\lambda} = \frac{1}{2}\delta g^{\mu\alpha}(g_{\alpha\nu,\lambda} + g_{\alpha\lambda,\nu} - g_{\nu\lambda,\alpha}) + \frac{1}{2}g^{\mu\alpha}(\delta g_{\alpha\nu,\lambda} + \delta g_{\alpha\lambda,\nu} - \delta g_{\nu\lambda,\alpha}) \,. \tag{2.8}$$

Here, the only non-vanishing terms are

$$\delta\Gamma^{0}_{00} = \dot{\psi} \,, \tag{2.9}$$

$$\delta\Gamma^{0}_{ij} = -\delta_{ij}\left[2(\psi + \phi)\mathcal{H} + \dot{\phi}\right] \,, \tag{2.10}$$

$$\delta\Gamma^{0}_{0i} = \delta\Gamma^{0}_{i0} = \delta\Gamma^{i}_{00} = \phi_{,i} \,, \tag{2.11}$$

$$\delta\Gamma^{i}_{j0} = \delta\Gamma^{i}_{0j} = -\delta^{i}_{j}\dot{\phi} \,, \tag{2.12}$$

where we define the conformal Hubble parameter by

$$\mathcal{H} \equiv \frac{\dot{a}}{a} = aH \,. \tag{2.13}$$

Next, the perturbed Ricci tensor and scalar are

$$\delta R_{\mu\nu} = \delta\Gamma^{\alpha}_{\mu\nu,\alpha} - \delta\Gamma^{\alpha}_{\mu\alpha,\nu} + \delta\Gamma^{\alpha}_{\beta\alpha}\Gamma^{\beta}_{\mu\nu} + \Gamma^{\alpha}_{\beta\alpha}\delta\Gamma^{\beta}_{\mu\nu} - \delta\Gamma^{\alpha}_{\beta\nu}\Gamma^{\beta}_{\alpha\mu} - \Gamma^{\alpha}_{\beta\nu}\delta\Gamma^{\beta}_{\alpha\mu} \,, \tag{2.14}$$

$$\delta R = \delta g^{\mu\nu}R_{\mu\nu} + g^{\mu\nu}\delta R_{\mu\nu} \,. \tag{2.15}$$

Then, the perturbed Ricci tensor is

$$\delta R_{00} = -k^2\psi + 3\left[\mathcal{H}(\dot{\psi} + \dot{\phi}) - \ddot{\phi}\right] \,, \tag{2.16}$$

$$\delta R_{ij} = -\phi\left(4k^2 + 12\mathcal{H}^2 + 6\dot{\mathcal{H}}\right) + \psi\left(k^2 - 12\mathcal{H}^2 - 6\dot{\mathcal{H}}\right) - 3\left[\mathcal{H}(5\dot{\phi} + \dot{\psi}) - \ddot{\phi}\right] \,, \tag{2.17}$$

$$\delta R_{0i} = \delta R_{i0} = k\left(\mathcal{H}\psi + \dot{\phi}\right) \,, \tag{2.18}$$

where the appearance of k reminds us that we are working in Fourier space. The Ricci scalar becomes

$$\delta R = -\frac{2}{a^2}\left\{2k^2\phi + \psi(-k^2 + 6\mathcal{H}^2 + 6\dot{\mathcal{H}}) + 3\left[\mathcal{H}(\dot{\psi} + 3\dot{\phi}) + \ddot{\phi}\right]\right\}. \tag{2.19}$$

Continuing, the perturbed Einstein tensors (with different indexations) are

$$\delta G_{\mu\nu} = \delta R_{\mu\nu} - \frac{1}{2}\delta g_{\mu\nu}R - \frac{1}{2}g_{\mu\nu}\delta R, \tag{2.20}$$

$$\delta G^\mu_\nu = \delta g^{\mu\alpha}G_{\alpha\nu} + g^{\mu\alpha}\delta G_{\alpha\nu}. \tag{2.21}$$

Combining all the above, we arrive to the expressions for the perturbed Einstein tensor

$$\delta G^0_0 = \frac{2}{a^2}\left[k^2\phi + 3\mathcal{H}\left(\mathcal{H}\psi + \dot{\phi}\right)\right], \tag{2.22}$$

$$\delta G^0_i = -\delta G^i_0 = -\frac{2k}{a^2}\left(\mathcal{H}\psi + \dot{\phi}\right), \tag{2.23}$$

$$\delta G^i_i = \frac{2}{a^2}\left\{k^2(\phi - \psi) + 3\psi\left(\mathcal{H}^2 + 2\dot{\mathcal{H}}\right) + 3\left[\mathcal{H}\left(2\dot{\phi} + \dot{\psi}\right) + \ddot{\phi}\right]\right\}, \tag{2.24}$$

$$\delta G^i_j = -\frac{k^2}{a^2}(\phi - \psi), \qquad \text{when } i \neq j. \tag{2.25}$$

2.4   Perturbed energy-momentum tensor

2.4.1   Non-interacting fluids

If we consider the Universe to contain only ideal fluids, the energy-momentum tensor equation (1.6) holds. Considering perturbations in density $\delta\rho = \rho - \bar{\rho}$ and pressure $\delta p = p - \bar{p}$, where the bar symbolizes the mean at the time, we obtain

$$\delta T^0_0 = -\delta\rho, \tag{2.26}$$

$$\delta T^i_0 = (\bar{\rho} + \bar{p})v^i, \tag{2.27}$$

$$\delta T^0_i = (\bar{\rho} + \bar{p})v_i, \tag{2.28}$$

$$\delta T^i_j = \delta p\, \delta^i_j + \bar{p}\,\Pi^i_j, \qquad \Pi^i_i = 0, \tag{2.29}$$

where $v^i = \mathrm{d}x/\mathrm{d}\tau$ is the peculiar velocity coming from perturbations in the 4-velocity u^{i2}, which can be treated as a first order quantity, and Π^i_j is the anisotropic stress perturbation traceless tensor, also of first order. We have to take this component into account when the different species interact with each other, in the contrary case it vanishes.

To make more sense of the variables we are using, we can recall the equation of state from Eq. (1.9) $w = p/\rho$, not necessarily a constant now, and define a sound speed through

$$c_s^2 \equiv \frac{\delta p}{\delta\rho}. \tag{2.30}$$

---

[2]Note that the perturbations in the 4-velocity δu^i exist independently only in the spatial coordinates; we tie the time coordinate to the condition $u^\mu u_\mu = 1$.

Thus, we arrive to the set of differential equations needed to be solved [38, 11]

$$k^2\phi + 3\mathcal{H}\left(\dot{\phi} + \mathcal{H}\psi\right) = -4\pi Ga^2\delta\rho \,, \tag{2.31}$$

$$k^2\left(\mathcal{H}\psi + \dot{\phi}\right) = 4\pi Ga^2(1+w)\rho\theta \,, \tag{2.32}$$

$$\ddot{\phi} + \mathcal{H}(\dot{\psi} + 2\dot{\phi}) + (2\frac{\ddot{a}}{a} - \psi\mathcal{H}^2) + \frac{k^3}{3}(\phi - \psi) = \frac{4\pi}{3}Ga^2c_s^2\delta\rho \,, \tag{2.33}$$

$$k^2(\phi - \psi) = 12\pi Ga^2(1+w)\rho\sigma \,, \tag{2.34}$$

with $\theta = ik_iv^i$ the velocity divergence[3] and σ relates to the anisotropic stress via

$$\sigma(\bar{\rho} + \bar{p}) = -\left(\frac{k_ik_j}{k^2} - \frac{1}{3}\delta_{ij}\right)\Pi_j^i \,. \tag{2.35}$$

In this set of equations, it is important to emphasize that, on the right hand side, we have to sum over all the species present in the Universe. Also, many times we consider non-interacting fluids like a two-fluid model including baryons and cold dark matter (they interact through gravity only) and, in the case of matter, we can assume $\sigma = 0$, giving the condition $\phi = \psi$ from Eq. (2.34).

We can include to this system the conservation of the energy-momentum tensor $T^\mu_{\nu;\mu} = 0$, with a semicolon representing a covariant derivative. Then, this equation also has to stand at first order,

$$\delta T^\mu_{\nu;\mu} = 0 = \delta T^\mu_{\nu,\mu} - \delta\Gamma^\alpha_{0\beta}T^\alpha_\beta - \Gamma^\alpha_{0\beta}\delta T^\alpha_\beta + \delta\Gamma^\beta_{\beta\alpha}T^\beta_0 + \Gamma^\beta_{\beta\alpha}\delta T^\beta_0 \,. \tag{2.36}$$

Working this equation (depending on the index ν) leads us to the following:

$$\dot{\delta} = -(1+w)\left(\theta - 3\dot{\phi}\right) - \mathcal{H}\left(\frac{\delta p - w\delta\rho}{\rho}\right) \,, \qquad (\nu = 0) \,, \tag{2.37}$$

$$\dot{\theta} = -\mathcal{H}(1 - 3w)\theta - \frac{\dot{w}}{1+w}\theta + \frac{1}{1+w}\frac{k^2\delta p}{\rho} + k^2(\psi - \sigma) \,, \qquad (\nu = i) \,, \tag{2.38}$$

where $\delta = \delta\rho/\bar{\rho}$.

In order to work analytically with these equations, we should restrict ourselves to special cases, in terms of scales and epochs. We can name two important scales for any fluid. These scales dictate qualitatively the status of perturbations' evolution, namely if they grow, remain constant or even decrease.

The first important scale is the *causal horizon* scale $k_{\text{causal}} = \mathcal{H}$. This scale places a limit where scales larger than it (or $k < k_{\text{causal}}$) cannot grow since this would break causality. Another important scale is the *sound horizon* scale $k_{\text{sound}} = \mathcal{H}/c_s$, and it represents a limit where smaller perturbations cannot grow since they are contained by the pressure of the fluid. We must notice that both horizon scales increase with time as the Universe expands, and since the sound speed is less than unity, the sound horizon scale is always smaller than the causal horizon. This means that, as the Universe expands, perturbations can enter the

---

[3]Sometimes it is useful to replace θ with the variable $V = (1+w)\theta$, since it helps to avoid singularities in Eqs. (2.37) and (2.38), see Ref. [19]. We will do so in Chapter 7.

causal horizon and start evolving. Another possibility is that c_s changes and increases, while diminishing the value of k_{sound}, allowing the evolution of perturbations that were trapped by pressure. One example of this is the passage from the radiation-dominated era to the matter-dominated era, which happened at the radiation-matter equality time, where perturbations of baryons where trapped by the radiation coupling and were free to grow afterward.

Table 2.1: Summary of characteristic scales and its limitations on the evolution of perturbations.

$k \lesssim k_{\text{causal}}$	$k_{\text{causal}} \lesssim k \lesssim k_{\text{sound}}$	$k_{\text{sound}} \lesssim k$
Perts. cannot grow	Perts. can grow	Perts. cannot grow

Another useful result that can be derived from Eqs. (2.37),(2.38) is, as it is commonly said, *baryonic matter falls into dark matter halos*, which means that the dark matter density perturbations source the matter density perturbations. This can be obtained by considering a sub-horizon scales, and a Universe filled only with baryonic matter and cold dark matter (which is approximately the case during the matter-dominated era), and considering that baryonic perturbations are much smaller than CDM perturbations, since the latter do not interact with radiation, so they could grow during the baryon-photon coupling epoch.

As a final remark of the subsection, we remind that the equations derived so far are valid for non-interacting fluids. This means they are a good way to picture the dark Universe through dark energy (if we assume it to be a fluid) and cold dark matter because both of them can be modeled as non-interacting fluids. So, if we choose a starting point within the matter-dominated era (using a well-motivated ansatz for the value of perturbations), we can integrate the equations up to the present time. However, when we need to include other species, as we will need in order to compute the matter power spectrum in the next chapter, we have to resort to the Boltzmann equation.

2.4.2   Mixing interacting species: the Boltzmann equation

When we consider that fluids can interact in other ways than gravity, like the interactions of baryons, photons, and neutrinos (an example being the Thompson scattering for baryons and photons), we must incorporate these effects into the energy-momentum tensor, and we can do so by working on their microscopic physics. We now follow the results from Ref. [38]. In order to be able to consider these effects, we need to solve the Boltzmann equation, Eq. (2.52), which takes into account all the species and their distinct properties, like photons, neutrinos (massive or not), baryons, cold dark matter, and dark energy (in the form of a cosmological constant or other simple EoS modifications where $w \neq -1$), and the possibility of curvature or not.

We start by focusing on the distribution function of the phase space of the particles $f(\mathbf{P}, \mathbf{x}, t)$, with $\mathbf{P}$ and $\mathbf{x}$ the conjugate momentum and the position of the particles, respectively. The conjugate momentum is the spatial part of the 4-momentum with lower indices, which is expressed as

$$P_i = a(1 - \phi)p_i \,, \tag{2.39}$$

where p^i is the physical momentum. In order to obtain more comfortable expressions, we use the variable $\varepsilon \equiv a(p^2 + m^2)^{-1/2}$, with m the mass of the particle. Thus, ε represents the scale factor times the energy of the particle, as measured by a comoving observer. We can also write its relation to P_0, the time coordinate of the 4-momentum (with lower indeces):

$$P_0 = -(1 + \psi)\varepsilon \ . \tag{2.40}$$

In the same fashion, we can define $q_i \equiv ap_i$, and we will refer to it as the comoving 3-momentum. It then follows that $P_i = (1 - \phi)q_i$. We can also split it in magnitude and direction: $q_i = qn_i$. Then, we rewrite $\varepsilon(q, \tau)^2 = q^2 + a^2m^2$. So, for massless particles, ε reduces to q.

We now consider the phase space as a six-dimensional space, which is sliced at an equal conformal time, a procedure which makes it not a strictly covariant quantity, but a very good approximation nonetheless. We also take this distribution as a background quantity plus a perturbation, denoting that by

$$f(x^i, P_j, \tau) = f^{(0)}(q)\left(1 + \Psi(x^i, q, n_j, \tau)\right) \tag{2.41}$$

(be careful to not confuse this phase-space perturbation Ψ with the metric perturbation ψ). We take the background part to be a Fermi-Dirac or a Bose-Einstein distribution, for fermions and bosons, respectively, or

$$f^{(0)}(\varepsilon(q, \tau)) = \frac{g_s}{h_P^3}\,\frac{1}{\exp\left(\frac{\varepsilon}{k_B T_0}\right) \pm 1}\ , \tag{2.42}$$

where g_s is the number of spin degrees of the particle ($g_s = 2s + 1$ with a spin s), h_P the Planck constant, k_B the Boltzmann constant, and $T_0 = aT$ is the temperature of the particles at present time. We take the plus $(+)$ sign for fermions and the minus $(-)$ sign for bosons. Then, energy-momentum tensor is obtained by computing

$$T_{\mu\nu} = \int dP_1 dP_2 dP_3 \sqrt{-g}\frac{P_\mu P_\nu}{P^0} f(x^i, P_j, \tau)\ , \tag{2.43}$$

with g the determinant of the metric. To linear order, we can obtain

$$\sqrt{-g} = a^{-4}(1 - \psi + 3\phi)\ , \tag{2.44}$$

$$dP_1 dP_2 dP_3 = (1 - 3\phi)q^2 dq d\Omega\ , \tag{2.45}$$

where $d\Omega$ is the solid angle differential, in reference to n_i. Continuing the calculations, using Eqs. (2.43)-(2.45) along with the plugging of the n_i and q_i variables, we arrive to the following expressions for the energy-momentum tensor [38]

$$T^0_0 = -a^{-4}\int q^2 dq d\Omega\,\sqrt{q^2 + m^2 a^2}\,f^{(0)}(q)\left(1 + \Psi\right)\ , \tag{2.46}$$

$$T^0_i = -a^{-4}\int q^2 dq d\Omega\,q\,n_i f^{(0)}(q)\,\Psi\ , \tag{2.47}$$

$$T^i_j = -a^{-4}\int q^2 dq d\Omega\,\frac{q^2 n_i n_j}{\sqrt{q^2 + m^2 a^2}}\,f^{(0)}(q)\left(1 + \Psi\right)\ . \tag{2.48}$$

21

Now, what we are left to specify is the evolution of the phase space distribution, which is described by the Boltzmann Equation,

$$\frac{\mathrm{d}f}{\mathrm{d}\tau} = C[f] = \frac{\partial f}{\partial \tau} + \frac{\partial f}{\partial x^i}\frac{\mathrm{d}x^i}{\mathrm{d}\tau} + \frac{\partial f}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}\tau} + \frac{\partial f}{\partial n_i}\frac{\mathrm{d}n_i}{\mathrm{d}\tau} \ . \tag{2.49}$$

where $C[f]$ represents collision processes that the particles experience, and their form can get very complex depending on the species considered and the level of accuracy intended in the physical description, also giving room for many parameters involved in these processes. It is for that reason that there are Boltzmann solvers like CLASS or CAMB, computational packages that solve the Boltzmann equation along with all the other perturbations equations and give all the theoretical observables as output. They include all the species mentioned in this thesis in the calculations, and more.

We can work the Boltzmann equation a bit more by noting that the last term in (2.49) is the multiplication of two first order quantities and thus can be neglected. We invoke the geodesic equation

$$P^0\frac{\mathrm{d}P^\mu}{\mathrm{d}\tau} + \Gamma^\mu_{\alpha\beta}P^\alpha P^\beta = 0 \ , \tag{2.50}$$

and, after some algebra, we obtain

$$\frac{\mathrm{d}q}{\mathrm{d}\tau} = q\dot{\phi} - \varepsilon n_i \partial_i \psi \ . \tag{2.51}$$

Wrapping all it up, we arrive to

$$\frac{1}{f^{(0)}}C[f] = \frac{\partial \Psi}{\partial \tau} + \mathrm{i}\,\frac{q}{\varepsilon}\,\Psi\,k^j n_j + \frac{\mathrm{d}\log f^{(0)}}{\mathrm{d}\log q}\left[\dot{\phi} - \mathrm{i}\,\frac{\varepsilon}{q}\,\psi\,k^j n_j\right] \ . \tag{2.52}$$

This equation is the one actually utilized in Boltzmann solvers. It is also fair to remember that these solvers *actually work on the synchronous gauge*, due to many calculations being faster to compute that way, and the consideration that this gauge is usually better for handling perturbations outside the causal horizon.

Chapter 3

The Large Scale Structure of the Universe

In order to describe the matter content of the Universe, it is important to picture it as a field that is nearly constant in which perturbations in density arise, then giving birth to galaxies. Now, we begin by defining the matter density contrast by

$$\delta(\mathbf{x}, t) \equiv \frac{\rho(\mathbf{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)} \,, \tag{3.1}$$

where $\rho(\mathbf{x}, t)$ is the local density in a given coordinate position $\mathbf{x}$ and cosmic time t, and $\bar{\rho}(t)$ is the mean density of the Universe at that time. The density contrast embeds the perturbation aspect just mentioned, and we can study it via its statistics, considering it a random field. For now, we will omit the time variable with the consideration that all spacial coordinates lie within the same time coordinate (and all quantities dependent on the matter density contrast). Therefore, we can define the 2-point correlation function as follows

$$\xi(\mathbf{r}) \equiv \langle \delta(\mathbf{x})\delta(\mathbf{x} + \mathbf{r}) \rangle \,, \tag{3.2}$$

with the braces denoting a spatial average. It is important to emphasize that we need to do a spatial average rather than averaging over many realizations, as usual in frequentist statistics[1], because the Universe is not an experiment we can reproduce. We may also refer to this as an ensemble average. Hence, it can also be written as

$$\xi(\mathbf{r}) = \frac{1}{V} \int \delta(\mathbf{x})\delta(\mathbf{x} + \mathbf{r})\mathrm{d}^3\mathbf{x} \,, \tag{3.3}$$

with V the integration volume. Later on, we will set it to the volume of a given galaxy survey.

It will now be essential to introduce the 3-dimensional Fourier transform convention to follow from now on [11].

$$\delta_{\mathbf{k}} = \frac{1}{V} \int \delta(\mathbf{x})\mathrm{e}^{-\mathrm{i}\mathbf{k}\mathbf{x}}\mathrm{d}^3\mathbf{x} \,, \tag{3.4}$$

$$\delta(\mathbf{x}) = \frac{V}{(2\pi)^3} \int \delta_{\mathbf{k}}\mathrm{e}^{\mathrm{i}\mathbf{k}\mathbf{x}}\mathrm{d}^3\mathbf{k} \,, \tag{3.5}$$

---

[1]In Chapter 4 we will cover the frequentist and Bayesian statistics basics.

where $\delta_{\mathbf{k}}$ is the Fourier conjugate of the matter density contrast (the same transformation rule applies for any function $f(\mathbf{x})$), as a function of the wave vector $\mathbf{k}$. Often, we would refer to $\mathbf{k}$ or rather its modulus k as the *scale* of a given quantity or formula, but, to avoid any confusion, we now clarify that the scale refers to the amplitude of the mode, via $\lambda = 2\pi/k$. This means that *large scales* are bound to *small k values* and vice-versa, because small k values dictate the behaviour of matter on large scales, as we will see shortly.

Now we can define the matter power spectrum through the 2-point correlation function of the Fourier transform of the density contrast

$$V\langle \delta_{\mathbf{k}}\delta_{\mathbf{k}'}^{*}\rangle = (2\pi)^3 P(\mathbf{k})\delta_D(\mathbf{k}-\mathbf{k}') \,, \tag{3.6}$$

with δ_D the Dirac delta function in Fourier space, defined by

$$\delta_D(\mathbf{k}) = \frac{1}{(2\pi)^3}\int e^{i\mathbf{k}\mathbf{x}}d^3\mathbf{x} \,. \tag{3.7}$$

Another important property is that we can relate the power spectrum to the correlation function via Fourier transformations, this is

$$P(\mathbf{k}) = \int \xi(\mathbf{x})e^{-i\mathbf{k}\mathbf{x}}d^3\mathbf{x} \,, \tag{3.8}$$

$$\xi(\mathbf{x}) = \frac{1}{(2\pi)^3}\int P(\mathbf{k})e^{i\mathbf{k}\mathbf{x}}d^3\mathbf{k} \,. \tag{3.9}$$

Now we can recover the time coordinate back to our calculations, in the form of the redshift. Therefore we can write the power spectrum $P(k,z)$, which can only depend on the wavenumber k due to the assumed isotropy of the Universe. Likely so, we can consider only the modulus of the space coordinate in the correlation function.

In galaxy surveys, one only obtains galaxies as counts, and from that, the density perturbations must be obtained. So, smoothing these shallow density peaks allows us to obtain the correlation function and then the power spectrum. An important quantity that uses a window function smoothing is the root-mean-square normalization of the power spectrum,

$$\sigma_R^2 = \frac{1}{2\pi^2}\int P(k)W_R^2(k)k^2 dk \,, \tag{3.10}$$

where R is a characteristic scale in Mpc/h units and W_R is the Fourier transform of a window function in real space, which is usually a top-hat function that smooths structures smaller than R. In that case, it reads

$$W_R(k) = \frac{3(\sin(kR) - kR\cos(kR))}{(kR)^3} \,. \tag{3.11}$$

It is usual to choose $R = 8\text{Mpc}/h$ since this value makes the normalization to be in the order of one (currently, $\sigma_8 = 0.8102 \pm 0.00601$ [9]). This number gives us a sense of linearity limit for structures, where values for σ_R above 1 alert that the scale is not suited to a linear treatment.

Another important property of the window function smoothing is the Poisson noise induced on the covariance of the density contrast: We begin by taking the smoothed density contrast δ_s,

$$\delta_s(\mathbf{x}) = V\delta(\mathbf{x})W(\mathbf{x}) \,, \tag{3.12}$$

where the subscript 's' stands for smoothed, meaning that we get a smoothed density contrast after applying the window selection function. This is important since δ consists in several sharp peaks in space (which are modelled through Dirac deltas). If we take the covariance of the Fourier transform of δ_s, we obtain an extra noise spectrum [11],

$$\Delta^2(\mathbf{k}) \equiv \langle \delta_{s,\mathbf{k}} \delta_{s,\mathbf{k}}^* \rangle = P(\mathbf{k}) + P_n \; . \tag{3.13}$$

Here, $P_n = V/N = 1/n$ is a Poisson noise, equal to the inverse of the number density of galaxies, and corresponds to the power spectrum of a sample of galaxies at random positions, or in a Poisson distribution. Loosely speaking, we usually take $\delta_{s,\mathbf{k}}$ as $\delta_{\mathbf{k}}$ in most calculations, but we take Eq. (3.13) into consideration whenever the covariance of the density contrast is needed to perform any statistical procedure.

3.1 The matter power spectrum

In this section, we describe the components of the power spectrum, and we can treat it on two parts: the primordial power spectrum, which we will see as a starting condition, and its evolving factors.

An important part of the cosmological view is the inflation epoch of the Universe, a process that gave origin to an extreme expansion after the Big Bang, flattening the Universe and shaping the primordial power spectrum, which can be parametrized by

$$P_{\text{prim}}(k) = A_s k^{n_s} \; , \tag{3.14}$$

where A_s is the amplitude of the power spectrum and n_s is called the spectral index of scalar perturbations, a number close to 1, that comes as a result from a power-law inflation[2]. If we approximate $n_s \approx 1$, the matter power spectrum is called the Harrison-Zel'dovich spectrum $P_{HZ}(k) \equiv A_s k$, which comes from an exponential inflation [11].

Then, the matter power spectrum in the linear regime can be described through the formula [11, 39]

$$P(k,z) = A_s k^{n_s} T^2(k) G^2(z) \; , \tag{3.15}$$

In this equation, we introduce $T(k)$, the transfer function, which encapsulates the behavior of the power spectrum over different scales. This quantity is paramount when running a Boltzmann solver, due to the complexity of the physics underlying the radiation era, in contrast to the decoupling over different scales of the dark components when dominating the Universe, in which case their perturbation's evolution can be obtained from Eqs. (2.31)-(2.34) and (2.37)-(2.38). A plot of the matter power spectrum can be seen in Fig. 3.1.

We also included $G(k,z)$, called the growth factor of density perturbations. It is defined by

$$G(k,z) \equiv \frac{\delta(k,z)}{\delta(k,0)} \; . \tag{3.16}$$

---

[2]Currently, the constrains on these parameters are $10^9 A_s = 2.105 \pm 0.030$ and $n_s = 0.9665 \pm 0.0038$ [9].
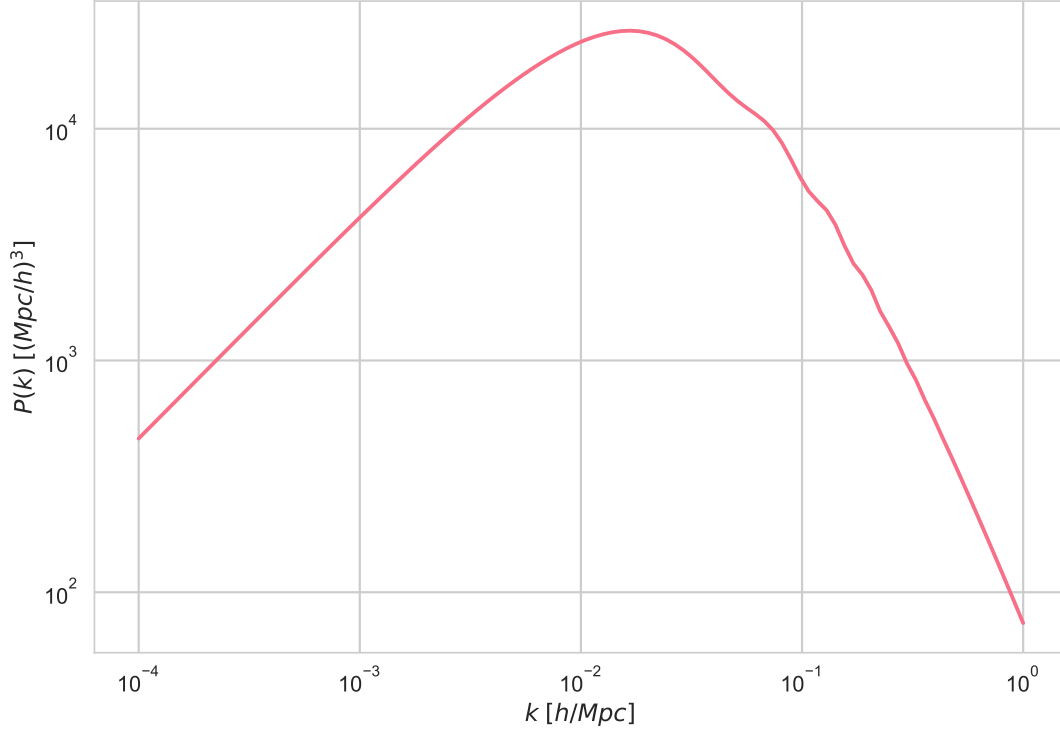
Figure 3.1: A plot of the matter power spectrum from linear perturbation theory as a function of k, at redshift $z = 0$. This was obtained from the CLASS code with its default parameters. Notice the almost scale-invariant part until $k \sim 10^{-2}$ h/Mpc, which corresponds to the primordial power spectrum. The peak occurs at $k_{\rm eq} \sim 0.015$ h/Mpc, the wavenumber that crosses the Hubble horizon at the radiation-matter equality time; wavenumbers greater than it have had time to properly evolve and form structure, after being unfrozen as they already crossed the Hubble horizon. On smaller scales, the *BAO wiggles* (product of the oscillations in the baryon-photon fluid) can be seen around $k = 1/r_s \sim 10^{-1}$ h/Mpc. Care must be taken when going into the range $k \gtrsim 10^{-1}$ h/Mpc, where one should consider switching to nonlinear perturbation theory to obtain accurate results (this plot does not consider nonlinear corrections).

In most cases, we can drop the dependence on the wavenumber k, because the behavior is almost constant across linear scales (although it depends on the model chosen), and we will adopt this choice from now on. The constant behavior comes from considering sub-horizon scales in a Universe composed only of cold dark matter and a cosmological constant (ΛCDM), which makes the perturbations equations, Eqs. (2.37)-(2.38), scale-independent. This consideration can be done when considering the Universe from the matter-dominated era onwards[3].

A related quantity to the the growth function Eq. (3.16) is the growth factor of matter

---

[3]Another more conservative start point is after the decoupling epoch.

perturbation, defined by

$$f(a) = \frac{\mathrm{d}\log\delta(a)}{\mathrm{d}\log a} \ .$$

(3.17)

This quantity has been important in cosmology since it is closely related to the velocity distribution of galaxies, as we will see for the RSD effect in Eq. (3.21). Another interesting fact is that the growth rate has a very good empirical fit $f = \Omega_m^\gamma$, originating the *growth index* of matter perturbations, an almost constant quantity for ΛCDM. In particular, $\gamma \approx 0.55$ [11].

3.2 Observational effects in the matter power spectrum

Several considerations and effects come in when measuring the matter power spectrum via the mapping of galaxy counts. In this section, we present some of the most important effects to take into account for a galaxy clustering survey, while the more technical ones (which address the redshift-binning of the measurements) are left for discussion on Chapter 5. This chapter is essential to set the grounds for the analysis of Chapter 5.

3.2.1 The Bias

It is important to emphasize that what we measure in a galaxy clustering survey are *galaxies*, and that means we can only trace the luminous part of the matter content in the Universe. However, we can assume that the galaxy density contrast δ_g is related to the total matter density contrast δ_m by a bias factor b,

$$b = \frac{\delta_g}{\delta_m} \ .$$

(3.18)

This procedure, while has been proven to work fairly well, is far from being an optimal description of the relation between the density contrasts of galaxies and dark matter. In a more general regime, b could depend on time and the scale, or the type of galaxy considered, like blue and red galaxies, which are younger and older, respectively. This consideration immediately gives us the following correction for the galaxy power spectrum, as a function of the total matter power spectrum $P(k)$

$$P_{\mathrm{obs}}(k) = b^2 P(k) \qquad (\text{w/ Bias}) \ .$$

(3.19)

3.2.2 Redshift Space Distortions

Galaxy distances in a galaxy clustering survey are measured through their redshifts. Then, the peculiar velocities of the galaxies as they tend to fall towards massive clusters produce differences in the observed power spectrum. This effect can be parametrized and included in the observed power spectrum. Before we can parametrize this effect, we must first get to know the velocities of galaxies, as a function of the quantities we have derived so far.

Velocity field and bulk flow

We start by obtaining the peculiar velocity as in Eq. (2.37), for the case of a baryonic matter fluid ($w = c_s = 0$ and considering a small scale $1/k$),

$$v^i = \frac{\mathrm{i}}{k_i}\dot{\delta}_k \ . \tag{3.20}$$

Separating the time part from the density contrast using the growth factor, Eq. (3.16), $\delta(k,\tau) = \delta_k G(\tau)$, and also using $\mathrm{d}/\mathrm{d}\tau = aH(a)\,\mathrm{d}/\mathrm{d}a$ we obtain

$$v^i = \frac{\mathrm{i}}{k_i}\delta_k aH(a)\frac{\mathrm{d}\ln G}{\mathrm{d}a} = \frac{\mathrm{i}}{k_i}a^2 H(a)f(a)\delta_k \ , \tag{3.21}$$

where we used the growth rate from Eq. (3.17). Now we consider local scales in the vicinity of the cluster, so we set $a \approx 1$ and $H \approx H_0$, getting to

$$\mathbf{v} = \mathrm{i}H_0 f\delta_k \frac{\mathbf{k}}{k^2} \ . \tag{3.22}$$

We then obtain the velocity in real space by inverse-Fourier transforming the former expression

$$\mathbf{v}(\mathbf{x}) = \mathrm{i}H_0 f\frac{V}{(2\pi)^3}\int \delta_k \frac{\mathbf{k}}{k^2}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\mathrm{d}^3 k \ . \tag{3.23}$$

Going a bit further, we can relate this velocity field to the power spectrum. We do so by considering the average of the squared speed in a volume V_R of radius R,

$$\langle v^2 \rangle_R = H_0^2 f^2 \frac{V^2}{(2\pi)^6}\int \langle \delta_k \delta_{k'}^* \rangle \frac{\mathbf{k}}{k^2}\frac{\mathbf{k'}}{k^2}W_R(k)W_R(k')\mathrm{d}^3 k\,\mathrm{d}^3 k'$$

$$= \frac{H_0^2 f^2}{2\pi^2}\int P(k)W_R^2(k)\mathrm{d}k \ , \tag{3.24}$$

$$\tag{3.25}$$

where we integrated over the angles and utilized the matter power spectrum definition from Eq. (3.6) and used a window function W_R as in Eq. (3.10). $\langle v^2 \rangle_R$ is called the bulk flow on the scale R. So, if a survey constrains the power spectrum and the bulk flow independently, it is possible to constrain the growth rate f.

Now, we continue to derive the redshift space distortion (RSD) effect on the power spectrum. We define the line-of-sight velocity $\mathbf{u}$ via

$$\mathbf{u}(r) = \mathbf{v}\frac{\mathbf{r}}{r} \ . \tag{3.26}$$

Now we define the redshift space vector $\mathbf{s}$ through

$$\mathbf{s} = \mathbf{r}\left[1 + \frac{\Delta u(r)}{r}\right] \ , \tag{3.27}$$

where $\Delta u(r) = u(r) - u(0)$. The importance of the redshift space comes from the fact that galaxy distances are measured on the redshift space, via the galaxy number counts in the

detector. Then, our task is to compare the different density contrasts in real and redshift space via a coordinate transformation. From now on, we divide velocities by H_0 to obtain Mpc units. We accept that the number of galaxies is the same in both spaces, in other words,

$$n(r)\mathrm{d}V_r = n(s)\mathrm{d}V_s \ , \tag{3.28}$$

where the subscrips r and s refer to real and redshift space, respectively. So, we can relate the volumes through a coordinate transformation with a Jacobian matrix J, whose determinant is

$$|J| = \left|\frac{\partial \mathbf{s}}{\partial \mathbf{r}}\right| = 1 + \frac{\mathrm{d}u}{\mathrm{d}r} \ . \tag{3.29}$$

Therefore, we write the volume as

$$\mathrm{d}V_s = s^2 \mathrm{d}s\mathrm{d}\theta\mathrm{d}\phi = \left(1 + \frac{\Delta u(r)}{r}\right)^2 |J|\mathrm{d}V_r \ . \tag{3.30}$$

Then, we can write the density contrast in redshift space by using Eqs.(3.28) and (3.30),

$$\delta_s = \frac{n(s)\mathrm{d}V_s}{n_0\mathrm{d}V_s} - 1 = \frac{n(r)}{n_0\left(1 + \frac{\Delta u(r)}{r}\right)^2 |J|} - 1 \ . \tag{3.31}$$

Therefore, we obtain (with approximations)

$$\delta_s \approx \delta_r - 2\frac{\Delta u(r)}{r} - \frac{\mathrm{d}u}{\mathrm{d}r} \ . \tag{3.32}$$

We now continue to express the velocity field $\mathbf{v}$ from Eq.(3.23), now including the bias in the parameter $\beta \equiv f/b$,

$$\mathbf{v} = \mathrm{i}H_0\beta \int \delta_{(g)k}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\frac{\mathbf{k}}{k}\mathrm{d}^3k \ , \tag{3.33}$$

where we omitted the constant $V/(2\pi)^3$ and included the subscript (g) to denote that we are considering a galaxy density contrast. Now we can derive $u(r)$ and $\mathrm{d}u/\mathrm{d}r$ from this equation,

$$u(r) = \mathrm{i}\beta \int \delta_{(g,r)k}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\frac{\mathbf{k}\cdot\mathbf{r}}{k^2 r}\mathrm{d}^3k \ , \tag{3.34}$$

$$\frac{\mathrm{d}u}{\mathrm{d}r} = -\beta \int \delta_{(g,r)k}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\left(\frac{\mathbf{k}\cdot\mathbf{r}}{kr}\right)^2\mathrm{d}^3k \ . \tag{3.35}$$

Therefore, it follows that

$$\delta_{(g,s)k} = \delta_{(g,r)k} + \beta \int \delta_{(g,r)k}\left(\frac{\mathbf{k}\cdot\mathbf{r}}{kr}\right)^2\mathrm{d}^3k \ . \tag{3.36}$$

If we consider small angular size surveys, we can assume that the cosine μ is constant, with

$$\mu \equiv \frac{\mathbf{k}\cdot\mathbf{r}}{kr} \ . \tag{3.37}$$

29

Finally, we get to the relation between galaxy density contrasts in redshift space and real space.

$$\delta_{(g,s)k} = \delta_{(g,r)k}(1 + \beta\mu^2) \ , \tag{3.38}$$

This last expression leads us to the next correction in the observed power spectrum, which we note is now not isotropic:

$$P_{\text{obs}}(k, \mu) = P(k, \mu)b^2(1 + \beta\mu^2)^2 \qquad (\text{w/ Bias and RSD}) \ , \tag{3.39}$$

where 'RSD' stands for Redshift Space Distortions. We remind the reader that we measure the *galaxy* power spectrum in redshift space, while the theoretical power spectrum is the total matter power spectrum in real space.

3.2.3 The Alcock-Paczynski Effect

Our next task is to include the Alcock-Paczynski effect in the observed matter power spectrum. This treatment addresses the change in volume in the power spectrum when changing the parameters, this is, how the change in parameters affects the conversion of redshifts and angles to distances. In concrete, we will study how this effect changes the wavevector and the volume the survey, and how this impacts the observed power spectrum.

We begin by considering two cosmologies: a reference cosmology (with a subscript 'r') and another cosmology (no subscript). We split the reasoning in a transverse and a radial part.

If we have an object of transverse comoving size d_M at redshift z, and we observe it subtends an angle θ, its angular diameter distance would be $d_A = d_M/(1+z)\theta$, as in Eq. (1.48). In this example, the angular diameter distance and the transverse comoving distance are cosmology-dependent, while the redshift and angle are not. So, if we take another cosmology, the fraction d_A/d_M would still hold constant. Then, if we recognize $d_M = \lambda_\perp = 1/k_\perp$ as our transverse scale, we would have

$$k_\perp = k_{r,\perp} \frac{d_{A,r}}{d_A} \ , \tag{3.40}$$

which explains the change of the transverse part of the wavevector when changing parameters.

In a similar way, we now consider a small object of small comoving size $\text{d}(d_c)$ subtending a small redshift range $\text{d}z$. Therefore, by Eq. (1.2.2) we obtain $\text{d}(d_c) = c\text{d}z/H(z)$. Thus, taking $\text{d}(d_c) = \lambda_\parallel$, the product $\lambda_\parallel H$ is a constant across cosmologies and we can express the parallel wavenumber correction:

$$k_\parallel = k_{r,\parallel} \frac{H}{H_r} \ . \tag{3.41}$$

Using what we just obtained, we can derive expressions for the magnitude of the wavenumber k and the cosine μ:

$$k = \sqrt{k_\perp^2 + k_\parallel^2} = R_{AP}k_r \ , \tag{3.42}$$

$$\mu = \frac{k_\parallel}{k} = \frac{H\mu_r}{H_r R_{AP}} \ , \tag{3.43}$$

where

$$R_{AP} = \frac{\sqrt{H^2 d_A^2 \mu_r^2 - H_r^2 d_{A,r}^2 (\mu_r^2 - 1)}}{H_r d_A} \ . \tag{3.44}$$

Analogously, we can obtain the change in volume (which is the multiplication of a parallel scale correction and two transverse scale corrections),

$$V = V_r \frac{H_r d_A^2}{H d_{A,r}^2} \ . \tag{3.45}$$

It is useful to write the matter power spectrum as $P(k) = V \delta_k^2$ in its true cosmology. That way, we can note that expressing the observed power spectrum (which uses a reference cosmology to transform redshifts) induces a correction factor of V_r/V. Consequently, the observed matter power spectrum now reads:

$$P_{\text{obs}}(k_r, \mu_r, z) = \frac{H(z) d_{A,r}^2(z)}{H_r(z) d_A^2(z)} P(k, \mu, z) b^2(z) \big(1 + \beta(z)\mu^2\big)^2 \qquad (\text{w/ Bias, RSD and AP}) \ , \tag{3.46}$$

where 'AP' stands for Alcock-Paczynski. We also made the redshift dependance explicit, allowing for a redshift-dependent bias $b(z)$. Also, note that the cosine μ in the RSD correction is also affected by the correction.

Using Baryon Acoustic Oscillations

So far, we can measure the observed power spectrum and, through velocity statistics on the correlation function, we can also obtain the growth rate $f(z)$ [4]. Our remaining task is, then, to obtain $H(z)$ and $d_A(z)$, which is accomplished by utilizing baryon acoustic oscillations. This is done by using the standard ruler property of the BAO scale at the drag epoch, r_s, which is a well-constrained quantity on CMB surveys [9]. Thus, one can simply obtain $H(z)$ and $d_A(z)$ using Eqs. (1.50) and (1.51), assuming a known formula to link $r_\parallel$ and $r_\perp$ to r_s. In practice, BAO surveys define parameters related to H and d_A, such as the transverse radial BAO dilation parameters $\alpha_\perp$ and $\alpha_\parallel$ [33],

$$\alpha_\perp \equiv \frac{H_r(z) r_{s,r}}{H(z) r_s} \ , \tag{3.47}$$

$$\alpha_\parallel \equiv \frac{d_A(z) r_{s,r}}{d_{A,r}(z) r_s} \ , \tag{3.48}$$

where 'r' also stands for a reference cosmology. These parameters are near to constant and can be included as free parameters in the analysis. Nonetheless, more general parametrizations exist, and are better suited to take in consideration several redshift bins, see [40].

---

[4]To avoid the bias, a more robust quantity to obtain is the combination $f\sigma_8$.

3.2.4 Redshift error supression

A simple but important effect to take into account is the redshift error coming from the survey. If this error comes as the standard deviation σ_z of a Gaussian distribution, then the error in comoving distance will be $\sigma_r = c\,\sigma_z/H(z)$, since $\mathrm{d}(d_c) = c\,\mathrm{d}z/H(z)$. Thus we consider that the position measures also distribute as a Gaussian centered in the true position r_0 with a standard deviation σ_r, or

$$f_{r_0}(r) = \frac{1}{\sqrt{2\pi}\sigma_r} \mathrm{e}^{-\frac{1}{2}[(r-r_0)/\sigma_r]^2} . \tag{3.49}$$

Then, the correlation function will be updated by marginalizing over r [11],

$$\xi(r_0) = \int_0^\infty \xi(r) f_{r_0}(r) \mathrm{d}r . \tag{3.50}$$

This can be done thanks to the correlation representing the probability of having galaxies separated by a certain distance. The Fourier transform of a convolution is the product of the Fourier transform of each term. The 1-D Fourier transform of the function f_{r_0} is[5]

$$F[f_{r_0}](k_\parallel) = \mathrm{e}^{-k_\parallel^2 \sigma_r^2} , \tag{3.51}$$

where we considered that the 1-D Fourier transform must be done in the line of sight component of a more general wavevector $\mathbf{k}$, and we approximated this Gaussian to be negligible for negative values, in order for it to be normalized in $(-\infty, \infty)$. Therefore, if we Fourier transform the 2-point correlation function into the matter power spectrum, and remembering $k_\parallel = k\mu$, we obtain

$$P_{\mathrm{obs}}(k, \mu) = P\, \mathrm{e}^{-k^2\mu^2\sigma_r^2} \qquad (\text{w/ } z\text{-error}) . \tag{3.52}$$

3.2.5 Shot noise and survey types

Our last effect addition is a shot noise P_s, scale independent, that makes up for white noise from the detector itself. This noise may or may not be significant.

Putting everything together, we have:

$$P_{\mathrm{obs}}(k_r,\, \mu_r,\, z) = P_{\mathrm{shot}}(z) + \frac{H(z)\,d_{A,r}^2(z)}{H_r(z)\,d_A^2(z)} P(k, \mu, z) b^2(z)\left(1 + \beta(z)\mu^2\right)^2 \mathrm{e}^{-k^2\mu^2\sigma_r^2}$$
$$(\text{w/ Bias, RSD, AP, } z\text{-error and Shot noise}) . \tag{3.53}$$

Galaxy clustering surveys are of two types: spectroscopic and photometric. Spectroscopic surveys utilize a dispersive element to separate the light and then observe features to obtain redshifts, while photometric surveys use color filters to split the light in different wavelengths. The practical difference is that spectroscopic surveys can detect fewer sources simultaneously

---

[5]See Wolfram MathWorld for more details.

than photometric surveys[6], but the redshift measurement is more accurate. Both types of survey have their advantages and they also target different types of galaxies.

The most notable examples of future galaxy clustering surveys are LSST, a ground-based telescope which is being built in the north of Chile which measures galaxies by photometric means, and Euclid, a space telescope of spectroscopic nature. In fact, our first project, on Chapter 5, delves into the use of the observed power spectrum to make forecasts on the parameter constraining power of LSST and uses all of the effects mentioned in this section.

---

[6]This can be parametrized by an *efficiency factor* that decreases the effective number density of galaxies.

Chapter 4

Bayesian Statistical Methods

Now that we have worked out the cosmological background needed for the upcoming projects, we shift our focus to the Bayesian statistical tools we are going to use for the different analyses. We follow Refs. [1] and [41]. Our first task is to explain why we rely so much on Bayesian statistics rather than the traditional, frequentist statistical methodology. Next, we go more into detail into the principal aspects of Bayesian tools: inference, forecasting and model comparison and their related numerical/approximation methods. We will frequently refer to a 'probability density function' by 'PDF'.

4.1   Frequentist and Bayesian experiments

A simple way to illustrate the importance of Bayesian statistics in cosmology is to inspect the Bayes' theorem

$$P(A|B) = P(B|A)\,\frac{P(A)}{P(B)}\ . \tag{4.1}$$

This equation relates the probability of two events, A and B. To be more specific, we can set these variables to $\mathbf{x}$ and $\boldsymbol{\theta}$, where $\mathbf{x}$ represents a data vector obtained from a certain experiment and $\boldsymbol{\theta}$ is a parameter vector that characterizes the experiment.

$$P(\boldsymbol{\theta}|\mathbf{x}) = P(\mathbf{x}|\boldsymbol{\theta})\,\frac{P(\boldsymbol{\theta})}{P(\mathbf{x})}\ . \tag{4.2}$$

This equation is a complete game-changer, since we can completely move the interpretation of the procedure we are considering: from seeing the probability of results given some parameters, to seeing the probability of the parameters given some results.

The word *experiment* is crucial to define the concept of probability in both branches of statistics:

- **Frequentist**: If you repeat an experiment many (infinite) times, the probability of a certain event/situation will be the *fraction of outcomes* that fulfills it. This relative frequency gives the frequentist philosophy its name.

- **Bayesian**: Probability is defined as the *degree of belief* one has in a situation or event to happen.

We can immediately see the lack of context put into these definitions, but it can be clarified by determining the nature of the experiment in question, especially on the possibility of the repetition of the experiment for where frequentists laid the foundations of the theory. To be fair, both approaches are appropriate for certain experiments, while many others would argue that the Bayesian approach is more general and more multipurpose (see Ref. [1]). Nonetheless, there are experiments that are inherently frequentist, like experimental high-energy physics, where particle collisions are studied via analysis of their frequencies. On the other hand, there are experiments completely Bayesian, like the study of our Universe, which is hard to see as an experiment, since we only have one. In spite of having only one Universe, we can study certain phenomena in an ensemble-average fashion like the 2-point correlation function of Eq. (3.2), a procedure which lies in a middle-ground between frequentist and Bayesian philosophies (even though the procedures utilized afterward for the analysis are entirely Bayesian).

4.2 Bayesian Inference

In this section we apply the Bayes theorem in order to get insights into the parameters of a certain model, which will remain fixed for now (we will consider several models in Sec. 4.4). We take a new look on Eq. (4.2),

$$\mathcal{P}(\boldsymbol{\theta}|\mathbf{x}) = L(\mathbf{x}|\boldsymbol{\theta}) \, \frac{\pi(\boldsymbol{\theta})}{\mathcal{E}(\mathbf{x})} \, , \tag{4.3}$$

where we put different labels and names to our probabilities. From now on we will consider these probabilities to be probability distributions, as is more common to see in cosmology. Here, $\mathbf{x} = (x^1, x^2, \ldots, x^N)$ are N random data, and $\boldsymbol{\theta} = (\theta^1, \theta^2, \ldots, \theta^n)$ are the n theoretical parameters of our physical model which (hopefully) describes the results. Each one of the terms in Eq. (4.3) is interesting to explain on its own, so we proceed to list their definitions,

4.2.1 Likelihood

The likelihood $L(\mathbf{x}|\boldsymbol{\theta})$ can be seen as the *sampling distribution* of our data. This is the quantity used in the frequentist regime to do inference, and is often written as function of the parameters $L(\boldsymbol{\theta}) \equiv L(\mathbf{x}|\boldsymbol{\theta})$. The likelihood function is usually the place to apply the central limit theorem in order to obtain a Gaussian distribution to approximate the data, which will be a correct strategy if there are enough measurements to satisfy the theorem. In the case of a Gaussian likelihood, its equation is

$$L(\mathbf{x}|\boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^N \det \mathbf{C}(\boldsymbol{\theta})}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}(\boldsymbol{\theta}))^T \mathbf{C}^{-1}(\boldsymbol{\theta})(\mathbf{x} - \boldsymbol{\mu}(\boldsymbol{\theta}))\right) , \tag{4.4}$$

where $\boldsymbol{\mu}(\boldsymbol{\theta})$ is the mean of the Gaussian distribution, representing the prediction of the model about the data, and $\mathbf{C}(\boldsymbol{\theta})$ is the covariance matrix of the data. As explicited in Eq. (4.4), $\boldsymbol{\mu}$ and $\mathbf{C}$ depend on the parameters of the model, but usually only one of them is affected by the parameters, depending on the nature of the observable and its produced data. In cosmology, random fields like the matter power spectrum, the CMB or weak gravitational lensing can be described by a covariance that depends on the cosmology, while other probes like SN Ia, BAO's or direct Hubble measures[1] are treated utilizing a parameter-dependent mean.

4.2.2 Prior probability

The prior probability $\pi(\mathbf{x})$ represents the knowledge we have about the parameters *before* seeing the data, like other experiments or theoretical findings. The prior can easily be turned into a topic of controversy, since we, in principle, can set whatever distribution we like, and if we do not like any distribution, we have to select a prior anyways, even if it is a constant (a *flat prior*) we have to specify boundaries for it to make it normalizable as any other probability distribution. The prior can heavily impact the results of any Bayesian analysis, especially on model selection, so one must be reasonable at selecting the priors and always explicit them and their goals. The most common example are *uninformative priors*, which are based on the sensitivity of the posterior to the parameters [42].

4.2.3 Posterior probability

The posterior probability $\mathcal{P}(\boldsymbol{\theta}|\mathbf{x})$ represents our state of knowledge *after* considering all the other quantities described. And by 'after', we mean in a logic sense rather than in a temporal way, as our prior knowledge can very well be based on newer experiments that complement an outdated theory and/or dataset we might attempt to analyze. In order to constrain our parameters, we focus on obtaining the best fit (the maximum posterior probability point) of parameters and their *confidence regions*[2]. These confidence regions consist a plot outlining the 68% and 95% probability regions[3], in other words, an area in a two-dimensional parameter space where there is a 68% or 95% probability of having the parameters. This region has a contour level line where the posterior probability distribution is a constant. The confidence region procedure is then done for every pair of parameters. A very similar plot is the confidence interval of one parameter, which is the 1-dimensional equivalent to the confidence region. We show an example of confidence regions on Fig. 4.1.

---

[1]In Chapter 7 we include direct $H(z)$ measurements through *cosmic chronometers*, which are evolving galaxies that can be related to a certain time on the history of the Universe via its characteristics, like *standard clocks*.

[2]Some authors (as in Ref. [43]) prefer to refer to the confidence regions (or confidence intervals) as a purely frequentist concept, as it is used for likelihood-based constraining (and is essentially a random variable too), while reserving the term *credible regions* for its Bayesian homologous. In this thesis, we will loosely utilize the more common 'confidence regions' term.

[3]The numbers 68 and 95 come from the case of a 1-dimensional Gaussian distribution of mean μ and variance σ^2, where the intervals defined by $\mu \pm 1\sigma$ and $\mu \pm 2\sigma$ enclose a 68% and 95% probability, respectively.
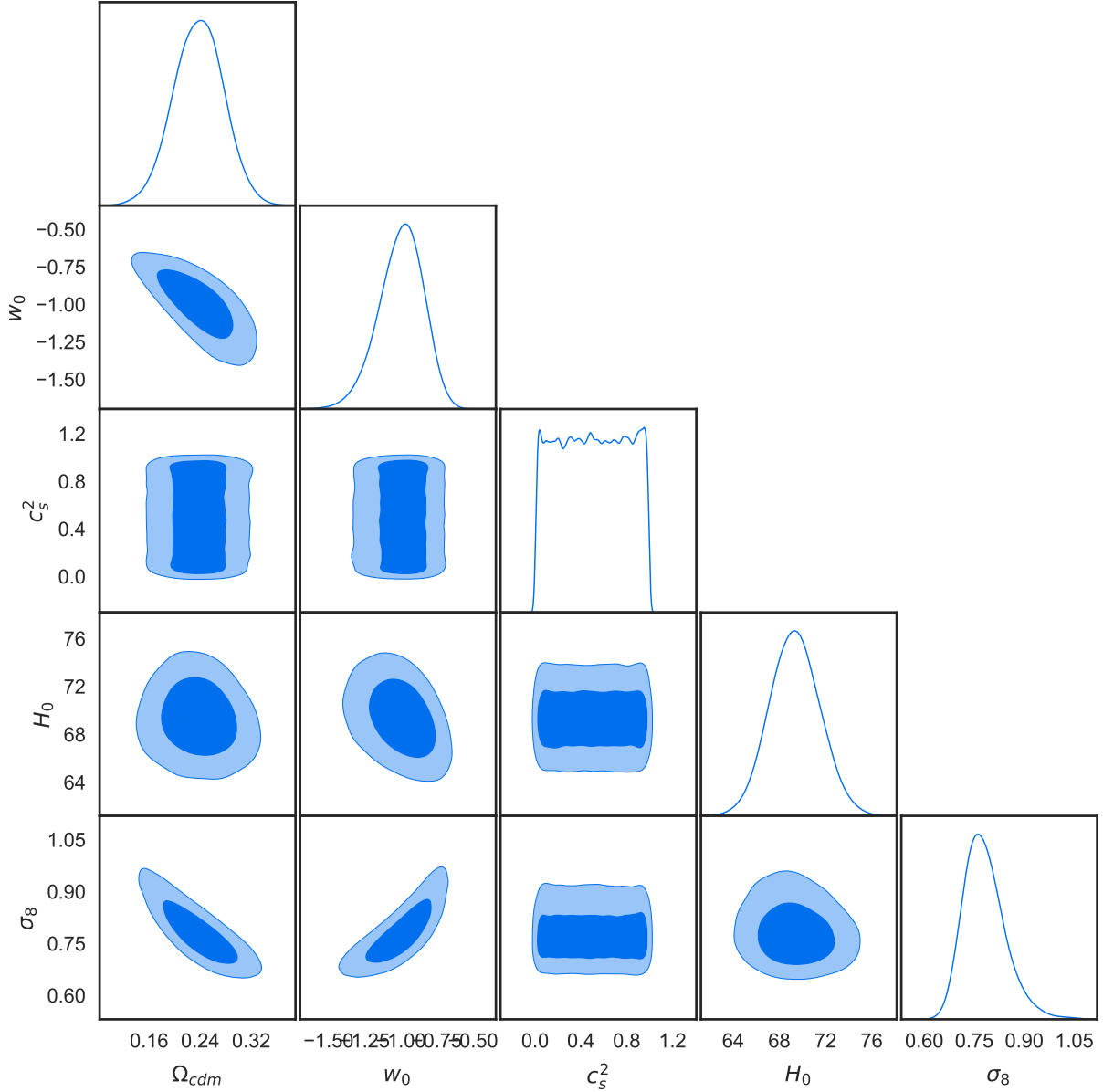
Figure 4.1: An example of confidence regions for the parameters of a model, also known as a *triangle plot* or *corner plot*. The regions enclose the 68% and 95% probability regions. The model used was ΛCDM with a free sound speed parameter, computed using the CLASS code, and the data used are the growth rate and Hubble parameter compilation used in the projects of Chapters 6 and 7, plus the Union2.1 SN Ia catalogue [3]. One can observe several degeneracies on the parameters (easily identifiable by their *banana shapes*). It is especially interesting to see that c_s^2 is not constrained at all by the data since we used a flat prior between 0 and 1 for that parameter.

4.2.4 The Evidence

The evidence $\mathcal{E}(\mathbf{x})$ is the central quantity for Sec. 4.4, since it represents the probability of having the data given the assumption of a theoretical model (for now we have kept the

model implicit). The evidence is also called *marginal likelihood* and *model likelihood*, since its resemblance to the likelihood. Nevertheless, for most inference problems it can be ignored since it only impacts the inference as a normalization factor for the posterior probability distribution. In that regard, it can be computed via

$$\mathcal{E}(\mathbf{x}) = \int L(\mathbf{x}|\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})\,\mathrm{d}^n\theta\;. \tag{4.5}$$

We can get rid of the evidence when doing inference by considering the log-posterior probability since it only represents a normalization factor, and most numerical methods rely on doing just that to explore the parameter space more efficiently as we will see shortly.

As mentioned in the posterior probability subsection 4.2.3, we base our parameter constraining efforts into *sampling* this probability distribution rather than evaluating it in a grid of n dimensions, because this becomes extremely difficult when the number of dimensions grows. As a rule of thumb, when the parameter dimensionality surpasses 3, one must resort to the more efficient sampling techniques. By sampling the distribution, we mean to use a Markov Chain Monte Carlo sampler in order to explore the parameter space in a very efficient way, defeating the old-fashioned method of grid-evaluation.

4.2.5 Markov Chain Monte Carlo methods

The use of Markov Chain[4] Monte Carlo[5] (hereinafter MCMC) methods exploded about 15 years ago when the CosmoMC package became available to the scientific community [37], a code that evolved to become the CAMB Boltzmann code. In practice, when constraining the parameters of a model, one employs a MCMC sampler to simulate a collection of parameter points $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_{(1)}, ..., \boldsymbol{\theta}_{(s)}\}$, such that they follow the posterior distribution (s will be the number of samples). Having obtained the chain points $\boldsymbol{\Theta}$, one can obtain a density estimation (with a kernel smoothing) to plot the confidence regions, by counting the samples that fall in an iso-probability region, be it the 68% or 95% of the points usually.

Another useful property of MCMC-chains is that marginalization of parameters become a trivial process. Traditionally, to marginalize over a certain parameter(s) involves integrating the posterior over that(those) parameter(s). For example, if we want to marginalize over all the parameters except the first one, we do it by computing

$$\mathcal{P}(\theta^1|\,\mathbf{x}) = \int \mathcal{P}(\boldsymbol{\theta}|\,\mathbf{x})\mathrm{d}\theta^2\dots\mathrm{d}\theta^n\;. \tag{4.6}$$

The quantity $\mathcal{P}(\theta^1)$ is called a *marginal posterior probability density*. This is extremely useful when some of the parameters are of no scientific interest, but nevertheless, have to be included in the analysis for completeness. These are called *nuisance parameters* and they usually are parameters of the experimental setup that are unrelated to the physics of the phenomenon observed. They can also be, say, astrophysical fitting formulas that are not important in a cosmological analysis.

---

[4]A 'Markov Chain' is a process where the state of the new step is only defined by the last step, i.e. it has no memory of past steps other than the last one.

[5]And 'Monte Carlo' refers to algorithms based on random sampling.

Continuing with the perks of MCMC sampling, one can reuse a chain in order to obtain the posterior probability distribution of any quantity that is a function of the parameters. We denote this quantity by $q = q(\boldsymbol{\theta})$ and its derived chain is

$$q(\boldsymbol{\Theta}) = \left\{ q(\boldsymbol{\theta}_{(1)}), ..., q(\boldsymbol{\theta}_{(s)}) \right\}, \tag{4.7}$$

this is, applying the function to every point of the chain returns a chain of the function. This procedure can be useful to compute the PDF for intermediate quantities in a calculation. This property is exploited in Chapter 7.

Having mentioned the principal characteristics of MCMC chains, we could not mention the most popular samplers like Metropolis-Hastings [44, 45], Hamilton Monte Carlo sampling and Gibbs sampling [46]. The simplest MCMC sampler is the former. Due to it being so straightforward to implement, we will use it in our projects. We now proceed to outline the algorithm structure to produce samples:

Metropolis-Hastings algorithm

1. We start on a random point in the parameter space (which can be drawn from the Prior PDF). We denote it by $\boldsymbol{\theta}_{(1)}$.
2. A proposed step, $\boldsymbol{\theta}_{(\text{prop})}$, is drawn from a step probability function (usually Gaussian) centered in the current point.
3. The proposed step is accepted with a probability equal to the ratio of the posteriors. If this is ratio is larger than 1, it is accepted immediately, i.e.

$$P(\text{Acceptance}) = \min \left(1, \frac{\mathcal{P}(\boldsymbol{\theta}_{(\text{prop})}|\, \mathbf{x})}{\mathcal{P}(\boldsymbol{\theta}_{(i)}|\, \mathbf{x})} \right). \tag{4.8}$$

4. If the proposed step is accepted, it gets included in the chain. In other words, $\boldsymbol{\theta}_{(i+1)} = \boldsymbol{\theta}_{(\text{prop})}$. On the contrary, if the proposed step is not accepted, the current point gets updated to *itself*. This is $\boldsymbol{\theta}_{(i+1)} = \boldsymbol{\theta}_{(i)}$.
5. Return to step 2 until sufficient points are stored.

Running an MCMC chain can be a tough labor since several requirements must be secured in order to obtain a *good* chain, in the sense that it truly lives up to represent the posterior PDF of the parameters.

A way to enhance the parameter space exploration in a chain is using a step that has a shape similar to the posterior. This can be achieved by choosing a Gaussian step, with a covariance equal to the covariance of another chain that has already finished. Also, the step probability function has to be tuned in size too (this can be done by multiplying the covariance by a scaling factor) in order for the sampler to get an appropriate acceptance rate for the steps, which has been found to be optimal at around a 23.4% [47]. Typically, a Metropolis-Hastings has at least a number of points on the order of 10^6, and that number has to increase when the parameter space is highly non-Gaussian, which makes a Gaussian step less efficient to explore the space.

Another thing to worry is the starting position of the chain because it takes some steps to arrive at the 'bulk' of the distribution, after being initialized on the outskirts of it. These points are not representative of the posterior distribution. So, in order to mitigate this effect, the first ≈ 500 points in the chain should be discarded, since they do not follow (or get to) the distribution just yet. This process is called *burn-in*.

To improve time efficiency, multiple computer cores or nodes can run multiple chains simultaneously, as this algorithm is completely parallelizable, as one can obtain a larger chain by adding each individual core-chain. One can take even more advantage of this if one finds a posterior that has multiple modes, considering that several chains might converge to different modes and stay there long enough to not explore the whole space during the given steps for each one.

Finally, in terms of the stopping criterion of the algorithm, one must check that the chain has converged. It is difficult to determine exactly when this happens, so one must be cautious and compute long chains, letting it travel several times over the posterior and also check that the chain is in a stationary state. Nonetheless, there are more quantitative ways to asses convergence, like performing a Gelman-Rubin diagnostic [48]. This is a method that contrasts the variance of the parameters inside chains to the *variance across chains*. This method, of course, needs us to have run several chains, which is typically done while parallelizing runs. The method can tell whether the chains have converged or not by checking if these variances are close to being equal, as would happen with long chains that started in different places in the parameter space.

4.3   Forecasting an experiment

One of the most interesting applications of the Bayesian theory is the possibility of forecasting the constraining power of upcoming experiments, and cosmology is a good field to apply these methods since cosmological surveys' specifications are publicly announced in advance. Forecasting in cosmology consists in taking a fiducial cosmology, like the best-fit ΛCDM parameters from a past survey, and simulate data accordingly with an approximation on the expected experimental uncertainties (in the form of a data covariance matrix) and perform Bayesian inference on this *mock dataset*, just as it was a real one. An example of this is on Chapter 5, where one can use a Boltzmann solver to obtain the theoretical matter power spectrum, and compute from it the observed galaxy power spectrum with the experimental specifications of a survey.

However, this approach has the emerging problem of it potentially being too time-consuming, as it could take weeks for a desktop computer to run an MCMC algorithm on the order of 10^6 samples. We bring up the desktop computer case since a person/team (without access to large computing facilities) outside the scientific committee of the survey might want to make forecasts on the survey's capabilities. A good idea might be to use an approximation (like in most physics branches) to the posterior PDF, as the whole idea of forecasting is fundamentally a science-based guess on the future constraints on the parameters of a cosmology.

4.3.1 The Fisher Matrix method

In order to forecast an experiment, we first set broad flat priors on the parameters of the model, in order to only be concerned with the computation of the likelihood. In order to forget several normalization constants, we prefer to work with the log-likelihood $\mathcal{L}(\boldsymbol{\theta}) \equiv \log(L(\mathbf{x}|\boldsymbol{\theta}))$. We start by Taylor-expanding the log-likelihood over the parameters up to second order around the maximum likelihood (ML) point $\boldsymbol{\theta}_{\mathrm{ML}}$,

$$\mathcal{L}(\boldsymbol{\theta}) \approx \mathcal{L}(\boldsymbol{\theta}_{\mathrm{ML}}) + \frac{1}{2} \left.\frac{\partial^2 \mathcal{L}}{\partial \theta^i \partial \theta^j}\right|_{\boldsymbol{\theta}_{\mathrm{ML}}} \cdot \Delta\theta^i \Delta\theta^j \ , \tag{4.9}$$

where the parameter distance from the ML point is $\Delta\theta^i = \theta^i - \theta^i_{\mathrm{ML}}$. With this equation, we can obtain an approximation to the likelihood after finding the maximum likelihood point by just computing a few derivatives, which is an enormous improvement over the more expensive MCMC methods. However, since we intend to do a forecast, we set the maximum likelihood cosmology as the fiducial one, and therefore we will only need to compute the second term of Eq. (4.9), and we can also neglect the zeroth-order log-likelihood, considering that it transforms to a normalization constant once we revert to the likelihood. In order to explore more characteristics on this approximation, we need to define the Fisher matrix.

The Fisher matrix (FM) was originally a frequentist quantity, defined by the mean value of the Hessian matrix, or

$$F_{ij} \equiv -\left\langle \frac{\partial^2 \mathcal{L}}{\partial \theta^i \partial \theta^j} \right\rangle \qquad \text{(Frequentist)} \ . \tag{4.10}$$

This mean is a data average. If we were dealing with real data, we should then take an ensemble average like we mentioned on the 2-point correlation function, Eq. (3.2). On the other hand, the Fisher Matrix used in cosmological forecasting applications is essentially different, as it is defined as the Hessian matrix evaluated at the ML point,

$$F_{ij} \equiv -\left.\frac{\partial^2 \mathcal{L}}{\partial \theta^i \partial \theta^j}\right|_{\boldsymbol{\theta}_{\mathrm{ML}}} \qquad \text{(Bayesian)} \ . \tag{4.11}$$

This equation is Bayesian in the sense that there is only one dataset to do inference, and hence it makes no sense to use a data-average. Continuing, we can easily implement the Bayesian Fisher matrix into the Taylor approximation, via

$$\mathcal{L}(\boldsymbol{\theta}) \approx \mathcal{L}(\boldsymbol{\theta}_{\mathrm{ML}}) - \frac{1}{2} F_{ij} \Delta\theta^i \Delta\theta^j \ . \tag{4.12}$$

Nevertheless, as noted in Ref. [11], in the case of forecasting, the two Fisher matrix definitions coincide, by considering that the likelihood is Gaussian, i.e. it follows Eq. (4.4). Specifically, if we average over many realizations on the production of mock datasets, it can be proven that the frequentist Fisher matrix coincides with its Bayesian counterpart [49]. The Fisher matrix for a Gaussian likelihood with mean $\boldsymbol{\mu}(\boldsymbol{\theta})$ and covariance $\mathbf{C}(\boldsymbol{\theta})$ is [41, 11]

$$\frac{1}{2} \mathrm{Tr}\left[\mathbf{C}^{-1}\mathbf{C}_{,i}\mathbf{C}^{-1}\mathbf{C}_{,j} + \mathbf{C}^{-1}\mathbf{M}_{ij}\right] \ , \tag{4.13}$$

where the commas represent partial derivatives with respect to the parameter that corresponds to the index. We also define $\mathbf{M}_{ij} \equiv \boldsymbol{\mu}_{,i}\boldsymbol{\mu}_{,j}^T + \boldsymbol{\mu}_{,j}\boldsymbol{\mu}_{,i}^T$.

Apart from its numerical convenience, the FM-approximated likelihood distributes as a Gaussian on the parameters, with a parameter covariance matrix equal to the inverse of the Fisher matrix, and centered on the ML point. Then, probably the most important property of this method, is the direct approximation of the parameter errors, via

$$\sigma_{\theta i}^2 = [\mathbf{F}^{-1}]_{ii} \; , \tag{4.14}$$

There are several other useful properties for the Fisher matrix, and we quote a good resume of them: "The 5 golden rules of *fisherology*", as portrayed in Ref. [11],

1. To *transform* variables, multiply the Fisher matrix on the right and on the left by the transformation Jacobian.

2. To *maximize* over some parameters, remove from the matrix the rows and the columns related to those parameters.

3. To *marginalize* over some parameters, remove from the inverse matrix the rows and the columns related to those parameters (being careful about the numerical instability pointed out above).

4. To *combine* Fisher matrices from independent experiments with the same fiducial model, sum the corresponding Fisher matrices, ensuring the same order of parameters, and, if necessary, inserting rows and columns of zeros for unconstrained parameters.

5. The *ellipsoidal confidence regions* have semiaxes lengths equal to the square root of the eigenvalues of the *inverse* Fisher matrix, while the semiaxes are oriented along the corresponding eigenvectors. The *area* of the ellipse (or volume of ellipsoid) is proportional to the square root of the determinant of the inverse Fisher matrix. The determinant of the Fisher matrix is an indicator of performance or a figure of merit.

The most important property of the above (for our subsequent work) is the property of combination of Fisher matrices, given that galaxy clustering forecasts separate the galaxy power spectrum into several redshift bins. Consequently, we can sum the Fisher matrices of each bin to obtain the total Fisher matrix for the complete survey.

4.4   Bayesian Model Comparison

In this section, we take on a more complex topic than Bayesian parameter inference: Bayesian model comparison. This part is usually more challenging since we take everything done before, and add to it the possibility of several models. Before that, our first task for this section is to employ Bayes' theorem to find the posterior probability $\mathcal{P}(M|\mathbf{x})$ of having a model M given some data

$$\mathcal{P}(M|\mathbf{x}) = \mathcal{E}(\mathbf{x}|M) \frac{\pi(M)}{\pi(\mathbf{x})} \; . \tag{4.15}$$

Here, $\mathbf{x} = (x^1, x^2, ..., x^N)$ is again the N random data, and now $\boldsymbol{\theta}^M = (\theta^1, \theta^2, ..., \theta^n)$ are the n theoretical parameters of the model M. This equation is the *big picture* of this section since it embodies the seeking of the probability of a certain model to be true.

Before the proper model comparison, we go a bit deeper on Eq. (4.15) by also writing the Bayes' equation following Eq .(4.3), but expliciting the presence of the given model M,

$$\mathcal{P}(\boldsymbol{\theta}\,|\,\mathbf{x}, M) = L(\mathbf{x}\,|\,\boldsymbol{\theta}^M, M)\,\frac{\pi(\mathbf{x}\,|\,M)}{\mathcal{E}(\mathbf{x}\,|\,M)}\ . \tag{4.16}$$

Here, the likelihood function $L(\mathbf{x}\,|\,\boldsymbol{\theta}^M, M)$, the prior probability $\pi(\boldsymbol{\theta}^M\,|\,M)$, the posterior $\mathcal{P}(\boldsymbol{\theta}\,|\,\mathbf{x}, M)$ now refer to the specific model M. Unlike we did in Sec. 4.2, we see the evidence as the most important quantity to characterize the probability of a model,

$$\mathcal{E}(\mathbf{x}\,|\,M) = \int L(\mathbf{x}\,|\,\boldsymbol{\theta}^M, M)\,\pi(\boldsymbol{\theta}^M\,|\,M)\,\mathrm{d}\theta^M\ . \tag{4.17}$$

We use Eq. (4.15) with two models M_1 and M_2: we compare them by taking the ratio of their probabilities:

$$\frac{\mathcal{P}(\,M_1\,|\,\mathbf{x})}{\mathcal{P}(\,M_2\,|\,\mathbf{x})} = \mathcal{B}_{12}\frac{\pi(M_1)}{\pi(M_2)}\ , \tag{4.18}$$

where $\pi(M_i)$ is the prior on each model, and we defined the Bayes factor via

$$\mathcal{B}_{12} = \frac{\mathcal{E}(\mathbf{x}\,|\,M_1)}{\mathcal{E}(\mathbf{x}\,|\,M_2)}\ . \tag{4.19}$$

For most cases, we can consider the prior probability of each model to be equal: if we have m contesting models, each one would have a weight of $1/m$. That way, the Bayes factor is the decisive quantity in Bayesian model comparison. However, this factor confronts only pairs of models, so it is more practical to rank the evidences of as many models as wanted to confront. Once one obtains the Bayes factors of two models, one can have a first guess on the severity of the difference by looking it up on the Jeffrey's scale on Table 4.1, which is a merge between quantitative and qualitative rankings on the Bayes factor, which is meant to not be taken too seriously.

Table 4.1: Jeffrey's Scale as in Ref. [1]. $\mathcal{B}_{12}$ is the Bayes factor of two models M_1 and M_2. The different levels represent different degrees of belief in that M_1 is the true theory.

$\log(\mathcal{B}_{12})$	Probability	Evidence
$0 \le \log(\mathcal{B}_{12}) < 1.0$	$0 \le P_1 < 0.75$	Inconclusive
$1.0 \le \log(\mathcal{B}_{12}) < 2.5$	$0.75 \le P_1 < 0.923$	Weak
$2.5 \le \log(\mathcal{B}_{12}) < 5.0$	$0.923 \le P_1 < 0.993$	Moderate
$5.0 \le \log(\mathcal{B}_{12})$	$0.993 \le P_1$	Strong

Obtaining the evidence is usually a difficult task like the posterior when doing Bayesian inference. As we mentioned, most cosmological models have several parameters (physical and nuisance), usually in the range $5 < n < 15$, but sometimes it can get up to hundreds of parameters. As a consequence, there are several methods to do a more efficient posterior probability integration. We now outline the (by far) most popular method for evidence calculation in cosmology, the Nested sampling algorithm [50].

4.4.1 Nested Sampling

Metropolis-Hastings and many other MCMC samplers are typically not good to compute evidences since they do not explore low probability areas at all, and most likely also missing information if the model is multi-modal. In response to that, Nested sampling is the most used evidence MCMC-integrator in cosmology[6]. We proceed to explain the framework of regular Nested Sampling, for means of simplicity.

Nested sampling works by transforming the parameter integral to an integral over the *enclosed prior mass X*, defined through

$$X(\lambda) = \int_{L>\lambda} \pi(\boldsymbol{\theta}) \mathrm{d}^n \theta \ . \tag{4.20}$$

where λ is a variable that represents a likelihood-level cut. Now, considering Eq. (4.5), we can rewrite it using X,

$$\mathcal{E}(\mathbf{x}) = \int_0^1 L(X) \mathrm{d}X \ . \tag{4.21}$$

This is a wonderful way of transforming a multidimensional integral into a one-dimensional one. Then, calculating this integral can be done as follows (in the most simple form):

The algorithm stores a chain of likelihoods L_i which starts at a value zero ($L_0 = 0$), a chain of prior masses X_i that starts at a value 1 ($X_0 = 1$), and a chain of weights w_i. The weights are the numerical differences in the prior mass in order to obtain the numerical integration of Eq. (4.21). This weight can then be set to $w_i = X_{i-1} - X_i$, where we consider that the prior mass is decreasing as we increase the likelihood. The integral is done by summing $w_i L_i$ to the stored evidence from the last step. Then, we can approximate and constrain the evidence by

$$\mathcal{E} \lessgtr \left(\sum_{i=1}^s w_i L_i \right) + L_{\max} X_s \ , \tag{4.22}$$

where s is the number of steps and $L_{\max}$ is the maximum likelihood. The right hand side of the equation represents an optimal additional step, such that $X_{s+1} = 0$ and $L_{s+1} = L_{\max}$.

At each step of the algorithm, we have a batch of b points in the parameter space (this defines the batch $\mathbf{B} = \{\boldsymbol{\theta}_{(1)}, ..., \boldsymbol{\theta}_{(b)}\}$). These points have been sampled from the prior in such a way that the likelihoods of these points are larger than L_{i-1}. The algorithm then chooses the parameter point of less likelihood in the batch to be set as the next point of the likelihood chain. Therefore, this process obtains a sample which is 'nested' in the likelihood-cut level from the last step.

Another interesting characteristic of the algorithm is the assignation of the prior mass X_i. The trick is to assume[7] the analogy of the process of sampling a Uniform(0,1) variable t_i so

---

[6]Nested sampling is usually used in the form of the MultiNest algorithm [51, 52], an improved version of Nested Sampling, which is more widely used in cosmology.

[7]This comes from the qualitative behavior of most evidences, where the sampling in X advances in a logarithmic fashion when the likelihood is Gaussian. For a more in-depth discussion see Ref. [50].

that the new point would be $X_i = t_i X_{i-1}$, to the process of sampling a parameter point from the prior probability so that $L_i > L_{i-1}$. Specifically, this is implemented by comparing the process of selecting the point of less likelihood from the batch, to selecting the larger value in the sampling of b Uniform variables in the range $(0, X_{i-1})$. Statistically, this progression in the prior mass can be shown to go as $X_i = \exp(-i/b)$.

Wrapping all this up, the algorithm is as follows:

Nested Sampling Algorithm

1. First, we start by computing the complete batch of b points from the prior probability. We also initialize the evidence $\mathcal{E} = 0$, the likelihood $L_0 = 0$ and the enclosed prior mass $X_0 = 1$.

2. Then, select the point of less likelihood $L_i = \min(L(\boldsymbol{\theta}_{(b)}))$ and set $X_i = \exp(-i/b)$ and the weight $w_i = X_{i-1} - X_i$, also increase $\mathcal{E}$ by $L_i w_i$.

3. Remove the point of less likelihood from the batch and replace it with a new sample with a higher likelihood (in the nested likelihood region), this is $L(\boldsymbol{\theta}_{(\text{new})}) > L_i$. The rest of the points do not need to be sampled again since they all fulfill the minimum likelihood requirement.

4. Repeat steps 2 and 3 for a fixed amount of steps or until a stopping criterion is triggered, such as a negligible increment on the likelihood, or considering that even the maximum likelihood of the batch would not increase the evidence by a given threshold.

5. (Optional) Finally, after the run ends, increase the evidence by

$$\frac{1}{b} \sum_{\boldsymbol{\theta} \in \mathbf{B}} L(\boldsymbol{\theta}) X_s \ . \tag{4.23}$$

This mean batch-likelihood is to approximately sum the gap mentioned on Eq. (4.22).

Nested sampling is usually fast, and needs a batch size of around $100 - 500$ to have good accuracy. The algorithm also produces a chain of parameters (the ones from each likelihood point $L_i = L(\boldsymbol{\theta}_i)$) that follow the posterior PDF, and this chain usually beats Metropolis-Hastings in terms of speed of convergence.

As a closing remark, we must mention that MultiNest improves on Nested sampling in that it approximates the nested likelihood regions with ellipsoidal nested regions. Additionally, it allows for multiple ellipses, in order to tackle problems of multimodal evidence computation, plus several other features. See [52] for detailed improvements upon Nested sampling.

Chapter 5

Non-Gaussian galaxy clustering LSST forecasts

This chapter presents a research work that is soon to be published, and will closely follow the following structure. We repeat some definitions on the observed power spectrum with some slight additions, in order to integrate the survey specifications in a straightforward way.

Abstract

We implement the Derivative Approximation of LIkelihoods (DALI) method to make forecasts of an LSST-like galaxy clustering survey for wCDM. DALI is an extension to the traditional Fisher Matrix approach, which can obtain a non-Gaussian approximation to the posterior probability density of the parameters of a theory. We present the results via the confidence regions for the parameters, which are compared to Markov Chain Monte Carlo samples. We found good fittings for the DALI method in terms of the confidence regions' size for all the redshift bins of the survey. At the same time, FM regions are vastly overestimated for the first redshift bins, covering wide non-physical regions. For higher redshift bins, neither the FM nor DALI obtains precise sizes for the region. The whole analysis was limited by numerical precisions, especially for the higher redshift bins, masking the potential of DALI.

5.1 Introduction

The large-scale distribution of galaxies embeds useful information about our Universe, especially related to its matter and energy content, and their dynamics. An important part of our current understanding of the Universe is dark energy, a useful concept to explain the current state of accelerated expansion, which was first found through Supernovae of type Ia observations [13, 14]. Dark energy is usually modeled in its simplest form as a fluid of negative pressure, and little is known about its fundamental nature, thus, a better explanation is

needed in order to construct a satisfactory model for our Universe, see Ref. [19] for a review on the topic. Consequently, it is important to predict if future surveys would be able to provide the data necessary to discern between contestant models and to constrain the values for their parameters.

Considering that larger and more precise galaxy surveys are in construction, like LSST [17] or Euclid [18], it is imperative to forecast how they might perform in constraining cosmological parameters, as these experiments will collect orders of magnitude of data than past generations. These surveys will map billions of galaxies along a wide redshift range and area coverage, probing the correlation function of galaxies and its associated power spectrum with unprecedented accuracy. These quantities portrait the physical information content from galaxies, as they can be related to cosmological models and their parameters.

Parameter constraining forecasts consist in obtaining confidence regions for the parameters of a cosmological model that describes the Universe. For this, mock data that mimics the survey's specifications are produced, following the reference or fiducial cosmology assumed. Also, one assumes that the posterior probability (or the likelihood, if one uses flat priors) of the parameters is well described by a Gaussian distribution. The inverse of the covariance matrix of this normal distribution is the Fisher matrix [53]. However, valuable information on the shape of the confidence regions is lost when performing the FM approximation, because it approximates the confidence regions to ellipses. The Derivative Approximation of LIkelihoods (DALI) [49, 54, 55] is an extension to the FM that includes derivatives of higher orders than those used in the FM approximation, enabling it to capture shapes beyond the ellipses produced by FM forecasts. In this work, we will apply the DALI method to forecast an LSST-like galaxy clustering survey. For that purpose, we utilize the CLASS Boltzmann solver [34, 35] to obtain the matter power spectrum and other background quantities.

The present work is structured as follows: In Sec. 5.2, we outline the theoretical considerations used to derive the observed galaxy power spectrum formulas. In Sec. 5.3 we present the FM and DALI approximations and their implementation for a galaxy clustering survey. In Sec. 5.4 we state the cosmological model and its considered parameters, plus the survey specifications, and other numerical considerations of interest, especially on the MCMC sampling. Next, in Sec. 5.5 we present the results obtained and corresponding analysis and, finally, we present our conclusions in Sec. 5.6.

5.2 The Observed Galaxy Power Spectrum

The observed power spectrum has several differences with the actual matter power spectrum. To start, one measures the light of galaxies, which can be tracers of the overall matter content, plus several effects from the peculiar velocities of galaxies in clusters, the change of the reference cosmology, among others. We now introduce the formula for the observed power spectrum [32, 56, 39] and briefly explain their associated effects:

$$P_{\text{obs}}(k, \mu, z) = \frac{d_A(z)_{\text{ref}}^2 H(z)}{d_A(z)^2 H(z)_{\text{ref}}} \Big(b(z) + f(z)\mu_{\text{AP}}^2 \Big)^2 P(k_{\text{AP}}, z) \exp\Big(-k_{\text{AP}}^2 \mu_{\text{AP}}^2 \sigma_r\Big) + P_{\text{shot}}(z) \ . \ (5.1)$$

The observed power spectrum is not independent of direction, since it depends on μ, the cosine of its direction to the line of sight. For simplicity, we now list the considered effects bounded to each multiplying factor, along with the definition of the variables not mentioned so far, see [57] for a more in-detail derivation of this formula.

- The first multiplying term, $\left(\frac{d_A(z)_{\mathrm{ref}}^2 H(z)}{d_A(z)^2 H(z)_{\mathrm{ref}}}\right)$, corresponds to the Alcock-Paczynski effect [58], which considers the change of volume when changing the cosmological parameters, due to the need of a cosmology to convert redshifts and angles to distances. The quantities $d_A(z)$ and $H(z)$ are in Mpc/h and h/Mpc units, respectively[1], with h the adimensional Hubble constant in the respective cosmology. It is defined by $100\,h\,\mathrm{km/s/Mpc} = H_0$, where H_0 is the Hubble constant in km/s/Mpc units. The subscript 'ref' denotes that the quantity is evaluated at a reference cosmology. In addition, k, μ are valued on the reference cosmology too, so they are affected by the Alcock-Paczynski effect as well, and we denote it via the 'AP' subscript. We can write $k_{\mathrm{AP}}, \mu_{\mathrm{AP}}$ explicitly:

$$k_{\mathrm{AP}}(k, \mu) = R_{\mathrm{AP}} k \tag{5.2}$$

$$\mu_{\mathrm{AP}}(k, \mu) = \frac{H(z)\mu}{H(z)_{\mathrm{ref}} R_{\mathrm{AP}}} \ , \tag{5.3}$$

where

$$R_{\mathrm{AP}} = \frac{\sqrt{H(z)^2 d_A(z)\mu^2 - H(z)_{\mathrm{ref}} d_A(z)_{\mathrm{ref}}(\mu^2 - 1)}}{H(z)_{\mathrm{ref}} d_A(z)} \ . \tag{5.4}$$

- The second term, $\left(b(z) + f(z)\mu_{\mathrm{AP}}^2\right)^2$, is the Kaiser formula [59], which accounts for the velocities in the line of sight of galaxies as they fall into clusters, which cause distortions in the observed redshift. This distortions in redshift space, where galaxies are measured appear in the power spectrum on real space. Here, we also considered the bias $b(z)$, defined as the ratio of galactic to total matter, and we will leave it as part of the survey's specifications, for it usually depends on the type of galaxy considered by the survey and can have certain redshift parametrizations motivated from simulations.

- Next, the exponential term, $\exp(-k_{\mathrm{AP}}^2 \mu_{\mathrm{AP}}^2 \sigma_r^2)$, is the suppress of power thanks to the uncertainty in redshift, since it affects the position in the line of sight where galaxies are detected. σ_r is the error in real space (or positional uncertainty), and relates to the error in redshift via

$$\sigma_r = \frac{\sigma_z}{H(z)} \ , \tag{5.5}$$

with σ_z the error in redshift, given by

$$\sigma_z^2 = (1+z)^2(\sigma_{0\gamma z}^2 + \sigma_{0v}^2) \ , \tag{5.6}$$

where $\sigma_{0\gamma z}$ is the photometric redshift measurement error, and σ_{0v} is an error from the intrinsic velocity dispersion of galaxies. In this work we will consider $\sigma_z \approx (1+z)\sigma_{0\gamma z}$, as σ_{0v} can be considered negligible compared to $\sigma_{0\gamma z}$ [2].

- The last term is the baseline instrumental shot noise $P_{\mathrm{shot}}(z)$ of the survey. We will assume no dependence on the cosmological parameters for it and, going further, we will neglect this nuisance term in our analysis.

---

[1]This is possible by setting the speed of light to $c = 1$ and dividing H by it.

5.3   Likelihood, Fisher Matrix, and DALI formalism

In a galaxy clustering survey, the data are the Fourier transforms of the modes, $\delta_i = \delta_{\mathbf{k}_i}$, which have mean 0 by construction. Their distribution is assumed Gaussian, with an associated covariance matrix equal to the power spectrum plus a Poisson noise term [11]

$$C_{ij} = \Delta_i^2 \delta_{ij} = (P_i + 1/\bar{n})\delta_{ij} , \qquad (5.7)$$

where δ_{ij} is a Kronecker delta. The Poisson noise represents the probability distribution of a random ensemble of galaxies. It arises from the consideration of a window smoothing of the galaxy counts to construct a continuous density function. This covariance matrix is diagonal, and it relates the power spectrum $P_i = P(\mathbf{k}_i)$ to the mean galaxy density $\bar{n}$ of each redshift bin. Technically, $\bar{n}$ is the selection function of the survey, which is an estimator of the number density of galaxies that the survey will measure, allowing for considering that a fraction of galaxies can be measured, be it due to its brightness or color or type. It is for that reason that it will be considered one of the survey parameters later on and, as such, it will not depend on the cosmological parameters.

The forecasts will be under the framework of [54], where the Gaussian likelihood in the data is approximated only by using the derivatives of the covariance matrix $\mathbf{C}$ with respect to the cosmological parameters $\boldsymbol{\theta}^2$. For the approximation, the data needs o have a null mean, and the density modes δ_i fulfill this requirement. Data without this mean can be easily transformed via subtraction of their mean.

Continuing, the Gaussian likelihood to be approximated reads

$$L = \frac{A}{\sqrt{\det \mathbf{C}}} \exp\left(-\frac{1}{2}\mathbf{x}^t \mathbf{C}^{-1} \mathbf{x}\right) . \qquad (5.8)$$

For a galaxy clustering case, it becomes

$$L_{\mathrm{GC}} = \frac{A}{\prod_i \Delta_i} \exp\left(-\frac{1}{2}\sum_i \frac{\delta_i^2}{\Delta_i^2}\right) . \qquad (5.9)$$

In order to move on to forecasting, we consider that the quantities in the likelihood need to be the mean of many mock data realizations in order to make a forecast free of random errors. This can be expressed as $L^{\mathrm{Forecast}} = \langle L \rangle$. Since we know the distribution of the data we are forecasting, we can average the squared density modes,

$$\langle \delta_i^2 \rangle = \Delta_i^2|_{\mathrm{ML}} = C_{ii}|_{\mathrm{ML}} , \qquad (5.10)$$

where the subscript 'ML' indicates that the quantity is evaluated in the maximum likelihood point, which is going to be a fiducial (or reference) cosmology. This last consideration is key, since not fixing the reference cosmology for the data would mean we are producing new dataset whenever evaluating the likelihood in different points, sabotaging the effort of forecasting. Hence, the forecasted log-likelihood will be

$$\mathcal{L} = -\frac{1}{2}\sum_i \log\left(\Delta_i^2\right) + \frac{\Delta_i^2|_{\mathrm{ML}}}{\Delta_i^2} , \qquad (5.11)$$

---

[2]In other types of problems, the data $\mathbf{x}$ can be the parameter-dependent quantity.

where $\mathcal{L} \equiv \log\left(L_{\rm GC}^{\rm Forecast}\right)$. From now on we will drop the labels 'GC' and 'Forecast' for brevity. The Fisher Matrix of a likelihood that depends on the parameters via its covariance matrix is given by [11]

$$F_{\alpha\beta} = -\left\langle \frac{\partial^2 \mathcal{L}}{\partial\theta^\alpha \partial\theta^\beta} \right\rangle \tag{5.12}$$

$$= \frac{1}{2} \text{Tr}\left[\mathbf{C}_0^{-1}\mathbf{C}_{,\alpha}\mathbf{C}_0^{-1}\mathbf{C}_{,\beta}\right] \tag{5.13}$$

$$= \frac{1}{2} \sum_{\mathbf{k}_i} \frac{P_{obs,\alpha}\, P_{obs,\beta}}{(P_{obs} + 1/\bar{n})^2} \,, \tag{5.14}$$

where we denoted $\mathbf{C}_0 \equiv \mathbf{C}|_{\rm ML}$. We also denoted partial derivatives respect to a parameter by a comma and the index of the parameter. Additionally, we must consider that the observed power spectrum and their derivatives are evaluated at the reference cosmology. The last passage was possible by remembering that the covariance matrix is diagonal. We invoked the wavevectors $\mathbf{k}_i$ in the last step because we can transform the sum of Eq. (5.14) into an integral over the Fourier volume, an integral on k and the cosine μ

$$F_{\alpha\beta} = \frac{V_{\rm bin}}{8\pi^2} \int\limits_{k_{\rm min}}^{k_{\rm max}} \int\limits_{-1}^{1} d\mu dk \frac{P_{obs,\alpha}\, P_{obs,\beta}}{(P_{obs} + 1/\bar{n})^2} k^2 \,. \tag{5.15}$$

Here, $k_{\rm max}$ and $k_{\rm min}$ are the maximum and minimum wavenumbers that the survey can measure. Also, $V_{\rm bin} = f_{\rm sky}V_{\rm shell}$ is the survey redshift bin volume in the reference cosmology, since we are integrating with the fiducial cosmology wavenumber k and cosine μ. Other way of seeing $V_{\rm bin}$, is considering it the expected bin volume, in a similar manner as the expected number density $\bar{n}$. The bin volume is equal to the fraction of the sky $f_{\rm sky}$ covered by the survey, multiplied by the volume of the redshift bin shell $V_{\rm shell}$, given by

$$V_{\rm shell} = \frac{4\pi}{3} \left(d_M(z)\right)^3 \Big|_{z_{\rm min}}^{z_{\rm min}} \,, \tag{5.16}$$

where $d_M(z) = (1+z)d_A(z)$ is the transverse comoving distance and $z_{\rm min}, z_{\rm max}$ are the ranges of the redshift bin. The log-likelihood in the FM approximation reads (with Einstein's sum notation)

$$\mathcal{L}_{\rm FM} = -\frac{1}{2}F_{\alpha\beta}\Delta\theta^\alpha \Delta\theta^\beta \tag{5.17}$$

The DALI method is an extension that goes beyond the ellipsoidal regions in the parameter space obtained from Fisher forecasts, potentially capturing *banana*, *boomerang* or *potato* shapes that are a common result in cosmology. The method features an expansion that is positive definite in the whole parameter space. In other words, by construction, it cannot yield negative probabilities. We use the approximation from Ref. [54], which features a log-likelihood expressed by

$$\begin{aligned}
\mathcal{L}_{\rm D2} = -\,&\frac{1}{4}D_{\alpha\beta}\Delta\theta^\alpha \Delta\theta^\beta \\
-\,&\frac{1}{4}D_{\alpha\beta\gamma}\Delta\theta^\alpha \Delta\theta^\beta \Delta\theta^\gamma \\
-\,&\frac{1}{16}D_{\alpha\beta\gamma\delta}\Delta\theta^\alpha \Delta\theta^\beta \Delta\theta^\gamma \Delta\theta^\delta \,,
\end{aligned} \tag{5.18}$$

where 'D2' stands for DALI-doublet. This denomination comes as we are considering derivatives of first and second order. There exists the DALI-triplet approximation that includes third-order derivatives as well, but we opted to be conservative in this sense. The DALI tensors of second, third and fourth rank, respectively, are defined by

$$D_{\alpha\beta} = \text{Tr}\left[\mathbf{C}_0^{-1}\mathbf{C}_{,\alpha}\,\mathbf{C}_0^{-1}\mathbf{C}_{,\beta}\right] , \tag{5.19}$$

$$D_{\alpha\beta\gamma} = \text{Tr}\left[\mathbf{C}_0^{-1}\mathbf{C}_{,\alpha}\,\mathbf{C}_0^{-1}\mathbf{C}_{,\beta\gamma}\right] , \tag{5.20}$$

$$D_{\alpha\beta\gamma\delta} = \text{Tr}\left[\mathbf{C}_0^{-1}\mathbf{C}_{,\alpha\beta}\,\mathbf{C}_0^{-1}\mathbf{C}_{,\gamma\delta}\right] . \tag{5.21}$$

We can see that the DALI-tensor of second rank in Eq. (5.19) is twice the Fisher matrix, i.e. $2F_{\alpha\beta} = D_{\alpha\beta}$. Now, in a complete analogous mechanism to the FM case, we recast these expressions directly into integrals on the Fourier volume,

$$D_{\alpha\beta} = \frac{V_{\text{bin}}}{4\pi^2} \int\limits_{k_{\min}}^{k_{\max}} \int\limits_{-1}^{1} d\mu dk \frac{P_{obs,\alpha}\,P_{obs,\beta}}{(P_{obs} + 1/\bar{n})^2}k^2 , \tag{5.22}$$

$$D_{\alpha\beta\gamma} = \frac{V_{\text{bin}}}{4\pi^2} \int\limits_{k_{\min}}^{k_{\max}} \int\limits_{-1}^{1} d\mu dk \frac{P_{obs,\alpha}\,P_{obs,\beta\gamma}}{(P_{obs} + 1/\bar{n})^2}k^2 , \tag{5.23}$$

$$D_{\alpha\beta\gamma\delta} = \frac{V_{\text{bin}}}{4\pi^2} \int\limits_{k_{\min}}^{k_{\max}} \int\limits_{-1}^{1} d\mu dk \frac{P_{obs,\alpha\beta}\,P_{obs,\gamma\delta}}{(P_{obs} + 1/\bar{n})^2}k^2 . \tag{5.24}$$

It is worth noting that we need to work with the partial derivatives of the observed power spectrum respect to the cosmological parameters $\boldsymbol{\theta}$ directly, and not to other intermediate variables because it's not straightforward to marginalize over them in the DALI approximation, unlike in the Fisher Matrix approach.

During all this section we considered independent redshift bins. Therefore, in order to obtain the full log-likelihood, we need to sum the log-likelihoods of each bin. Analogously, to obtain the full FM and DALI approximations we need to sum the Fisher matrices and DALI-tensors, respectively.

5.4 Methodology

5.4.1 Cosmological parameters

The cosmology considered is a wCDM model, which corresponds to a Universe with a dark energy perfect fluid with equation of state parameter w, constant over redshift. The free parameters we consider are

$$\boldsymbol{\theta} = \{\Omega_c, \Omega_b, h, n_s, \ln\left(10^{10}A_s\right), w\} , \tag{5.25}$$

where Ω_c is the cold dark matter density parameter, Ω_b the baryon density parameter, h the adimensional Hubble constant, n_s and A_s are the spectral index and amplitude of primordial

perturbations, respectively, and w is the EoS parameter of dark energy just mentioned. Their fiducial values are

$$\boldsymbol{\theta}_0 = \{0.25, 0.05, 0.7, 0.96, 3.1, -1\} \, . \tag{5.26}$$

It is worth nothing that the choice of Ω_c and Ω_b over the measurable quantities $\Omega_c h^2$ and $\Omega_b h^2$ was made purely to improve the stability of the derivatives of the power spectrum, and so the confidence regions[3].

In addition to the free parameters, the parameters Ω_γ and Ω_n correspond to the density parameters of photons and relativistic neutrinos, respectively. They have fixed values of $\Omega_\gamma = 5\mathrm{e}^{-5}$ and $\Omega_n = 3.5\mathrm{e}^{-5}$. By demanding $1 = \Omega_c + \Omega_b + \Omega_{\mathrm{de}} + \Omega_\gamma$, we derive a value for the density parameter of dark energy, which ensures we remain in a flat Universe. Therefore, its fiducial value is $\Omega_{\mathrm{de,ML}} \approx 0.7$.

The priors for the analysis are flat, and they mean to represent the allowed physical regions for each parameter. Not every parameter needs these constraints and, consequently, for them, we consider priors broad enough that allow for the likelihood to cover the parameter space without restriction. We list these bounds on Table 5.1.

Lower limit	Parameter	Upper limit
0	Ω_c	1
0	Ω_b	1
0	h	-
-	n_s	-
-	$\ln\left(10^{10} A_s\right)$	-
-	w	$-1/3$

Table 5.1: Flat prior constrains. When not mentioned, we consider priors broad enough so that the likelihood alone creates the bound.

In order to compute the DALI-tensors, one needs the observed power spectrum of Eq. (5.1) and, for that purpose, as mentioned in the Introduction, we utilize the CLASS code[4] to solve the Boltzmann equations considering all the species in the Universe and also to obtain the functions $P(k, z)$, $d_A(z)$, $H(z)$ and $f(z)$, which also depend on the cosmological parameters $\boldsymbol{\theta}$. Any other cosmological parameter not mentioned in this section is left to the default CLASS' value.

5.4.2 Survey specifications

As mentioned before, we use an LSST-like survey with 7 independent redshift bins, as specified in [2]. The redshift bins and other survey parameters are shown in Table 5.2. This table

---

[3]If wanted, one can obtain the other set of parameters after getting the samples by using a parameter transformation on them.

[4]http://class-code.net/

has values for the mean of the redshift bin, z-mean, the range in redshift z-range$= z_{\min}$-$z_{\max}$, the expected number density per bin $\bar{n}$, and the limits of the wavenumber, $k_{\min}$ and $k_{\max}$.

z-mean	z-range	$\bar{n}\ [h^3\mathrm{Mpc}^{-3}]$	$k_{\min}\ [h/\mathrm{Mpc}]$	$k_{\max}\ [h/\mathrm{Mpc}]$
0.31	0.20-0.46	0.15	0.0071	0.08
0.55	0.46-0.64	0.10	0.0050	0.09
0.84	0.64-1.04	0.064	0.0040	0.11
1.18	1.04-1.32	0.036	0.0035	0.14
1.59	1.32-1.86	0.017	0.0030	0.17
2.08	1.86-2.30	0.0069	0.0028	0.23
2.67	2.30-3.00	0.0022	0.0026	0.31

Table 5.2: The LSST-like survey parameters associated to each redshift bin according to [2]. In addition to them, the survey has a sky coverage $f_{\mathrm{sky}} = 0.58$, a photometric redshift measurement error $\sigma_{0\gamma z} = 0.04$, and a bias $b(z) = 0.84z$.

Each quantity inside a bin, be it the power spectrum or any background function, is approximated to the mean redshift value.

5.4.3 MCMC sampling and FoM

After the computation of the FM and the DALI-tensors, and considering broad flat priors, we perform several Metropolis-Hastings Markov Chain Monte Carlo (MCMC-MH) [44, 45] runs of the unapproximated likelihood, which is explicitly parameter-dependent:

$$\mathcal{L}(\boldsymbol{\theta}) = -\frac{V_{\mathrm{bin}}}{8\pi^2} \int\limits_{k_{\min}}^{k_{\max}} \int\limits_{-1}^{1} d\mu\, dk\, \ell(\boldsymbol{\theta}) k^2 \ , \tag{5.27}$$

where we used

$$\ell(\boldsymbol{\theta}) \equiv \log\left(P_{obs}(\boldsymbol{\theta}) + 1/\bar{n}\right) + \frac{P_{obs}(\boldsymbol{\theta}_0) + 1/\bar{n}}{P_{obs}(\boldsymbol{\theta}) + 1/\bar{n}} \ . \tag{5.28}$$

This expression corresponds to Eq. (5.11), after transforming the sum of modes into an integral.

We produce a chain for each redshift bin and for the total survey. The purpose of this is to observe the performance of DALI at each bin since combining all of them might result in a loss of shape information. It's important to mention that MCMC sampling is much slower than DALI, but it can be parallelized in a more straightforward way in a computer cluster, the reason being that DALI only needs to compute the first and second order derivatives for the observed power spectrum at the reference cosmology, and store those power spectrum derivatives as interpolations to be called later by the integrals of Eqs. (5.22)-(5.24). On the other hand, the MCMC method relies on computing the order of 10^6 likelihood evaluations.

In order to optimize the exploration of the parameters, the step needed in the algorithm is set up to be Gaussian, with a covariance equal to the Fisher matrix corresponding to the

bin, times a scaling factor to get a proper acceptance rate of $\sim 23\%$ [47]. For the total survey, we use the full Fisher matrix as the covariance of the proposal step, scaled accordingly as well.

The approximated methods, FM and DALI, are also MCMC-sampled in order to obtain their confidence regions, with an overwhelming speed difference in comparison to the direct likelihood MCMC-sampling. With that in mind, we sample all the regions without considering the priors, and then we filter them by invoking the prior boundaries, discarding the points that fall off the priors. This procedure will be of importance in the next section. We must note, however, that the priors do not affect the direct MCMC sampling region since CLASS needs *physical values* in order to compute the *physical quantities* we need. The only exception to this rule is the w parameter, which is not physically bound to the upper value of $-1/3$, since that limit is to have an expanding Universe only, and is supported by CLASS on higher values than it.

In order to asses the size of the different regions obtained, we define a Figure of Merit (FoM) as **the logarithm of the area of the 95% confidence region** of each parameter pair in a given redshift bin, or

$$\mathrm{FoM}_{\alpha\beta,z} \equiv \log\Big(A_{95\%}(\theta^{\alpha}, \theta^{\beta}, z) \Big) \, , \tag{5.29}$$

where we explicitly named the parameters θ^{α} and θ^{β}, plus the redshift bin at z. This FoM illustrates the logarithm of the inverse constraining power of each method. Hence, using the FoM will allow us to compare the relative sizes of the FM and DALI approximations to the MCMC-MH regions. We choose the logarithm as it is fairer at ranking relative sizes than considering the ratio of areas[5].

5.4.4 CLASS high precision parameters

In order to compute the DALI tensors, we need to set extremely high precision parameters for CLASS. This is due to the necessity of having an accurate and densely-valued power spectrum in k, as the derivatives of the power spectrum had to be numerically robust for many modes in order to compute an integral in two dimensions. Additionally, the background quantities need to be interpolated to find accurate values to properly compute the AP effect.

The consideration of the AP effect is very important for the need for high precision parameters. The integrals for the DALI tensors depend on the reference cosmology, and the AP effect changes k and μ slightly inside the observed power spectrum. As CLASS produces a power spectrum in a grid of k values, we need to compute high-density interpolations on k in order to properly account for the AP effect.

Another cause for needing high precision parameters is the step for the numerical derivatives with respect to the cosmological parameters. In order to obtain adequate Fisher Matrix regions, we need to set a step which will vary from bin to bin. For high-constraining bins, a

---

[5]This is the same argument as for why the logarithmic evidence is used as opposed to the evidence in Bayesian model comparison.

more narrow step is better-suited to obtain the FM, going as small as a relative step of 10^{-4} of the parameters' reference values.

We provide in the Appendix 5.7 the specific values of precision parameters for the code.

The direct likelihood MCMC sampling, however, did not need such precision optimization, since the nature of the sampling 'averages' the lack of precision put into the power spectrum computation, allowing for faster exploration of the parameter space. Nonetheless, the DALI method was still faster than the MCMC runs, even when the latter was much more parallelized.

5.5 Results and Discussion

In this section, we report the plots obtained from the analysis. To be brief, we show only the *triangle plots* of the redshift bins 1, 4, 7 and the total survey. These plots contain the parameter confidence regions derived from the three different methods considered. The choice of showing those redshift bins is important to inspect the different regimes that the approximations exhibit.

Figs. 5.1 and 5.2 show the behavior of the approximations and the samples of the first redshift bin. We observe that the FM regions are vastly overestimated in comparison to the MCMC samples, covering non-physical areas that are forbidden by the prior limits. The difference between the two plots is the enforcement of the priors on top of the approximations in Fig. 5.2, which reduces considerably the extent of the confidence regions for the FM method. It is important to note that deleting the points of the FM-sample that go off the flat boundaries in one parameter often also constrains other parameters. One notable example of this is the w parameter, which has a marginalized probability density that goes down to a value of -8 in the 'prior-free' case, while it only goes to -4 with the priors on other parameters. The DALI approximation also has a larger area covered than the MCMC samples but to a lesser extent. We can see that DALI tends to have the shape of the samples, but stretched in a certain direction.

The Fisher matrix method is known (apart from its speed and simplicity) as a way to obtain the smallest possible covariance for the parameters any estimator can produce, which translates very roughly to obtaining the smaller confidence regions possible. The Cramer-Rao inequality is the theorem that represents this limit but, nevertheless, it was proven on Ref. [55] (in its Appendix B) that this inequality is valid only if the likelihood and its derivatives exist and are finite for all the parameter combinations. In the case of the first bin (and also the second which is not shown), this is not fulfilled as the FM spans unphysical regions, justifying the size of their confidence regions.

Next, Fig. 5.3 shows the regions obtained on the fourth redshift bin. Priors do not affect the regions noticeably as before. Here we can see that the FM regions are a lot less oversized than before, and the DALI regions are also much more fit to the sample-derived region. We do observe a certain running on the parameters, which may be caused by an extra mode in
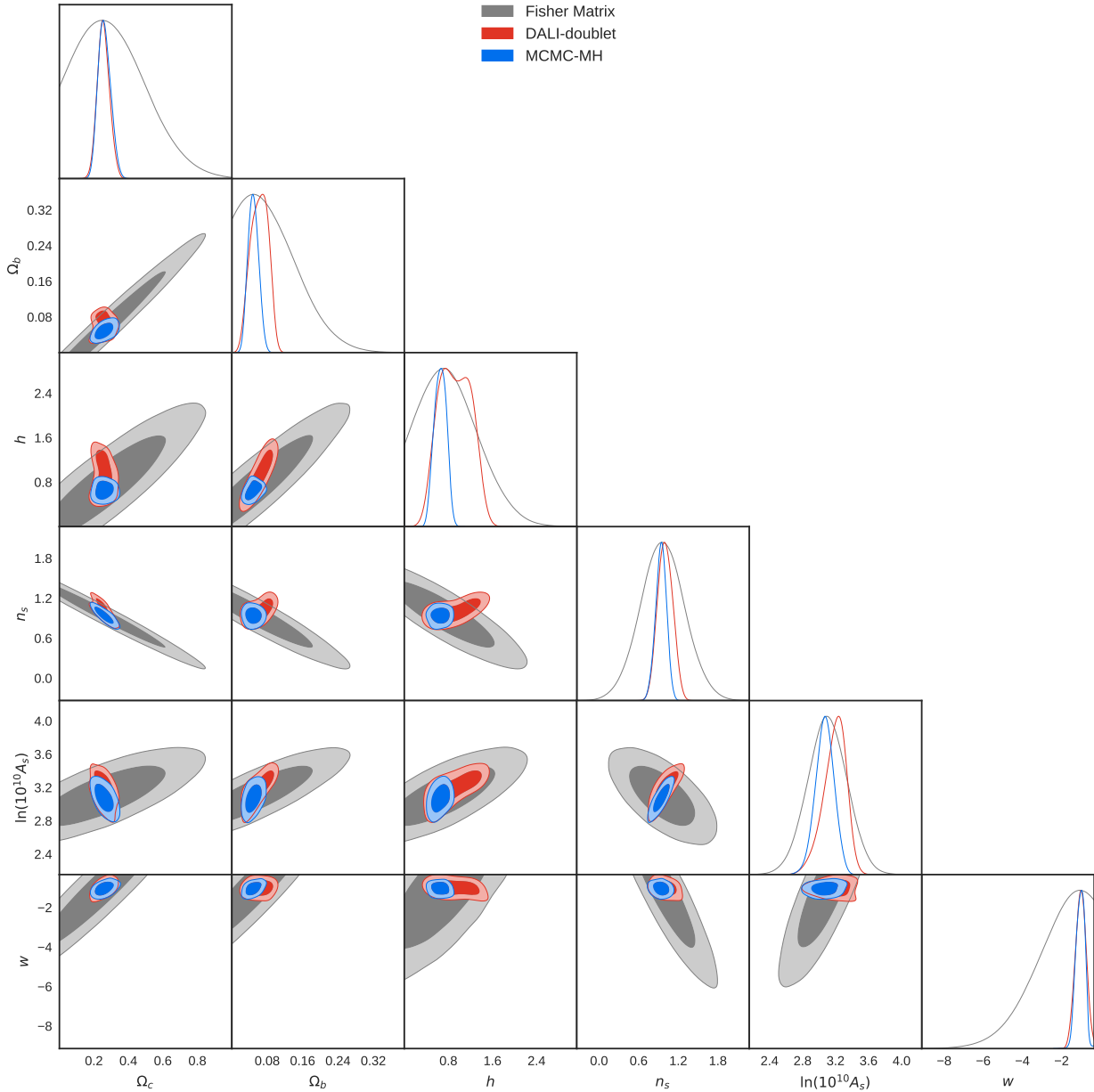
Figure 5.1: Confidence regions of the first redshift bin at $z = 0.31$ without priors.

the distribution that was not properly explored, or was explored only by some of the *walkers*[6] of the chain, as there were 20 of them. The FM regions are slightly larger than the MCMC samples while not exploring unphysical regions.

We can justify the size of the FM regions on redshift bins 3 through 5 by considering that there are parameter combinations that make the solver crash. We denominate these parts of the parameter space as 'uncomputable regions', which serve the same purpose as unphysical regions at proving that the Cramer-Rao inequality does not apply *strictly*, since the incidence of these uncomputable parameters is slight, especially on the latter bins.

---

[6] A walker in an MCMC run is an individual chain produced by a single core of a computer. The total chain is obtained by concatenation of each walker chain.
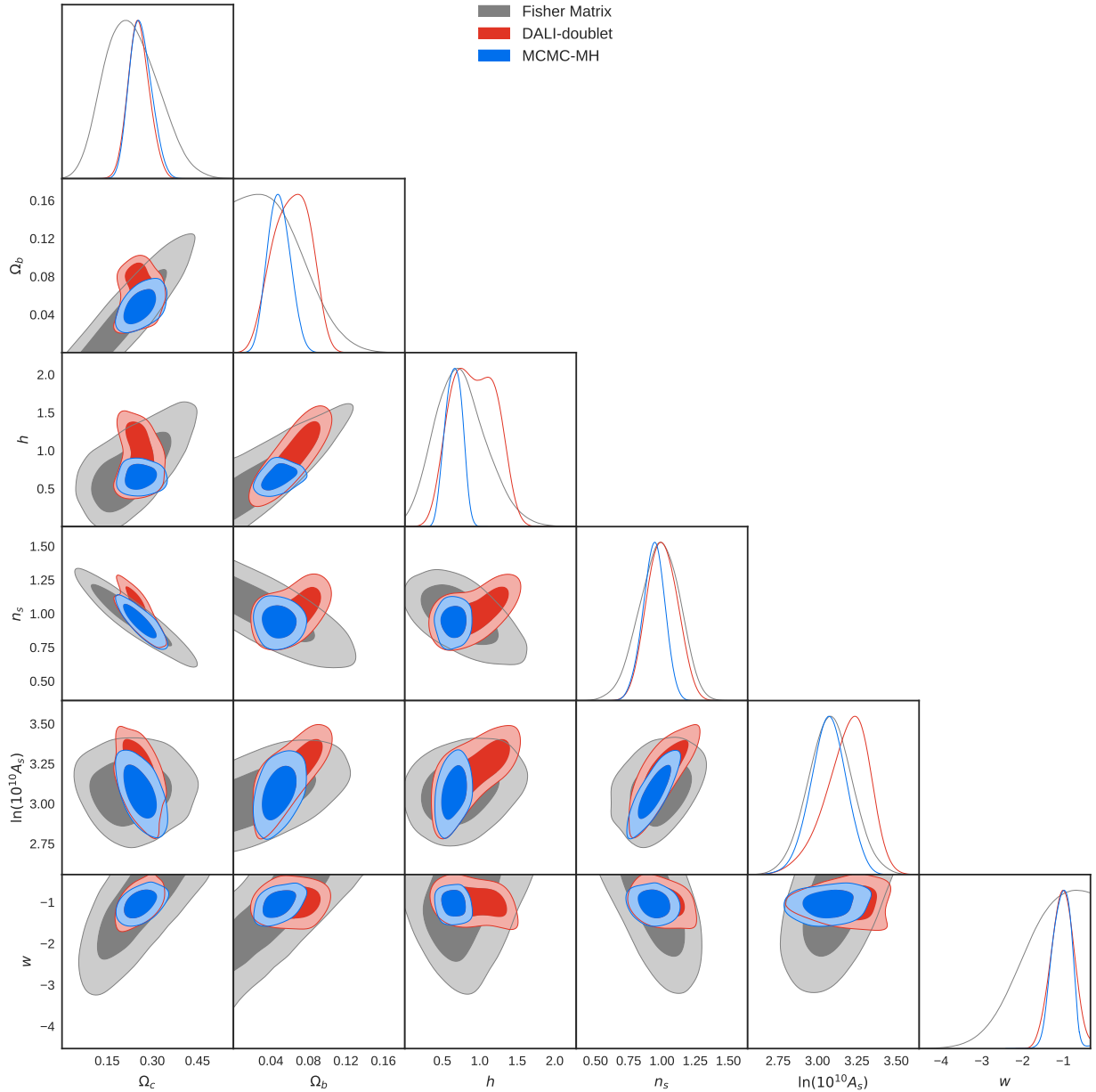
Figure 5.2: Confidence regions of the first redshift bin at $z = 0.31$, now including the prior limits, which act by filtering every sampled point outside of them.

Fig. 5.4 shows the confidence regions of the last bin. We observe that DALI does not capture the shape or size of the distribution as well as before, while the FM gets closer to determining the size of the distribution while staying larger than it.

For the last two redshift bins, the constraining power is much stronger and the MCMC sampler does not explore unphysical nor uncomputable regions. Despite that, the FM regions are still larger (although slightly) than the MCMC samples. The reason might be unsatisfactory as, up to this stage, we are limited by the precision of the computation. This part of the analysis became very sensitive to the numerical step for the derivatives with respect to the parameters, as we could obtain smaller FM regions while sacrificing any orientation detection
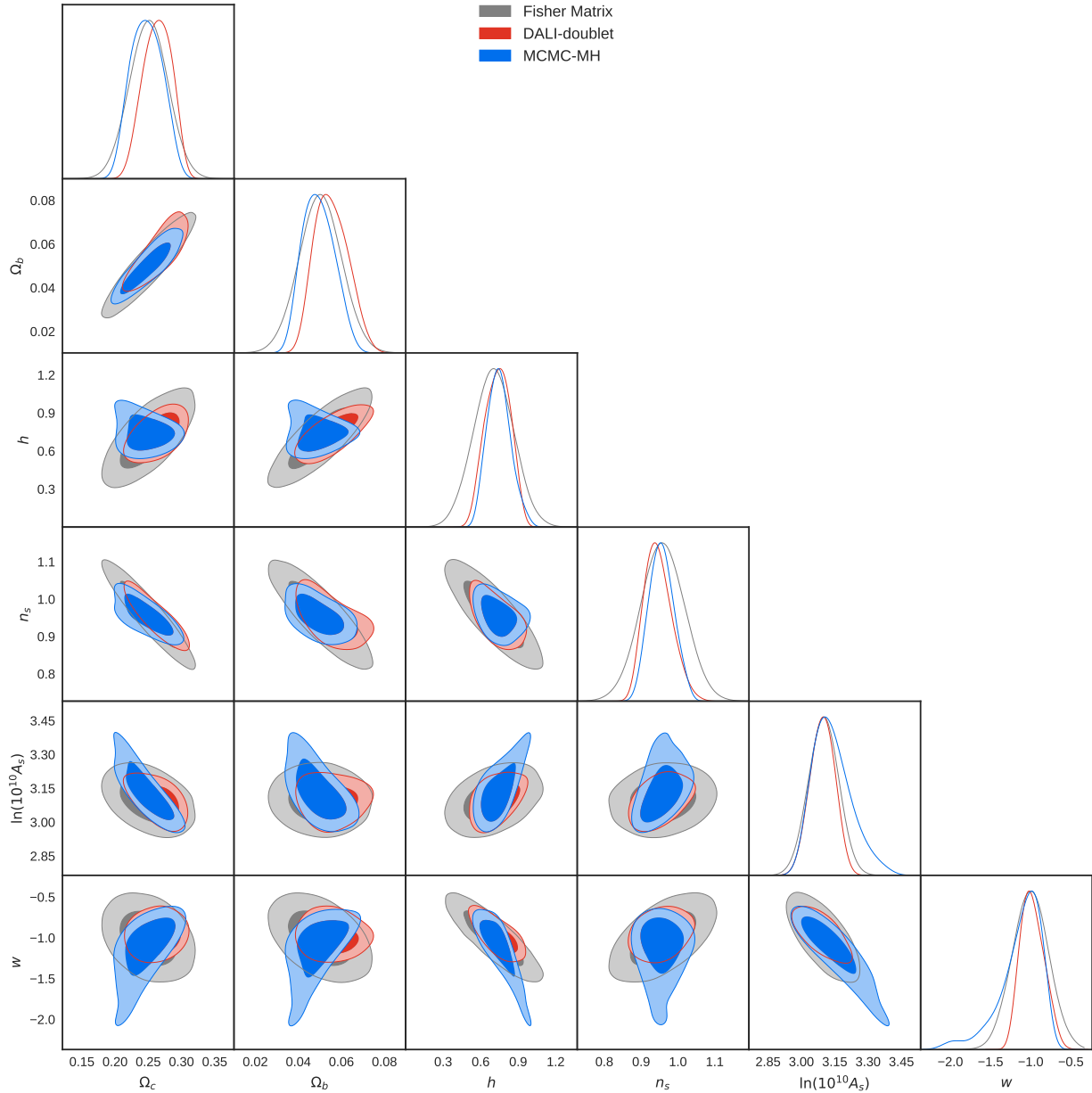
Figure 5.3: Confidence regions of the fourth redshift bin at $z = 1.18$.

by increasing the derivative's step. We chose orientation against size for the Fisher matrices, as this allowed us to obtain better constraints with DALI, which is even more susceptible to the step of the second-order derivatives, demanding them to be very high in order to retain its size. When attempting to use a smaller step, the derivatives become unstable, resulting in regions several times smaller than the ones reported in Fig. 5.4.

One possibility of an explanation for this is that the maximum wavenumber k_{max} on the latter bins is larger, and hence includes more of the BAO wiggles and non-linear scales on the power spectrum, which may be more difficult to take derivatives from, especially if one considers large numerical steps for it. In addition, different scales might react differently to the parameters. Hence, a good size for a step on small scales can be very off for larger scales,
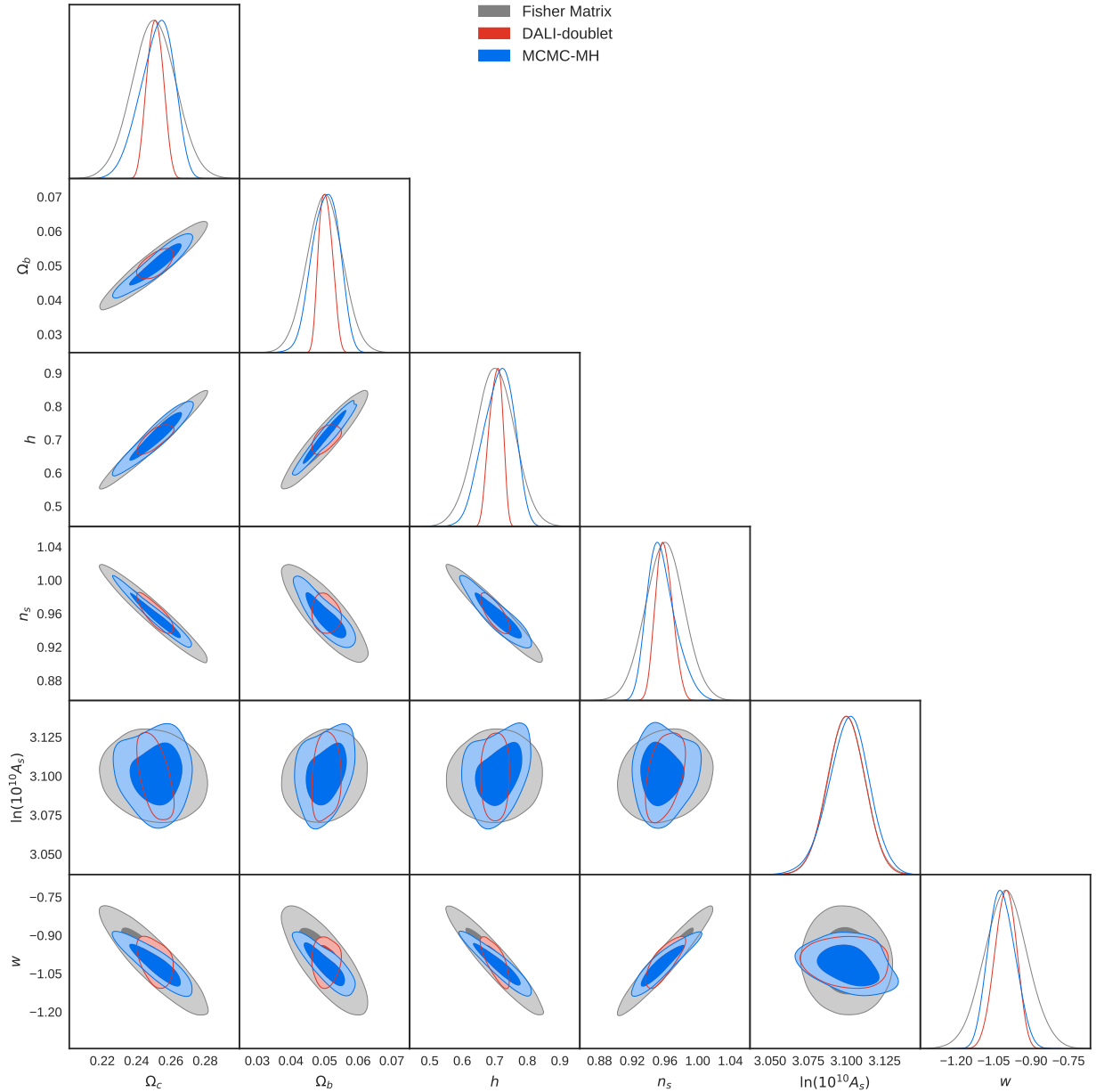
Figure 5.4: Confidence regions of the seventh redshift bin at $z = 2.67$.

and higher redshift bins have a larger difference between the minimum and maximum k.

In Fig. 5.5 we exhibit a summary of the analysis. It utilizes the FoM defined in 5.4.3. We observe an omnipresent oversizing of the FM regions, while the DALI regions are much more close in size to the MCMC samples, evidencing its better performance at forecasting. We can see that the difference between the FM-FoM and the MCMC-FoM is fairly even-distributed over the different parameter combinations. On the other hand, we see a clear diminishing of it for higher redshift bins, while they are more constraining. DALI follows the same trends just mentioned. It is important to mention that, for the fifth redshift bin, the FM obtains a closer constraining level to the samples than DALI.
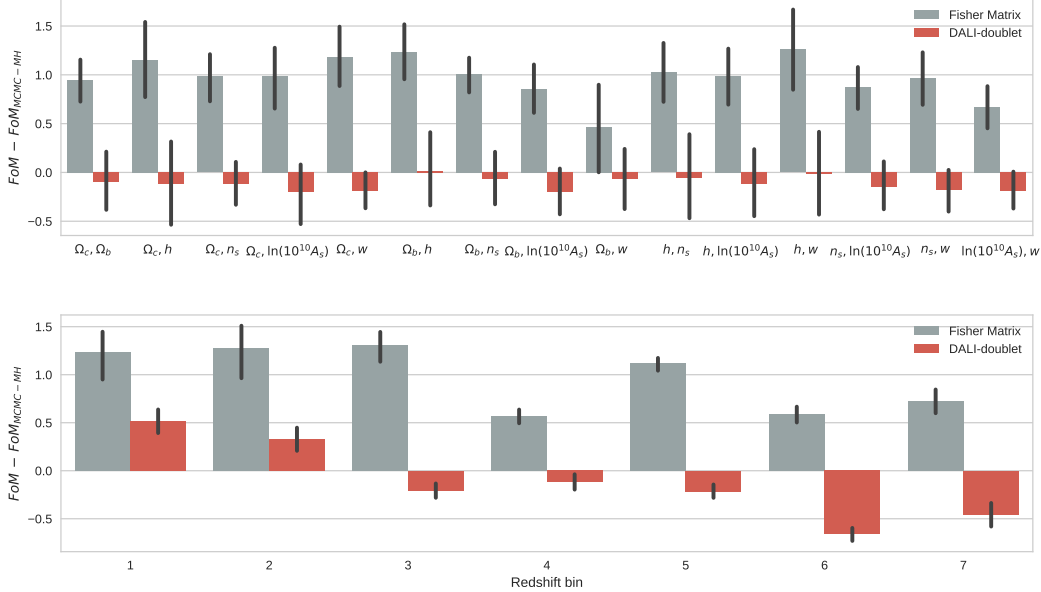
Figure 5.5: Summary of the FoM results. In order to better compare to the MCMC FoM, subtracted it. Hence, a larger FoM is bound to a larger area of the confidence region and less constraining power. In the upper panel, we have a mean over redshift bins, while on the lower panel we have a mean over the parameter combinations. The black lines delimit the standard deviation of the averaging over redshift bin or parameter combination, respectively.

Lastly, Fig 5.6 shows the total constraints for the survey. The confidence regions are similar to those of the last redshift bin, as it is by far the most constraining, which made the FM method closer in size than DALI to the MCMC samples. To obtain the total Fisher matrix and DALI-tensors, we summed them up without considering the prior filtering that is applied to each separate bin. Notice that the Ω_b, h region has a small 'island' region, which was not expected. The MCMC chain used for the total survey was several times longer than the ones per-bin, and so we do not attribute this extra mode to an issue of poor sampling.

5.6 Summary and conclusion

In this work, we performed galaxy clustering forecasts on the constraining power of LSST, a photometric galaxy clustering survey currently in construction. For this purpose, we utilized the Fisher Matrix (FM) method along with the DErivative Approximation of LIkelihoods (DALI) [49, 54], which is a method that extends the FM by including higher-order derivatives on the log-likelihood of the model. We contrasted both approximations with Metropolis-Hastings MCMC samples.

DALI is a method conceived to go beyond the FM by capturing a variety of shapes in the confidence regions of the parameters, shapes other than the ellipses product of a Fisher forecast, such as the usual *boomerangs* or *bananas*. Therefore, our goal was originally to inspect if there were shapes to capture using the DALI method that would result in
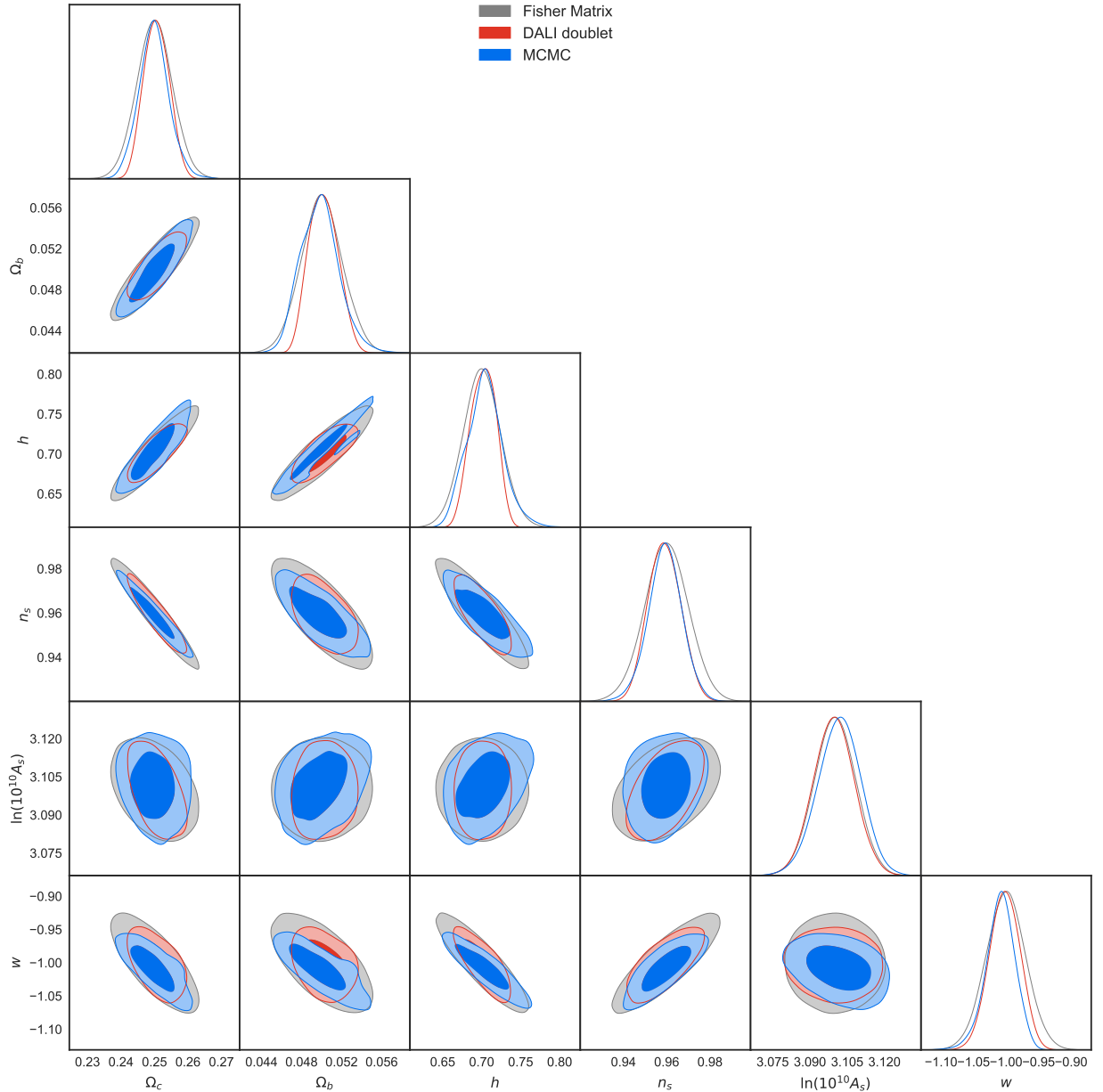
Figure 5.6: Confidence regions for the overall survey.

an improvement to the FM. However, during the production of the forecasts, we noticed limitations in terms of the numerical precision required in order to obtain accurate shape detection with DALI. The greater order derivatives on the likelihood of DALI translates into derivatives of the power spectrum, which get numerically difficult when considering the Alcock-Paczynski effect, as this changes the 'grid' of the integrals that need to be computed for the likelihood, forcing us to compute high-density interpolations on the power spectrum. Consequently, we were forced to augment the numerical step taken for the second-order derivatives of parameters, restricting us to make more accurate constraints only in terms of the size of confidence regions, compared to the FM regions.

For low and medium redshifts, this is for $z \leq 1.86$, we observed that DALI is closer in size

and shape to the MCMC samples, especially on the lower redshifts ($z \leq 0.64$), where the FM approximation largely overestimates the size of the regions, considering large portions of unphysical parameters. We applied a flat prior on top of this likelihood in order to improve their constraints. As the redshift increases, the regions for both the FM and DALI decrease in size. For high redshift bins, where $z \geq 1.86$, DALI gets considerably smaller than the MCMC constraints, being even with the FM in terms of our Figure of Merit, which is a logarithmic measure of confidence region area. This means that the DALI method underestimates the regions as much as the FM overestimates them.

Our conclusion is that the DALI-extension to the FM is a good tool that enhances the accuracy of forecasts. Apart from its shape capturing power, we found that it can estimate sizes of confidence regions better. Nonetheless, numerical difficulties fogged these insights on higher redshifts, where the appropriate step for second-order numerical derivatives on the parameter space is smaller than the step allowed by our accuracy.

Appendix

5.7   CLASS precision file

The crucial CLASS precision parameters used for this work were

```
k_per_decade_for_pk = 50000
k_per_decade_for_bao = 50000
back_integration_stepsize = 1.5e-4
```

These parameters have the goal of giving a very dense output for both the power spectrum and the background quantities. This ended in a very high RAM-usage and therefore should be run in a computer cluster rather than a desktop one.

In addition to the crucial parameters, we used the file `pk_ref.pre` which comes included with CLASS. It comprises a set of high fidelity power spectrum precision parameters.

Chapter 6

The Internal Robustness of Growth Rate data

This chapter comprehends a work carried in an internship at the IFT-UAM under the guidance of Savvas Nesseris during the months of April and May of 2018, culminating in the article of Ref. [60], submitted to Physical Review D[1].

Abstract

We perform an *Internal Robustness* analysis (iR) to a compilation of the most recent $f\sigma_8(z)$ data, using the framework of Ref. [61]. The method analyzes combinations of subsets in the dataset in a Bayesian model comparison way, potentially finding outliers, subsets of data affected by systematics or new physics. In order to validate our analysis and assess its sensitivity we performed several cross-checks, for example by removing some of the data or by adding artificially contaminated points, while we also generated mock datasets in order to estimate confidence regions of the iR. Applying this methodology, we found no anomalous behavior in the $f\sigma_8(z)$ dataset, thus validating its internal robustness.

6.1   Introduction

During the last twenty years a plethora of observations suggests that the Universe is undergoing a phase of accelerated expansion at late times. In order to explain this phenomenon, the concept of attractive gravity had to be revised either by introducing a new form of matter dubbed dark energy, see Ref. [19] for a review, or by altering explicitly the laws of gravity [62]. However, the simplest way to account for a phase of accelerated expansion of the Universe within the framework of Friedmann-Lemaître-Robertson-Walker (FLRW) cosmology is to simply introduce a cosmological constant (Λ). While this model gives rise to severe

---

[1]`https://journals.aps.org/prd/`

coincidence and fine-tuning problems, current cosmological observations are still compatible with a Universe that is filled by a dark energy component that has the same characteristics of the cosmological constant [63].

Nonetheless, these cosmological observations are not accurate enough at the moment to either constrain any potential time evolution of the cosmological constant, which might lead dark energy to cluster, evolve in time, or any modifications of gravity. Despite the fact that the two aforementioned classes of theories can in principle be arbitrarily similar [64, 65], it is still necessary to be able to discriminate between the currently available models.

Future surveys such as Euclid [18] and LSST [66], both of which will gather orders of magnitude more data than current surveys. Thus, it would be interesting to constrain the dynamical features of gravity and test our assumptions. One way to do this is via the growth of matter density perturbations $\delta_m = \delta\rho_m/\rho_m$ and its logarithmic derivative the growth rate $f = d\ln\delta_m/d\ln a$. In practice, most of the growth rate measurements are made via the peculiar velocities obtained from Redshift Space Distortions (RSD) measurements [59] coming from galaxy redshift surveys. In general, these surveys measure the perturbations of the galaxy density δ_g, which can be related to (dark) matter perturbations through the bias b via $\delta_g = b\,\delta_m$. Thus, initial growth rate measurements measured the growth rate f divided by the bias factor b leading to the parameter $\beta \equiv f/b$. This parameter is very sensitive to the value of the bias which can vary in the range $b \in [1, 3]$ (conservatively speaking), thus making difficult to combine β from different surveys and as a result leading to unreliable datasets of $\beta(z_i)$.

As a result, a more reliable parameter was sought and this was found in the combination $f(z)\sigma_8(z) \equiv f\sigma_8(z)$, which can be shown to be independent of the bias, and can be measured either via weak lensing or RSD observations. Still, the current measurements of $f\sigma_8(z)$ (presented in later sections) come from a plethora of different surveys with different assumptions and systematics, thus an approach to study the statistical properties and robustness of the data is imperative.

One such approach is the so-called "Internal Robustness", pioneered in Ref. [61]. This is a fully Bayesian approach which is not only sensitive to the local minimum like a standard χ^2 comparison, but also to the entire likelihood and can in principle detect the presence of systematics in the dataset. The main goal of this approach is to identify systematic-contaminated data-points, which can then be further analysed and potentially excluded if they cannot be corrected.

In this work we present an application of the "Internal Robustness" approach to the currently available growth-rate data in the form of $f\sigma_8(z)$ with the aim to examine the dataset for systematics and outlier points in a fully automated manner. The layout of our manuscript is as follows: in Sec. 7.2 we provide the basic elements of FLRW cosmology related to our models, in Sec. 6.3 we briefly review the "Internal Robustness" method and its application to the growth data, while in Sec. 7.4 we provide the compilation of growth data used in our analysis and finally, we discuss our results in Sec. 7.6.

6.2 Basic equations

In this section we present the basic equations required in our analysis. We begin with the Hubble parameter in a flat ΛCDM universe (with a constant equation of state parameter for dark energy $w = -1$), given by

$$H(a)^2 = H_0^2 \left[\Omega_{m,0} a^{-3} + (1 - \Omega_{m,0}) \right] \ , \tag{6.1}$$

where H_0 is the Hubble constant, and $\Omega_{m,0}$ is the present day value of the matter density parameter and a is the scale factor. The matter density can then also be expressed as a function of the scale factor:

$$\Omega_m(a) = \frac{\Omega_{m,0} a^{-3}}{H(a)^2 / H_0^2} \ . \tag{6.2}$$

Under the assumption of flat ΛCDM model, the angular diameter distance takes an analytical expression, given by:

$$\frac{H_0}{c} \frac{\sqrt{\Omega_{m,0}}}{2a} d_A(a) = {}_2F_1 \left[\frac{1}{2}, \frac{1}{6} ; 1 + \frac{1}{6} ; \left(1 - \frac{1}{\Omega_{m,0}} \right) \right]$$
$$- \sqrt{a} \cdot {}_2F_1 \left[\frac{1}{2}, \frac{1}{6} ; 1 + \frac{1}{6} ; \left(1 - \frac{1}{\Omega_m(a)} \right) \right] \ , \tag{6.3}$$

where ${}_2F_1$ is the hypergeometric function. The matter density perturbations in Fourier space $\delta_m(a, k)$ depend on the underlying cosmological model; for the ΛCDM scenario, the linear matter perturbations grow according to

$$\delta_m''(a) + \left(\frac{3}{a} + \frac{H'(a)}{H(a)} \right) \delta_m'(a) - \frac{3}{2} \frac{\Omega_m(a)}{a^2} \delta_m(a) = 0 \ . \tag{6.4}$$

The equation above has an analytical solution for the growing mode, given by [67, 68, 69]

$$\delta_m(a) = a \cdot {}_2F_1 \left(\frac{1}{3}, 1 ; \frac{11}{6} ; a^3 (1 - \frac{1}{\Omega_{m,0}}) \right) \ . \tag{6.5}$$

Note that the dependence on the wave number k disappears because of the assumption of small scales approximation.

We define the growth rate f and the root mean square (RMS) normalization of the matter power spectrum σ_8 as:

$$f(a) = \frac{d \log \delta_m}{d \log a}, \tag{6.6}$$

$$\sigma_8(a) = \sigma_{8,0} \frac{\delta_m(a)}{\delta_m(1)} \ . \tag{6.7}$$

As already mentioned in Sec. 6.1, a more robust and reliable quantity that is measured by redshift surveys is the combination of the growth rate $f(a)$ and the RMS $\sigma_8(a)$:

$$f\sigma_8(a) = a \frac{\delta_m'(a)}{\delta_m(1)} \sigma_{8,0} \ . \tag{6.8}$$

Equation (6.8) will be our key quantity, which will be tested with the most recent data available in the following sections.

6.3 Formalism

Here we report the basic equations that will be used to perform our analysis and we refer to [61] for the details on the derivation of the internal robustness. We can compare two different models by considering the ratio of their probabilities:

$$\frac{\mathcal{P}(M_1|\mathbf{x})}{\mathcal{P}(M_2|\mathbf{x})} = \mathcal{B}_{12}\frac{\pi(M_1)}{\pi(M_2)} \, , \tag{6.9}$$

with $\mathcal{P}$ and π symbolizing posterior and prior probabilities, respectively. The Bayes factor is defined as

$$\mathcal{B}_{12} = \frac{\mathcal{E}(\mathbf{x}|M_1)}{\mathcal{E}(\mathbf{x}|M_2)} \, . \tag{6.10}$$

If we assume that the prior probabilities of having two different models are the same, then the Bayes factor alone will help us to favor or disfavor a particular model. If $\mathcal{B}_{12} > 1$ then the data favors the model M_1, if it is less than 1, then M_2 is favored.

However, the robustness test needs a further assumption, that is: the data have to come from two different distributions. The reason is two fold: first the total evidence can be factorised as the product of the two evidences and, second, which is the underlying meaning of the robustness test, we would like to prove that data are reliable. If the data is partitioned into two subsets, say $\{\mathbf{x_1}, \mathbf{x_2}\}$ and we assume that they come from two models, say M_1, M_2, then the Bayes factor becomes

$$\mathcal{B}_{12} = \frac{\mathcal{E}(\mathbf{x}|M_1)}{\mathcal{E}(\mathbf{x_1}|M_1)\mathcal{E}(\mathbf{x_2}|M_2)} \, . \tag{6.11}$$

Finally, we can define the internal robustness as

$$\mathrm{iR}_{12} = \log\mathcal{B}_{12} = \log\left(\frac{\mathcal{E}(\mathbf{x}|M_1)}{\mathcal{E}(\mathbf{x_1}|M_1)\mathcal{E}(\mathbf{x_2}|M_2)}\right) \, . \tag{6.12}$$

This approach allows us to detect if a subset of the data follows another cosmological model or if a specific survey is affected by systematics and hence altering the measurement itself.

However, the assumption of having two different models is not strictly mandatory and we will choose the same cosmological model for both subsets. In this work we will set the cosmological model to be ΛCDM and the parameters for both subsets to $\boldsymbol{\theta} = (\Omega_{m,0}, \sigma_{8,0})$. Hereinafter, we drop the M superscript and the $1, 2$ subscripts, since we only consider one cosmological model.

Our analysis invokes priors on the parameters, for that we choose a flat prior in the $[0, 1]$ range for $\Omega_{m,0}$ in order to allow for all physical values possible for the matter density. On the other hand, the choice of a prior for $\sigma_{8,0}$ is less evident; since the prior directly affects the evidence value, so we are going to consider three priors for $\sigma_{8,0}$, to assess the impact of the prior selection on the internal robustness. We choose the following three cases:

- **Narrow flat prior**: this is a typical flat prior in the range $[0.3, 1.5]$.
- **Broad flat prior**: this is a flat prior in the range $[0, 10]$, which allows for high values of $\sigma_{8,0}$.
- **Gaussian prior**: the third prior to consider is a Gaussian distribution centered on 0.8150, with a standard deviation of 0.0087, based on the Planck 2015 results (TT, TE, EE+lowP+lensing [63]).

It is clear that we only allow for positive values of $\sigma_{8,0}$ in order to remain physical.

The data considered are $f\sigma_8(z)$ measurements (with $z = -1+1/a$ being the redshift of the measurement), and we assume a Gaussian likelihood for the data with a covariance matrix $\mathbf{C}$. We represent the observed data in different redshifts as $\mathbf{m} = (m(z_1), \ldots, m(z_n))$ and its theoretical prediction as $\boldsymbol{\mu}(\boldsymbol{\theta}) = (\mu(z_1), \ldots, \mu(z_n))$, which depends on the cosmological model and parameters. We also take into account the redshift correction of [70], which features a correction factor of

$$\text{fac}(z^i) = \frac{H(z^i)\, d_A(z^i)}{H^{\text{ref},i}(z^i)\, d_A^{\text{ref},i}(z^i)} \, . \tag{6.13}$$

with the label ref, i stating that the cosmology considered is the reference cosmology used on the corresponding data point on redshift z^i. Hence, the corrected theoretical prediction is

$$\mu_c^i = \frac{\mu^i}{\text{fac}(z^i)} \, . \tag{6.14}$$

We are now in the position to define the data vector with the corresponding modification:

$$\mathbf{x} = \mathbf{m} - \boldsymbol{\mu}_c \, . \tag{6.15}$$

Then, the chi-squared is

$$\chi^2 = \mathbf{x}^T \mathbf{C}^{-1} \mathbf{x} \, , \tag{6.16}$$

which is related to the likelihood via $L = e^{-\chi^2/2}/\sqrt{(2\pi)^n |\mathbf{C}|}$.

To speed up the computations, we note that the $\sigma_{8,0}$ parameter can be marginalised theoretically [71, 72]. We rewrite the χ^2:

$$\chi^2 = \mathbf{m}^T \mathbf{C}^{-1} \mathbf{m} - 2\mathbf{m}^T \mathbf{C}^{-1} \boldsymbol{\mu}_c + \boldsymbol{\mu}_c^T \mathbf{C}^{-1} \boldsymbol{\mu}_c \, . \tag{6.17}$$

The corrected theoretical prediction marginalised over $\sigma_{8,0}$ will be $\boldsymbol{\nu} = \boldsymbol{\mu}_c/\sigma_{8,0}$. Then, the χ^2 can be rewritten as

$$\chi^2 = \xi_{mm} - 2\xi_{m\nu}\sigma_{8,0} + \xi_{\nu\nu}\sigma_{8,0}^2 \, , \tag{6.18}$$

where the single terms are:

$$\begin{aligned}
\xi_{mm} &= \mathbf{m}^T \mathbf{C}^{-1} \mathbf{m} \, , \\
\xi_{m\nu} &= \mathbf{m}^T \mathbf{C}^{-1} \boldsymbol{\nu} \, , \\
\xi_{\nu\nu} &= \boldsymbol{\nu}^T \mathbf{C}^{-1} \boldsymbol{\nu} \, .
\end{aligned} \tag{6.19}$$

The posterior probability distribution marginalised over $\sigma_{8,0}$ is

$$\mathcal{P}(\Omega_{m,0}) = \int L(\Omega_{m,0}, \sigma_{8,0})\pi(\Omega_{m,0}, \sigma_{8,0}) \mathrm{d}\sigma_{8,0} \, . \tag{6.20}$$

We now consider two cases for the prior probability on $\sigma_{8,0}$: a flat prior between $[a, b]$ and a Gaussian prior with mean s and variance ε^2. Let us start by considering the flat prior case. The integration of the posterior is:

$$
\begin{aligned}
\mathcal{P}_f(\Omega_{m,0}) &= \frac{1}{(b-a)\sqrt{(2\pi)^n|\mathbf{C}|}} \int_a^b e^{-\chi^2/2} d\sigma_{8,0} \\
&= \frac{1}{(b-a)\sqrt{(2\pi)^n|\mathbf{C}|}} \exp\left(-\frac{1}{2}\left[\xi_{mm} - \frac{\xi_{m\nu}^2}{\xi_{\nu\nu}}\right]\right) I_f \ ,
\end{aligned}
\tag{6.21}
$$

where the quantity I_f is:

$$
I_f = \int_a^b e^{-\frac{\xi_{\nu\nu}}{2}(\xi_{m\nu}/\xi_{\nu\nu} - \sigma_8)^2} d\sigma_8 = \sqrt{\frac{\pi}{2\xi_{\nu\nu}}} \mathrm{erf}\left(\frac{\xi_{m\nu} - \mathrm{x}\xi_{\nu\nu}}{\sqrt{2\xi_{\nu\nu}}}\right) \Bigg|_a^b .
\tag{6.22}
$$

For the Gaussian prior case we find, by discarding negative values:

$$
\begin{aligned}
\pi_g(\sigma_{8,0}) &= \frac{e^{-\frac{1}{2}(s-\sigma_{8,0})^2/\varepsilon^2}}{\int_0^\infty e^{-\frac{1}{2}(s-\sigma_{8,0})^2/\varepsilon^2} d\sigma_{8,0}} \\
&= \frac{e^{-\frac{1}{2}(s-\sigma_{8,0})^2/\varepsilon^2}}{\sqrt{\pi\varepsilon^2/2}\,[1 + \mathrm{erf}(s/\sqrt{2\varepsilon^2})]} = A_g\, e^{-\frac{1}{2}(s-\sigma_{8,0})^2/\varepsilon^2} \ ,
\end{aligned}
\tag{6.23}
$$

where we implicitly defined the normalization constant A_g.

The posterior probability distribution function then reads:

$$
\begin{aligned}
\mathcal{P}_g(\Omega_{m,0}) &= \frac{A_g}{\sqrt{(2\pi)^n|\mathbf{C}|}} \int_0^\infty e^{-[\chi^2 + (s-\sigma_{8,0})^2/\varepsilon^2]/2} d\sigma_{8,0} \\
&= \frac{A_g}{\sqrt{(2\pi)^n|\mathbf{C}|}} \exp\left(-\frac{1}{2}\left[\xi_{mm} - \frac{\xi_{m\nu}^2}{\xi_{\nu\nu}}\right]\right) I_g \ ,
\end{aligned}
\tag{6.24}
$$

where I_g is equal to:

$$
\begin{aligned}
I_g &= \int_0^\infty e^{-\frac{\xi_{\nu\nu}}{2}(\xi_{m\nu}/\xi_{\nu\nu} - \sigma_8)^2} e^{-\frac{1}{2}(s-\sigma_{8,0})^2/\varepsilon^2} d\sigma_8 \\
&= \sqrt{\frac{\pi\varepsilon^2}{2(\varepsilon^2\xi_{\nu\nu} + 1)}} \exp\left(-\frac{(\xi_{m\nu} - \xi_{\nu\nu}s)^2}{2\xi_{\nu\nu}(\varepsilon^2\xi_{\nu\nu} + 1)}\right) \left[1 + \mathrm{erf}\left(\frac{\xi_{m\nu}\varepsilon^2 + s}{\varepsilon\sqrt{2(\varepsilon^2\xi_{\nu\nu} + 1)}}\right)\right] \ ,
\end{aligned}
\tag{6.25}
$$

which is the multiplication of the exponentials of two Gaussians, which is also the exponential part of a Gaussian distribution.

6.4 Data Considerations

6.4.1 The Dataset

The growth rate dataset is based on the Gold-2017 compilation from [70], consisting of 18 independent measurements of $f\sigma_8(z)$, obtained from redshift space distortion measurements from a variety of surveys. Among these surveys, it is important to note that the three WiggleZ [73] measurements are correlated, and their covariance matrix is

$$\mathbf{C}_{\text{WiggleZ}} = 10^{-3} \begin{pmatrix} 6.400 & 2.570 & 0.000 \\ 2.570 & 3.969 & 2.540 \\ 0.000 & 2.540 & 5.184 \end{pmatrix}. \tag{6.26}$$

In addition to the Gold-2017 compilation, we update it with 4 recent measurements from [74]. These points have a covariance matrix given by

$$\mathbf{C}_{\text{SDSS-IV}} = 10^{-2} \begin{pmatrix} 3.098 & 0.892 & 0.329 & -0.021 \\ 0.892 & 0.980 & 0.436 & 0.076 \\ 0.329 & 0.436 & 0.490 & 0.350 \\ -0.021 & 0.076 & 0.350 & 1.124 \end{pmatrix}. \tag{6.27}$$

Our final dataset will be constituted of $N = 22$ data points, shown in Table 6.1, the possible combinations of subsets from the them is $2^{N-1} - 1 = 2097151$, and we analyze all of the subsets[2]. The analysis is possible thanks primarily to the marginalization over $\sigma_{8,0}$, as shown in Sec. 6.3.

6.4.2 Mock Data

An important feature of this work is the comparison of confidence regions for the probability distributions of the internal robustness (iR-PDF). To obtain the confidence regions, we generate mock datasets based on the form

$$f\sigma_8^{\text{mock}}(z_i) = f\sigma_8(z_i | \, \boldsymbol{\theta}^{\text{bestfit}}) + \mathcal{N}^{\text{random}} \,, \tag{6.28}$$

meaning that the mock growth rate data is generated from the best fit parameters $\boldsymbol{\theta}^{\text{bestfit}}$, which are obtained using the complete dataset and minimizing the posterior probability (which takes the prior into account). The $\mathcal{N}^{\text{random}}$ term is evaluated by assuming a Gaussian noise with zero mean and standard deviation equal to those given by the data $\sigma_{f\sigma_8}(z_i)$.

The main reason of comparing the results obtained by using the data and the mock catalogues is to compare directly the iR-PDF to the confidence regions. If the iR-PDF from the data falls off the confidence regions, then we can state that either the data set is not internally robust, meaning that the dataset could be affected by systematics, or that a better

---

[2]Note that we do not count the combination of the full dataset with the empty set $\varnothing$.

Table 6.1: Compilation of the $f\sigma_8(z)$ measurements used in this analysis along with the reference matter density parameter Ω_{m_0} (needed for the growth correction) and related references.

z	$f\sigma_8(z)$	$\sigma_{f\sigma_8}(z)$	$\Omega_{m,0}^{\text{ref}}$	Ref.
0.02	0.428	0.0465	0.3	[75]
0.02	0.398	0.065	0.3	[76],[77]
0.02	0.314	0.048	0.266	[78],[77]
0.10	0.370	0.130	0.3	[79]
0.15	0.490	0.145	0.31	[80]
0.17	0.510	0.060	0.3	[81]
0.18	0.360	0.090	0.27	[82]
0.38	0.440	0.060	0.27	[82]
0.25	0.3512	0.0583	0.25	[83]
0.37	0.4602	0.0378	0.25	[83]
0.32	0.384	0.095	0.274	[84]
0.59	0.488	0.060	0.307115	[85]
0.44	0.413	0.080	0.27	[73]
0.60	0.390	0.063	0.27	[73]
0.73	0.437	0.072	0.27	[73]
0.60	0.550	0.120	0.3	[86]
0.86	0.400	0.110	0.3	[86]
1.40	0.482	0.116	0.27	[87]
0.978	0.379	0.176	0.31	[74]
1.23	0.385	0.099	0.31	[74]
1.526	0.342	0.070	0.31	[74]
1.944	0.364	0.106	0.31	[74]

physical model is required in order to better describe the data. In other words, the mock data confidence regions portrait acceptable offset levels from the best fit cosmology obtained from the complete dataset.

For each choice of the prior on $\sigma_{8,0}$, we generate 1000 mock datasets. Then, we sample each one of these datasets into 14000 subset combinations, distributed as follows: 2000 samples for subsets in which the smaller subset size (hereinafter SSS) is 11, another 2000 for subsets with SSS=10, and so on, until SSS=4. We stop at SSS=4 because the number of samples would be larger than the available combinations. As mentioned, the goal is to explore different subset sizes in an equal manner, with the further consideration that, for larger SSS value, we have more possible combinations.

As an ultimate test, we produced mock datasets based on the Planck 2015 best fit parameters [63], for which the parameters are $\Omega_{m,0} = 0.3121$ and $\sigma_{8,0} = 0.815$. The

idea behind this is to check whether the tension on measurements of $\sigma_{8,0}$ between Cosmic Microwave Background (CMB) surveys like Planck and galaxy clustering surveys, see Refs. [70, 88, 89, 90, 91, 92, 93, 94, 95], could be due to inconsistencies in the data themselves.

6.4.3 Cross-Checks

In order to ensure that the method is stable and sensitive to the dataset, we decided to opt for two extra cross-checks on our analysis. In brief, the cross-checks have been done using the narrow flat prior only, where we expect the method to be more sensitive to the final results. The cross-checks are:

- **Data removal**: we select one of the combinations with lowest internal robustness and SSS. Then, we remove the data points corresponding to the smaller subset, and evaluate again the complete internal robustness analysis with the new dataset. Clearly this procedure forces us to generate a new mock dataset with its own best fit. The SSS value will now range from SSS=4 to the maximum SSS possible. In order to be consistent with the number of points, each SSS will be constituted of 2000 sample subsets.
- **Data contamination**: we deliberately choose to *contaminate* the first data point of the dataset in order to have a worse iR-PDF. This contamination has been implemented by moving the data point by 5σ away from its actual value. In other words, the new first point is constructed as

$$f\sigma_8^{\text{cont}}(z_1) = f\sigma_8(z_1) + 5\,\sigma_{f\sigma_8}(z_1)\,. \tag{6.29}$$

By moving one of the point by 5σ away from its actual position, we expect the iR-PDF to be affected and fall off the confidence regions, clearly the mean iR has to decrease.

6.5 Results and Discussion

The first results are the complete inspection of the dataset, comprising all the possible subset combinations. In Fig. 6.1 we show the iR-PDF in the form of violin plots, arranged by the smaller subset size (SSS) of each subset combination. The three figures differ by the prior used.

From Fig. 6.1 we can see that the internal robustness increases with the SSS. This results was somehow expected as a larger data subsets are less prone to manifest outliers, if the data is free of systematic effects. We also see that the broad prior (middle panel in Fig. 6.1) has much larger iR than the narrow prior (upper panel in Fig. 6.1). The difference in the iR value is of the order two regardless of the SSS considered. However, the shape of the distributions changes for SSS<3. We can also see that for the Gaussian prior, the distributions are more stretched for small SSS (SSS<3) and more clumped up for medium and larger SSS.

In Fig. 6.2 we show the confidence regions of the mock datasets, as reported in Sec. 7.4, along with the iR-PDF of the corresponding prior. The dataset black lines were obtained
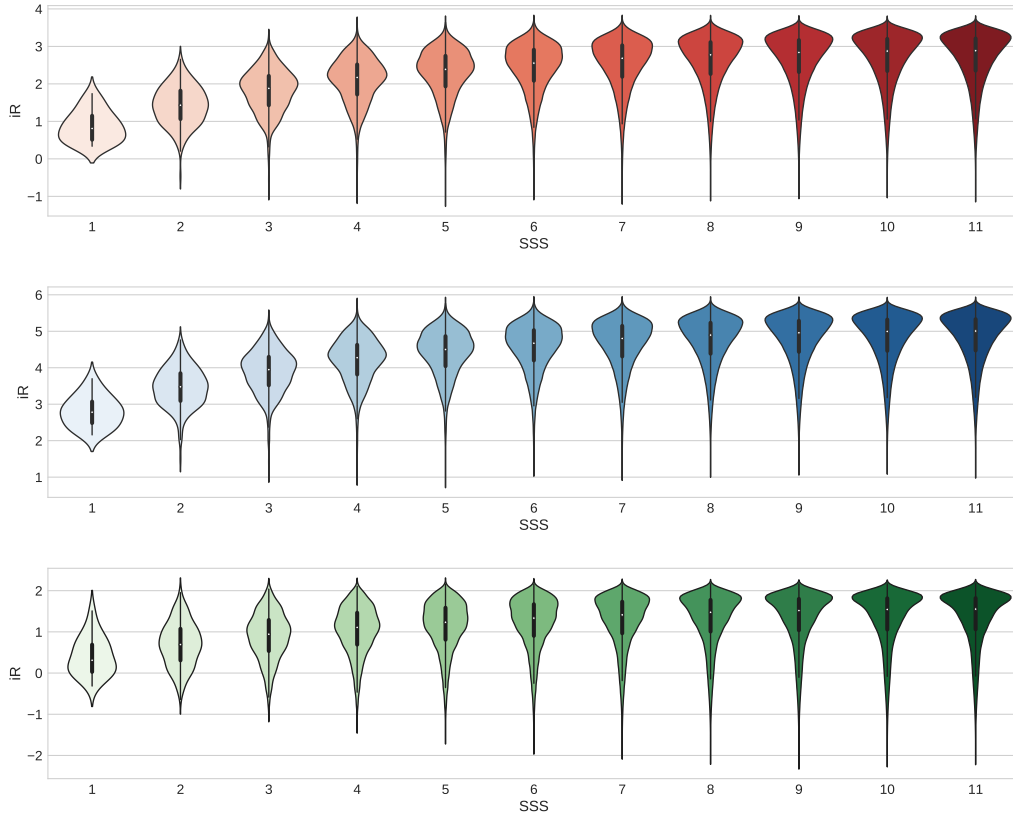
Figure 6.1: Violin plots of the internal robustness distributions, for each of the smaller subset sizes of each partition. We employed the narrow flat prior (upper panel), broad flat prior (middle panel) and Gaussian prior (lower panel). The white dots are the mean value of the internal robustness, the bold black line is the 1σ region and the thinner black line is the 2σ region.

from samples that were equal in size in each SSS as the mock data. We observe that, with the 3 types of prior, the data iR-PDF is within the confidence levels obtained in all the ranges of the internal robustness. This validates the dataset, discarding systematic contamination and any other irregularities detectable within the iR formalism.

In Fig. 6.3 we have plot the confidence regions with the Planck-based cosmology. We see that the confidence regions are nearly identical to those of Fig. 6.2, with the exception of the Gaussian prior case, where the iR-PDF gets closer to the 1σ region with the Planck-based cosmology mock data. We recall that the Gaussian prior was also chosen based on the Planck 2015 results, so this result is not controversial, although it was not automatically expected, unless we consider the prior to be more constraining than the likelihood alone.

6.5.1 Cross-checks results

As mentioned in the previous section, we decided to make a cross-check analysis to ensure that both method and dataset gave sensible results. The first check consisted of removing
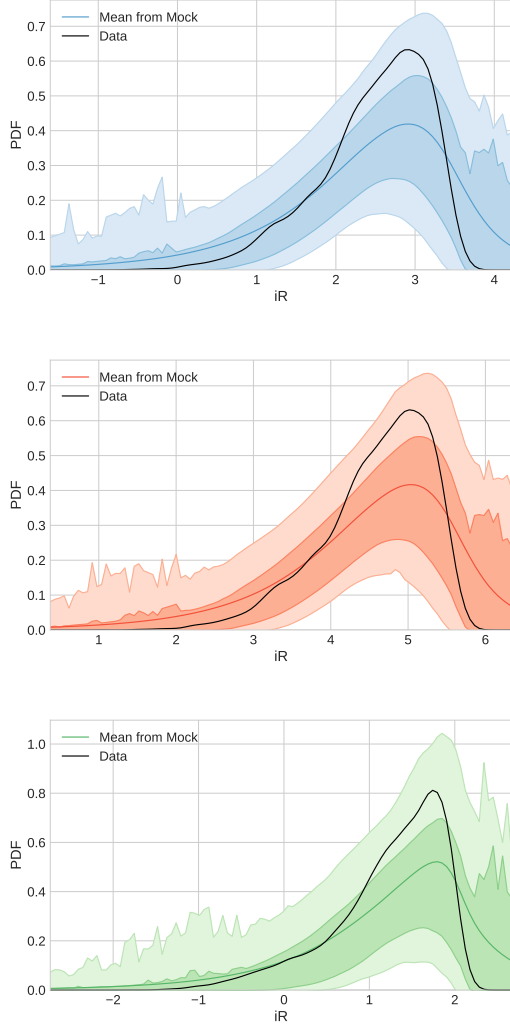
Figure 6.2: Internal Robustness PDF and confidence regions from mock data based on the best fit cosmology, using each prior. They are: narrow flat prior (upper panel), broad flat prior (medium panel) and Gaussian prior (lower panel).
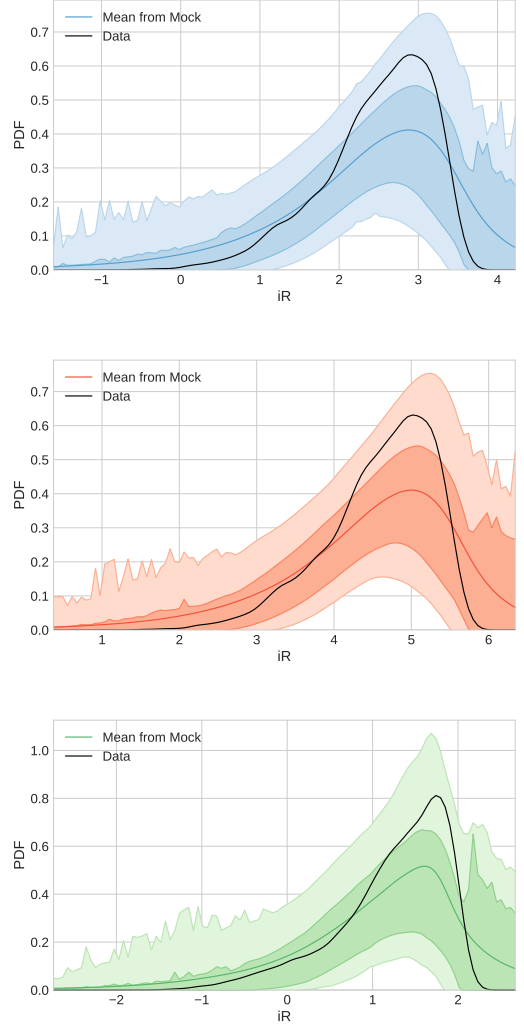
Figure 6.3: Same as Fig. 6.2, but the base parameters for the mock data are from Planck 2015. Priors used: narrow flat prior (upper panel), broad flat prior (medium panel) and Gaussian prior (lower panel).

data points from the subset that gave the lowest robustness. In our analysis we found that the lowest SSS that gave a negative lowest robustness was constituted of 2 points (hence SSS=2) and the data points falling into this subset were the second and sixth data in the table 6.1. We decided to name this subset 'r26'.

Our second cross-check was to take the first data point[3] and move 5 σs away from its actual position. The new dataset is denominated 'c1'. In Fig. 6.4 we show the iR-PDFs for the cross-check datasets. We can see from the figure that, for the 'r26' dataset, the iR for

---

[3]There is no particular reason of choosing the first point. Since the dataset is statistically robust, we are allowed to take randomly any point on the list.
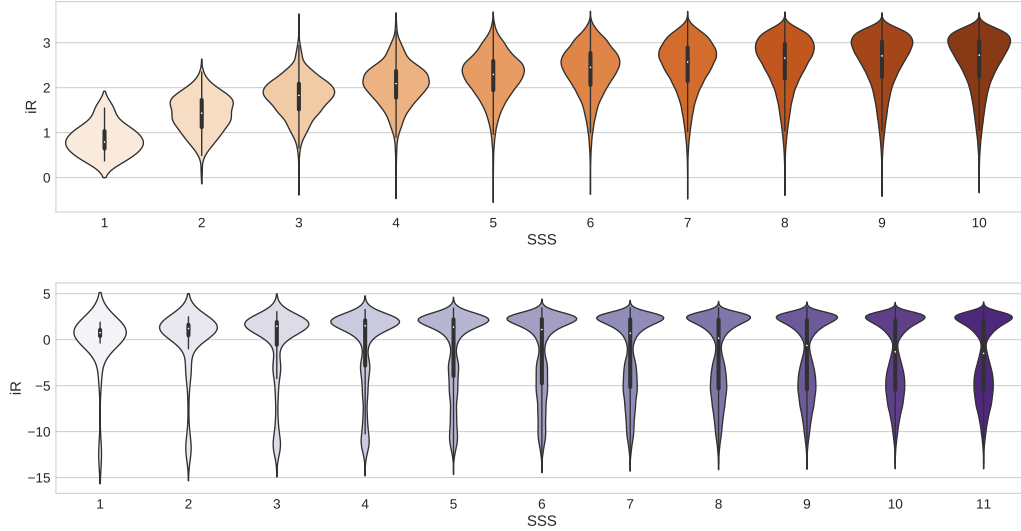
Figure 6.4: Same as Fig. 6.1, but considering the 'r26' dataset (upper panel) and the 'c1' dataset (lower panel), both with a narrow flat prior.
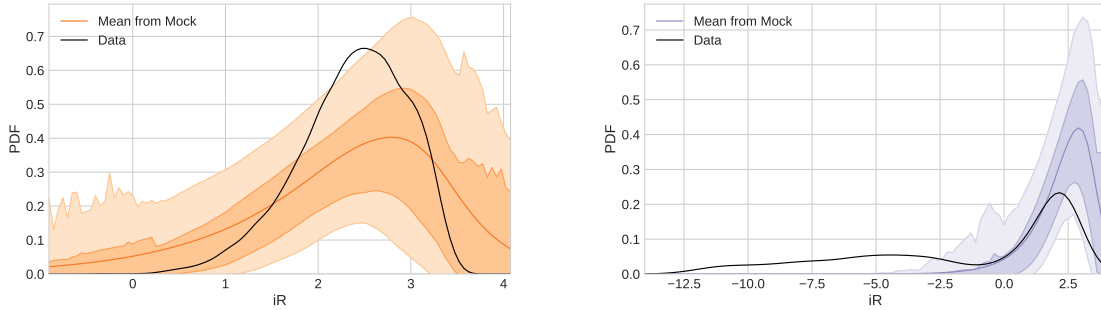
Figure 6.5: Same as Fig. 6.2, but considering the 'r26' dataset (left panel) and the 's1' dataset (right panel), both with a narrow flat prior. The mock data used to generate the confidence regions from the upper panel come from the best fit of the 'r26' dataset, while the ones for the lower panel are from the original dataset.

SSS>6 has a higher minimum but the maximum iR is lower, as well as the mean iR values are smaller with respect to the full dataset. This is probably due to, when improving a dataset by adding robust points, the iR is expected to increase. On the other hand, the improvement in the minimum iR indeed comes from choosing to remove the points with lower iR on the original dataset.

For the second cross-check, i.e. the contamination of one datapoint, we can see immediately that the internal robustness method detects the inconsistency caused by the contaminated datapoint, by exhibiting a bimodal shape and a decrease of the iR value when we increase the SSS. These are two features that are not proper of a robust dataset.

Finally, in Fig. 6.5 we show the confidence regions for the cross-check datasets. We can see that the confidence regions for the removal cross-check do not fully contain the iR-PDF.

74

The reason is that, by removing some of the data points, the iR for lower SSS is more affected than those with a higher SSS. This is clear if we consider that the effect of dropping data points are more significant for a small dataset rather than a large one, assuming they have the similar weights. The anomaly in the iR-PDF for low SSS can be interpreted as the result of an artificial forcing to avoid small iR values.

For the 'c1' dataset we clearly see that the deviation from the confidence regions from the mock data confirms the efficacy of the methodology presented.

6.6   Summary and conclusion

In this work we implemented the "Internal Robustness" of Ref. [61] to the currently available growth-rate data in the form of $f\sigma_8(z)$ shown in Table 6.1, with the aim to examine the dataset for systematics and outlier points in a fully automated manner. The "Internal Robustness" is a fully Bayesian approach which is not only sensitive to the local minimum like a standard χ^2 comparison, but also to the entire likelihood and can in principle detect the presence of systematics in the dataset. The method works by analyzing combinations of subsets in the dataset in a Bayesian model comparison way, potentially finding groups of outliers, data affected by systematics or groups that might follow different physics.

The main goal of our approach was to identify systematic-contaminated data-points, which can then be further analyzed and potentially excluded if they cannot be corrected. Furthermore, in order to validate our analysis and assess its sensitivity we also performed several cross-checks, for example by removing some of the data points or by artificially contaminating some points, while we also generated mock datasets in order to estimate confidence regions of the iR.

We found that, in the first case, when removing the two points with the least robustness the iR for SSS>6 has a higher minimum but the maximum iR is lower, as well as the mean iR values are smaller with respect to the full dataset. In the second case, by adding an artificially contaminated point which was $\sim 5\sigma$ away from its actual value, we found that the internal robustness method indeed detected the inconsistency caused by the contaminated datapoint.

In conclusion, implementing the iR methodology we found that the $f\sigma_8(z)$ dataset, used in our analysis and shown in Table 6.1, is internally robust showing no anomalous behavior, thus ensuring its internal robustness. This is interesting when discussing the tension of the Planck 15 CMB data and the low redshift measurements coming from galaxy surveys, as we can make sure that the discrepancy does not originate from inconsistencies in the data.

Chapter 7

Comparing Dark Energy models with Hubble versus Growth Rate data

This chapter presents a research work submitted to Physical Review D, which corresponds to Ref. [96]. This work is a team effort with aid from the student Javier Silva, who mainly developed the perturbation equations for the different models considered, especially the analytical ones, plus help with coding extensions.

Abstract

In this work we perform an analysis on the recently proposed conjoined cosmic growth and cosmic expansion diagram [97] to compare several dark energy models using the Figure of Merit showed in [98], which consists in the inverse of the 1σ confidence region in the $f\sigma_8(z)-H(z)$ plot. Our analysis also consists of comparing the models by performing different statistical criteria: Bayes factor [99], the Bayesian Information Criteria [100] and the Akaike Information Criterion [101]. We also developed a 3-dimensional Figure of Merit to account simultaneously for the errors on the growth rate and the Hubble parameter. The main idea is to consider several cosmological models and compare them with the different statistical criteria in order to highlight the differences and the accuracies of each single criterion.

7.1   Introduction

Recent observations [13, 14] pointed out that the Universe seems to be in a phase of accelerated expansion. These evidences have led cosmologists to revise the theory of the expansion of the Universe either by introducing a new form of matter called dark energy [19] or by modifying directly the laws of gravity [62].

Within the framework of Friedmann-Lemaître-Robertson-Walker (FLRW) cosmologies,

such accelerated expansion can be generated by adding up a simple cosmological constant Λ to the total budget of the Universe. Even though the latter gives rise to severe coincidence and fine-tuning problems, observations seem in agreement with such an explanation [102, 63, 90]. Over the years a series of dark energy models have been considered in order to solve, or at least alleviate, the theoretical problems related to dark energy.

Alternative theories of gravity came naturally as a consequence of the incapability of having a self-consistent model of dark energy. This class of models intends to modify General Relativity (GR) and to explain the observed acceleration of the Universe as a purely geometrical effect due to the weakening of gravity at very large scales.

The important question here is whether the two scenarios can be distinguished. It is well known that any Hubble expansion can be generated by choosing an appropriate equation of state for the dark energy, see [103]. Over the years there have been claims that it is possible to distinguish alternative theories of gravity from dark energy models by using growth data; however the last assumption is not always true unless the expansion history is fixed, [104]. Nonetheless, recent works have proposed to study the cosmic growth versus cosmic expansion history conjoined diagram, the $f\sigma_8 - H$ plot, to put constraints on the parameter space of cosmological models, or to compare different models directly [97]. Model comparison using this approach has already been investigated in [105, 98]. The advantage of the $f\sigma_8 - H$ plot over other probes relies on the degeneracy break of the history curves when comparing different models or the parameter space, since it contrasts a geometrical observable, $H(z)$, to a non-geometrical one, $f\sigma_8(z)$, which is a pure gravitational effect.

Using this approach, dark energy models were compared using the $f\sigma_8 - H$ plot [98] through the FoM defined as the inverse area of the 1σ confidence region in the conjoined diagram. In this work, we follow a similar approach and we compare the models using different statistical tools: the standard Bayesian evidence [99], the Bayesian Information Criteria [100] (BIC), the Akaike Information Criterion [106] (AIC) and the FoM. Furthermore, we considered an extension to the FoM which we define 3-FoM, which considers both errors on $f\sigma_8(z)$ and $H(z)$.

Anticipating the results, we find that the FoM is a fairly good estimator of the errors, however, its extension, the 3-FoM, captures simultaneously the growth of matter and the expansion history making it more stable over different models. The criteria BIC and AIC_c penalize substantially models with extra parameters.

The work is structured as follows: in Section 7.2 we report the basic equations that will be used in our work, whereas in Section 7.3 we list the cosmological models that will be compared, and the link between H and $f\sigma_8$ measurements. In Section 7.4 we show the datasets used in the analysis and the statistical methodology is reported in Section 7.5. In Section 7.6 we report the results of our analysis and discuss them.

7.2   Basic equations

By gravitational collapse, matter forms structures in the universe, which are called perturbations $\delta\rho(a, k)$, where k represents the scale in Fourier space. These perturbations grow over time according to the characteristics of the fluid: EoS, pressure perturbation δp and anisotropic stress σ.

The growth of perturbations for a general fluid is governed, assuming homogeneity and isotropy, by the differential equations [38]

$$\delta' = \quad 3 \quad (1+w)\,\phi' - \frac{V}{H\,a^2} - \frac{3}{a}\left(\frac{\delta p}{\rho} - w\delta\right)\,, \tag{7.1}$$

$$V' = \quad - \quad (1-3w)\frac{V}{a} + \frac{k^2}{H\,a^2}\frac{\delta p}{\rho} + (1+w)\frac{k^2}{H\,a^2}\psi$$

$$\quad - \quad (1+w)\frac{k^2}{H\,a^2}\sigma\,, \tag{7.2}$$

where the primes denote derivatives with respect to a, $\delta = \delta\rho(a, k)/\rho(a)$ is the density contrast, $V = \mathrm{i}\,k_j T_0^j/\rho(a)$ is the scalar velocity perturbation, and T_j^i is the energy-momentum tensor. The quantities ψ and ϕ are the gravitational potentials in the Newtonian gauge. These potentials follow

$$k^2\phi = -4\pi G a^2 \sum_j \rho_j \left(\delta_j + \frac{3aH}{k^2}V_j\right)\,, \tag{7.3}$$

$$k^2\,(\phi - \psi) = 12\pi G a^2 \sum_j (\rho_j + p_j)\,\sigma_j\,, \tag{7.4}$$

where the sum runs over all the species in the Universe. We will then have sets of equations of the form of Eqs. (7.1) and (7.2) depending on the number of species present in the Universe. For non-relativistic particles, i.e. cold dark matter and baryons, we just need to set $w = \delta p = \sigma = 0$. However, in this work we consider general dark energy models as well. There is no unique way to parametrize these quantities as they depend directly on the specific model considered.

For simplicity, in this work, we consider only two components, and they are pressureless dark matter and a dark energy fluid. In the next section, we describe the different dark energy models.

Since we want to test our models with observations, we need to obtain a measurable quantity; the real observable is $f\sigma_8(a)$, defined as the product of the growth rate of matter perturbations $f(a) = \mathrm{d}\ln\delta_m(a)/\mathrm{d}\ln a$ and the root mean square (RMS) of matter density perturbations measured in a sphere of $8\,h^{-1}\mathrm{Mpc}$, defined as $\sigma_8(a) = \sigma_{8,0}\delta_m(a)/\delta_m(a = 1)$. We then have:

$$f\sigma_8(a) = \sigma_{8,0}\,a\frac{\delta'_m(a)}{\delta(a = 1)}\,, \tag{7.5}$$

where $\sigma_{8,0}$ is the RMS measured today. This quantity is more reliable than $f(a)$ alone due its independancy of the bias b, which is the ratio of baryon perturbations to total matter perturbations, i.e. $\delta_b = b\,\delta_m$.

7.3 Models

Here we list the models considered in the analysis. Throughout this work, we assume that all the dark models have zero anisotropic stress, $\sigma = 0$. Consequently, the two gravitational potentials are equal $\phi = \psi$.

7.3.1 ΛCDM

This corresponds to the simplest and most accepted cosmological model. It assumes a constant EoS parameter exactly equal to -1. We consider two different cases in the ΛCDM scenario.

ΛCDM: this model refers to flat ΛCDM (without spatial curvature) where we set the curvature parameter $\Omega_{k,0} = 0$, hence the Hubble parameter reads

$$H^2 = H_0^2 \Big[\Omega_{m,0} a^{-3} + (1 - \Omega_{m,0}) \Big].$$

Furthermore, the cosmological constant Λ has zero perturbations, hence the system of equations simplifies and the gravitational potentials only depend on pressureless matter. For small scales, Eqs. (7.1) - (7.4) reduce to a single second-order differential equation for matter density contrast to which an analytical solution[1], expressed in terms of hypergeometric functions, can be found, see Appendix 7.8 for more details.

Consequently, we will have one model with two variants: ΛCDM and ΛCDM-a, using the numerical and analytic solution, respectively. However, for consistency reasons, we decided to use the full set of differential equations Eqs. (7.1) - (7.4), leaving to the appendix the results obtained by using the analytical solutions as a further test.

Finally, the parameters of both models are:

$$\boldsymbol{\theta}_{\Lambda\text{CDM}} = (\Omega_{m,0},\ H_0,\ \sigma_{8,0})\,. \tag{7.6}$$

ΛCDM-nf: this model corresponds to a non-flat (to which we use the label '-nf') ΛCDM where we allow for the curvature parameter to vary. Then, the Hubble parameter takes the form:

$$H^2 = H_0^2 \Big[\Omega_{m,0} a^{-3} + (1 - \Omega_{m,0} - \Omega_{\text{de},0}) a^{-2} + \Omega_{\text{de},0} \Big].$$

The cosmological constant still has zero perturbations, however the differential equation for matter perturbations does not have an analytical solution, hence we solve numerically Eqs. (7.1) - (7.4).

The parameters of the model are:

$$\boldsymbol{\theta}_{\Lambda\text{CDM-nf}} = (\Omega_{m,0},\ \Omega_{\text{de},0},\ H_0,\ \sigma_{8,0})\,. \tag{7.7}$$

---

[1] We denote analytically-solved models by using the label 'a'.

7.3.2 w**CDM**

This model is an extension of the ΛCDM model in which a constant EoS w is set as a free parameter. If the EoS parameter of dark energy is no longer constant and equal to -1, then dark energy may have perturbations and its growth will be fully characterized by the values of w and $c_s{}^2$. Clearly, if dark energy has perturbations, these will affect the growth of matter perturbations through the gravitational potential Eqs. (7.3) - (7.4). As a consequence, we identify four different cases.

w**CDM:** this model corresponds to flat wCDM where perturbations in the dark energy sector have been switched off; the Hubble parameter reads:

$$H^2 = H_0^2\Big[\Omega_{m,0}a^{-3} + (1 - \Omega_{m,0})a^{-3(1+w)}\Big]. \tag{7.8}$$

If we decide to ignore *a priori* the dark energy perturbations, then the growth of matter density is still governed by a second order differential equation, and it is still possible to find an analytical solution to the matter density contrast, see Appendix 7.8 for more details. As a consequence we have the wCDM and wCDM-a solutions to this model. As for the ΛCDM case, we also consider the full numerical solutions from Eqs. (7.1) - (7.4) and leaving the results from the analytical solution to the appendix.

Finally, the parameters of both models are

$$\boldsymbol{\theta}_{\mathrm{wCDM}} = (\Omega_{m,0},\ w,\ H_0,\ \sigma_{8,0})\ . \tag{7.9}$$

w**CDM-nf:** this model corresponds to a non-flat wCDM; the Hubble parameter reads

$$H^2 = H_0^2\Big[\Omega_{m,0}a^{-3} + (1 - \Omega_{m,0} - \Omega_{\mathrm{de},0})a^{-2} + \Omega_{\mathrm{de},0}a^{-3(1+w)}\Big]. \tag{7.10}$$

Here we set dark energy perturbations to zero. However, due to the complexity of the Hubble parameter, analytical solutions for the matter density contrast do not exist and we solve numerically the system of Eqs. (7.1) - (7.4).

We have the following free parameters for the model:

$$\boldsymbol{\theta}_{\mathrm{wCDM\text{-}nf}} = (\Omega_{m,0},\ \Omega_{\mathrm{de},0},\ w,\ H_0,\ \sigma_{8,0})\ . \tag{7.11}$$

w**CDM-p:** this model is a flat wCDM for which we allow perturbations (this addition is symbolized by '-p') in the dark energy sector. The Hubble parameter is given by Eq. (7.8). However, we now have two sets of equations (7.1) - (7.2), for pressureless matter and for the dark energy fluid. Analytical solutions can also be found in some special limits, see Appendix 7.8. However, as for the other cases, we also use the full numerical solutions from the equation of perturbations.

As mentioned earlier, the growth of the perturbations of one species depends on the characteristics of the fluid, which are given by w, δp and σ. For pure pressureless matter,

---

[2]We remind the reader that we assume the anisotropic stress of any dark energy model to be zero.

$w = \delta p = \sigma = 0$. For a dark energy fluid, these are not known. In this work we assume zero anisotropic stress $\sigma = 0$, and the pressure perturbation to be given by [64]:

$$\delta p = c_s^2 \rho \delta + \frac{3aH(c_s^2 - c_a^2)}{k^2}\rho V \,, \tag{7.12}$$

where $c_a^2 \equiv \dot{p}/\dot{\rho}$ is the adiabatic sound speed of the fluid that can be expressed as

$$c_a^2 = w - \frac{\dot{w}}{3H(1+w)} = w - \frac{w'}{3(1+w)} \,, \tag{7.13}$$

and for a constant EoS, the adiabatic sound speed yields $c_a^2 = w$.

The free parameters of the models (wCDM-p and wCDM-p-a) are

$$\boldsymbol{\theta}_{\text{wCDM-p}} = (\Omega_{m,0}, \, w, \, c_s^2, \, H_0, \, \sigma_{8,0}) \,. \tag{7.14}$$

w**CDM-nf-p:** this model correspond to non-flat wCDM for which we allow perturbations in the dark energy sector; the Hubble parameter takes the form in Eq. (7.10) and the perturbations will be solved numerically for both matter and dark energy. Thus, the parameter set of the model is

$$\boldsymbol{\theta}_{\text{wCDM-nf-p}} = (\Omega_{m,0}, \, \Omega_{\text{de},0}, \, w, \, c_s^2, \, H_0, \, \sigma_{8,0}) \,. \tag{7.15}$$

7.3.3 Chevallier-Polarski-Linder (CPL)

This class of models [26, 27] can be considered an extension to wCDM models in which the equation of state depends on the scale factor. The simplest extension is a Taylor expansion around the present time $a = 1$, giving

$$w(a) = w_0 + w_a(1 - a) \,. \tag{7.16}$$

Hence, giving two extra parameters: w_0, which is the present time EoS parameter and w_a which represents the variation over time of $w(a)$. We identify four different models using this parametrization. The quantity $\hat{w}$ is the effective EoS parameter, given by

$$\hat{w}(a) = \frac{1}{\ln a} \int_1^a \frac{w(x)}{x} \mathrm{d}x \,. \tag{7.17}$$

CPL: this corresponds to the simplest scenario where the Hubble parameter does not depend on curvature and we set dark energy perturbations to zero. Then, the Hubble parameter reads

$$H^2 = H_0^2 \left[\Omega_{m,0} a^{-3} + (1 - \Omega_{m,0}) a^{-3(1+\hat{w}(a))} \right]. \tag{7.18}$$

There is no exact analytic expression for the matter density contrast when the EoS parameter takes the form of Eq. (7.16). Here we solve numerically Eqs. (7.1) - (7.4). This way, the parameters are

$$\boldsymbol{\theta}_{\text{CPL}} = (\Omega_{m,0}, \, w_0, \, w_a, \, H_0, \, \sigma_{8,0}) \,. \tag{7.19}$$

CPL-nf: in this model we allow the curvature parameter to vary. Then, Hubble parameter becomes

$$H^2 = H_0^2 \left[\Omega_{m,0} a^{-3} + (1 - \Omega_{m,0} - \Omega_{\mathrm{de},0}) a^{-2} + \Omega_{\mathrm{de},0} a^{-3(1+\hat{w}(a))} \right]. \qquad (7.20)$$

We set dark energy perturbations to zero and solve numerically the Eqs. (7.1) - (7.4) for pure pressureless matter only. The free parameters of the model are:

$$\boldsymbol{\theta}_{\mathrm{CPL\text{-}nf}} = (\Omega_{m,0},\, \Omega_{\mathrm{de},0},\, w_0,\, w_a,\, H_0,\, \sigma_{8,0}) \,. \qquad (7.21)$$

CPL-p: the Hubble parameter is given by Eq. (7.18), the equation of perturbations will be solved numerically by using Eqs. (7.1) - (7.4). The characteristics of the dark energy fluid are given by Eq. (7.12), with the further assumption that the adiabatic sound speed $c_a^2 = w$; the former is somehow required in order to stabilize the growth of dark energy perturbations when it crosses the phantom regime [64]. Thus, the parameters of the model are

$$\boldsymbol{\theta}_{\mathrm{CPL\text{-}p}} = (\Omega_{m,0},\, w_0,\, w_a,\, c_s^2,\, H_0,\, \sigma_{8,0}) \,. \qquad (7.22)$$

CPL-nf-p: the Hubble parameter takes the form in Eq. (7.20). We solve numerically Eqs. (7.1) - (7.4) for pressureless matter and dark energy. The characteristic of the dark energy fluid are given by Eq. (7.12), with the further assumption of $c_a^2 = w$ and $w' = 0$ at crossing.
The parameter set for the model is

$$\boldsymbol{\theta}_{\mathrm{CPL\text{-}nf\text{-}p}} = (\Omega_{m,0},\, \Omega_{\mathrm{de},0},\, w_0,\, w_a,\, c_s^2,\, H_0,\, \sigma_{8,0}) \,. \qquad (7.23)$$

7.4 Data

The Hubble parameter data for the analysis are the *cosmic chronometers* compilation used in [107], which consists in 31 independent measurements of $H(z)$, obtained from evolving galaxies at different redshifts [108].

The growth rate dataset is based on the compilation used in [60], which is an updated version of the 'Gold-2017' dataset from [70]. The dataset consists of 22 independent measurements of $f\sigma_8(z)$, obtained through baryon acoustic oscillations and weak lensing surveys. Among these surveys, it is important to note that the three WiggleZ [73] and the four SDSS-IV [74] measurements are correlated, and their covariance matrices are given by Eqs. (6.26)-(6.27).

7.5 Methodology

To perform the analysis, both datasets are assumed to have Gaussian likelihood distributions, this is the probability of the data given a set of parameters. The datasets are assumed to be independent, thus their conjoined likelihood is the product of each dataset's likelihood. In

z	$H(z)$	$\sigma_H(z)$	Ref.	z	$H(z)$	$\sigma_H(z)$	Ref.
0.07	69.0	19.6	[109]	0.4783	80.9	9.0	[108]
0.09	69.0	12.0	[110]	0.48	97.0	62.0	[111]
0.12	68.6	26.2	[109]	0.593	104.0	13.0	[112]
0.17	83.0	8.0	[110]	0.68	92.0	8.0	[112]
0.179	75.0	4.0	[112]	0.781	105.0	12.0	[112]
0.199	75.0	5.0	[112]	0.875	125.0	17.0	[112]
0.2	72.9	29.6	[109]	0.88	90.0	40.0	[111]
0.27	77.0	14.0	[110]	0.9	117.0	23.0	[110]
0.28	88.8	36.6	[109]	1.037	154.0	20.0	[112]
0.352	83.0	14.0	[112]	1.3	168.0	17.0	[110]
0.3802	83.0	13.5	[108]	1.363	160.0	33.6	[113]
0.4	95.0	17.0	[110]	1.43	177.0	18.0	[110]
0.4004	77.0	10.2	[108]	1.53	140.0	14.0	[110]
0.4247	87.1	11.2	[108]	1.75	202.0	40.0	[110]
0.4497	92.8	12.9	[108]	1.965	186.5	50.4	[113]
0.47	89.0	49.6	[114]				

Table 7.1: The 31 cosmic chronometer data points used in this analysis along with their related references. The $H(z)$ and $\sigma_H(z)$ data are in units of $\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

terms of the traditional chi-squared, defined by $\chi^2 \equiv -2\log L$, where L is the likelihood of the current model, it is simply given by the sum of each dataset's chi-squared, or

$$\chi^2 = \chi_H^2 + \chi_{f\sigma_8}^2 \tag{7.24}$$

Where the subscripts '$f\sigma_8$' and 'H' indicate growth and expansion contributions, respectively.

Let us suppose that there are n measurements of H or $f\sigma_8$, so we represent the observed data in different redshifts as $\mathbf{m} = (m(z_1), \ldots, m(z_n))$ and its theoretical prediction as $\boldsymbol{\mu}(\boldsymbol{\theta}) = (\mu(z_1), \ldots, \mu(z_n))$, which depend on the cosmological model and parameters. We define the data vector as

$$\mathbf{x}_s = \mathbf{m}_s - \boldsymbol{\mu}_s , \tag{7.25}$$

with the subscript 's' denoting the data source: H or $f\sigma_8$. However, in the case of growth measurements, we need to take into account a redshift correction, which is featured in Ref. [70]. This correction consists in the following factor

$$\mathrm{fac}(z_i) = \frac{H(z_i)\,d_A(z_i)}{H^{\mathrm{ref},i}(z_i)\,d_A^{\mathrm{ref},i}(z_i)} \tag{7.26}$$

where $d_A(z)$ is the angular diameter distance and the superscript 'ref, i' indicates that the reference cosmology is taken on the corresponding data point at redshift z_i. With this procedure, we arrive at the corrected growth theoretical prediction:

$$\mu_c^i = \frac{\mu_{f\sigma_8}^i}{\mathrm{fac}(z_i)} . \tag{7.27}$$

For all the datapoints, the reference model used is ΛCDM, and one can note that the product $H(z)\,d_A(z)$ is independent of H_0 and $\sigma_{8,0}$ for all models considered. We list the reference values for $\Omega_{m,0}$ of each datapoint in Table 6.1. Using the corrected prediction, the data vector for $f\sigma_8$ is

$$\mathbf{x}_{f\sigma_8} = \mathbf{m}_{f\sigma_8} - \boldsymbol{\mu}_c \ . \tag{7.28}$$

Therefore, the chi-squared are constructed through

$$\chi_s^2 = \mathbf{x}_s^T \mathbf{C}_s^{-1} \mathbf{x}_s \ , \tag{7.29}$$

where $\mathbf{C}_s^{-1}$ the inverse of the covariance matrix of the dataset. In the case of cosmic expansion, the covariance matrix is diagonal and equal to each datapoint's variance. The total cosmic growth covariance matrix is given by a diagonal matrix with the measurements' variance, with the insertion of the WiggleZ matrix and SDSS-IV matrices, given by Eqs. (6.26) and (6.27).

Parameter	Flat prior limits
$\Omega_{m,0}$	$[0, 1]$
$\Omega_{\mathrm{de},0}$	$[0, 1.7]$
w_0	$[-3.5, -1/3]$
w_a	$[-2.5, -1/3 - w_0]$
c_s^2	$[0, 1]$
H_0 [Mpc/km/s]	$[35, 110]$
$\sigma_{8,0}$	$[0.3, 1.5]$

Table 7.2: Ranges of the flat priors used for each parameter. Note that w_a depends on the value of w_0 to define its upper bound. This is to ensure that $w(a) < -1/3$ in order to have acceleration on the expansion of the Universe.

We now proceed to present the methods used to compare different dark energy models. We use five methods in total:

Evidence: The first method is the standard Bayesian model comparison via evidence computation $\log(E)$ [99], where the evidence is defined via

$$E(\mathbf{m}\,|\,M) = \int L(\mathbf{m}\,|\,\boldsymbol{\theta}_M, M)\,\pi(\boldsymbol{\theta}_M\,|\,M)\,\mathrm{d}\boldsymbol{\theta}_{\mathrm{M}} \ . \tag{7.30}$$

The former quantity determines the probability of a given model M to be true, given the data $\mathbf{m}$. As already mentioned, the likelihood function $L(\mathbf{m}\,|\,\boldsymbol{\theta}_M, M)$ is Gaussian on the data $\mathbf{m}$, and the prior probability for the parameters, $\pi(\boldsymbol{\theta}_M\,|\,M)$. If we assume the prior probabilities $\pi(M)$ to be the same for each model, then the Evidence completely defines the ranking of the cosmological models.

All throughout the analysis, and specifically for the evidence computation, we adopted standard flat priors for all the parameters, with boundaries reported in Table 7.2. Despite that, the only special treatment was made on w_a, for which we used an upper bound that depends on the value of w_0 in order to guarantee a phase of accelerated expansion [26].

Furthermore, we use the same priors for all models that have the free parameter, as we are more concerned about the statistical methods used.

The computation is performed using the package Nestle [115], a Python implementation of the MultiNest algorithm [51, 52]. This algorithm is an efficient and robust way of computing the evidence integral, a numeric task that becomes too large to be grid-integrated. MultiNest also produces a Markov chain that can be reused as the MCMC-sample for the next method below.

Figure of Merit: With this method, the models are ranked by their FoM's defined in [98], which corresponds to the inverse of the 1σ confidence region area in the conjoined $f\sigma_8 - H$ plot given a redshift range. The likelihood is used to MCMC-sample in the parameter space of each model, and this parameter chain is used to get the 1σ range of $f\sigma_8(z_i)$ for $i \in 1, ..., n$. If there are sufficient z_i points, a spline can be constructed to connect the points in the $f\sigma_8 - H$ plane, keeping $H(z_i)$ fixed to its mean value. This method is viable because $f\sigma_8$ is much less constrained than H in all the models tested, and $H(z)$ increases monotonically with z for each model. The redshift range is, in principle, defined between $z = 0$ to $z_{max} = 2$, to include the whole redshift data range. We will also show how the FoM varies when z_{max} changes.

3-FoM: Here, we propose an extension of the previous method, in which we now consider the 1σ range of $H(z_i)$ (as opposed as in the last method where it was omitted). For a z_i point we obtain the values plus the associated confidence levels of the Hubble parameter and the growth rate, i.e.

$$H(z_i)^{+\sigma_{H(z_i)+}}_{-\sigma_{H(z_i)-}} \quad \text{and} \quad f\sigma_8(z_i)^{+\sigma_{f\sigma_8(z_i)+}}_{-\sigma_{f\sigma_8(z_i)-}} .$$

With these values we compute the ellipsoidal area on each redshift point z_i, as an approximation for the 2-dimensional confidence region in the $f\sigma_8(z_i), H(z_i)$ space

$$A_e(z_i) = \frac{\pi}{4} \Big(\sigma_{H(z_i)+} + \sigma_{H(z_i)-} \Big) \Big(\sigma_{f\sigma_8(z_i)+} + \sigma_{f\sigma_8(z_i)-} \Big) . \tag{7.31}$$

The 3-FoM is defined as the inverse of the ellipsoidal volume quantity in the $f\sigma_8, H, \sigma_H$ space.

$$
\begin{aligned}
V_e &= \int A_e(z) \mathrm{d}H(z) \\
&= - \int_{z=0}^{z=2} A_e(z) \frac{H'(z)}{(1+z)^2} \mathrm{d}z \\
&\simeq - \sum_i A_e(z_i) \frac{H'(z_i)}{(1+z_i)^2} \Delta z .
\end{aligned}
\tag{7.32}
$$

If there are many equispaced z_i points, the previous quantity corresponds to the volume enclosed in Fig. 7.1.

BIC: The fourth method is the Bayesian Information Criterion [100, 116] which is given by:

$$\mathrm{BIC} = 2\ln(N_{\mathrm{data}})n_{\mathrm{pars}} - 2\ln L_{\mathrm{max}} . \tag{7.33}$$

This method still considers the maximum likelihood L_{max}, however it tends to penalise models with several parameters through the direct dependence of n_{pars} . Its formulation aims at
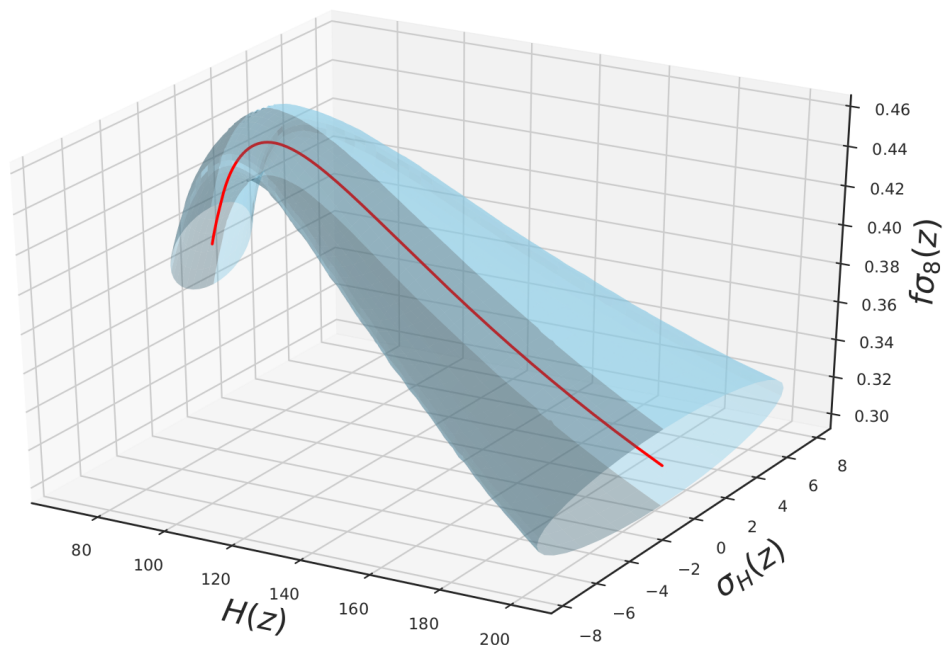
Figure 7.1: 3-FoM plot for a ΛCDM model. This shows the volume to be integrated for the calculation of the 3-FoM.

approximating the evidence (specifically, $-2\log(E)$) of the model to be tested, hence the favored model is the one with the lowest BIC value.

AIC$_c$. The last statistical method is the corrected Akaike Information Criterion (AIC$_c$) [101]. This method is similar to the BIC method because it still penalises models with several parameters, however the penalisation is weighted with the number of data. Contrary to the BIC test, the AIC$_c$ tends to favor one model if the data set is large enough. The criterion is given by:

$$\mathrm{AIC}_c = 2n_{\mathrm{pars}} - 2\ln L_{\mathrm{max}} + \frac{n_{\mathrm{pars}}(n_{\mathrm{pars}} + 1)}{N_{\mathrm{data}} - n_{\mathrm{pars}} - 1}\,. \tag{7.34}$$

This equation, derived in [101], accounts for a correction term when the number of data is small, unlike the original Akaike Information Criterion [106]. As before, the test should also be similar to the value of $-2\log(E)$, which means that the lower AIC$_c$ is, the more favored is the model.

7.6   Results and Discussion

In this section, we discuss the results found for each model and we compare the values of the criteria used. As mentioned previously the goal of the work is to accurately test the common criteria found in literature and to highlight their differences.

In general, we are not interested in the specific value of the criterion found for a particular model but rather their difference between two models. This difference will tell us which is the model that is able to better reproduce the data.

86

Model	$\log(E)$	FoM	3-FoM	BIC	AIC_c	H_{max}
ΛCDM	-21.87	0.192	0.027	51.35	34.02	201.71
ΛCDM-nf	-22.09	0.145	0.012	58.84	35.91	210.06
wCDM	-23.50	0.124	0.013	59.31	36.39	204.89
wCDM-p	-23.01	0.124	0.014	67.21	38.78	204.47
wCDM-nf	-23.33	0.125	0.010	66.57	38.14	207.54
wCDM-nf-p	-23.29	0.122	0.010	74.80	40.99	207.36
CPL	-24.16	0.129	0.014	67.18	38.75	208.97
CPL-p	-24.14	0.127	0.014	74.87	41.05	208.69
CPL-nf	-24.53	0.119	0.010	74.51	40.69	204.65
CPL-nf-p	-24.53	0.120	0.010	82.62	43.52	204.33

Table 7.3: Results of the different methods for each model. We also show $H_{max} = H(z = 2)$ to compare the extension of the integration in the H-dimension for the FoM and 3-FoM methods.

For the first criterion, i.e. the evidence E, we use Jeffrey's scale to assess the difference of the logarithmic evidences for two particular models.

The other two criteria, i.e. BIC and AIC_c, are directly connected to the likelihood of the models and hence they can be used as model selection tests. Since they come from a Taylor expansion around the maximum likelihood estimator of the likelihood function, they can be connected to Jeffrey's scale, however, this interpretation must be taken with care, see [117] for a detailed discussion. Generally, we can still consider the difference $|\Delta\text{BIC}| = |\text{BIC}_2 - \text{BIC}_1|$, where the index 2 refers to the model with the higher value of BIC and the index 1 to the one with the lower, as a good model selection test. Specifically, if $|\Delta\text{BIC}| \leq 2$, then there is no evidence in support of a model, if $2 < |\Delta\text{BIC}| \leq 6$, then there is a positive evidence in favor of the model with the smaller value, whereas if $|\Delta\text{BIC}| > 6$, the evidence is considered to be strong. The same discussion applies to the AIC_c criterion, where in this case we have: if the difference is less than 2, then both models are able to reproduce the data with the same accuracy, if $|\Delta\text{AIC}|$ is between 2 and 4, then there is a positive evidence for the model with the lower AIC_c, instead if $|\Delta\text{AIC}| > 10$, then the model with the larger AIC_c is strongly disfavored, see [118].

The last two criteria considered in this work are the FoM, defined as the inverse of the enclosed area at 1σ level for $f\sigma_8(z)$, and the 3-FoM defined as the inverse of the enclosed volume at 1σ level in both $f\sigma_8(z)$ and $H(z)$. It is clear the FoM and its extension (3-FoM) are not criteria able to favor/disfavor a model, but rather they give an estimation on the sensitivity of the parameters according to the data used. In practice, a larger FoM and/or 3-FoM means that the model is better constrained by the data.

Fig. 7.2 (top panel) shows the reconstruction of the $H(z) - f\sigma_8(z)$ assuming flat and non-flat ΛCDM as the cosmological model. The shaded areas are obtained directly from the 1σ errors of the parameters given by the MCMC samples. The best fit of the parameters are reported in Tab. 7.4. For this particular model the addition of an extra parameter,
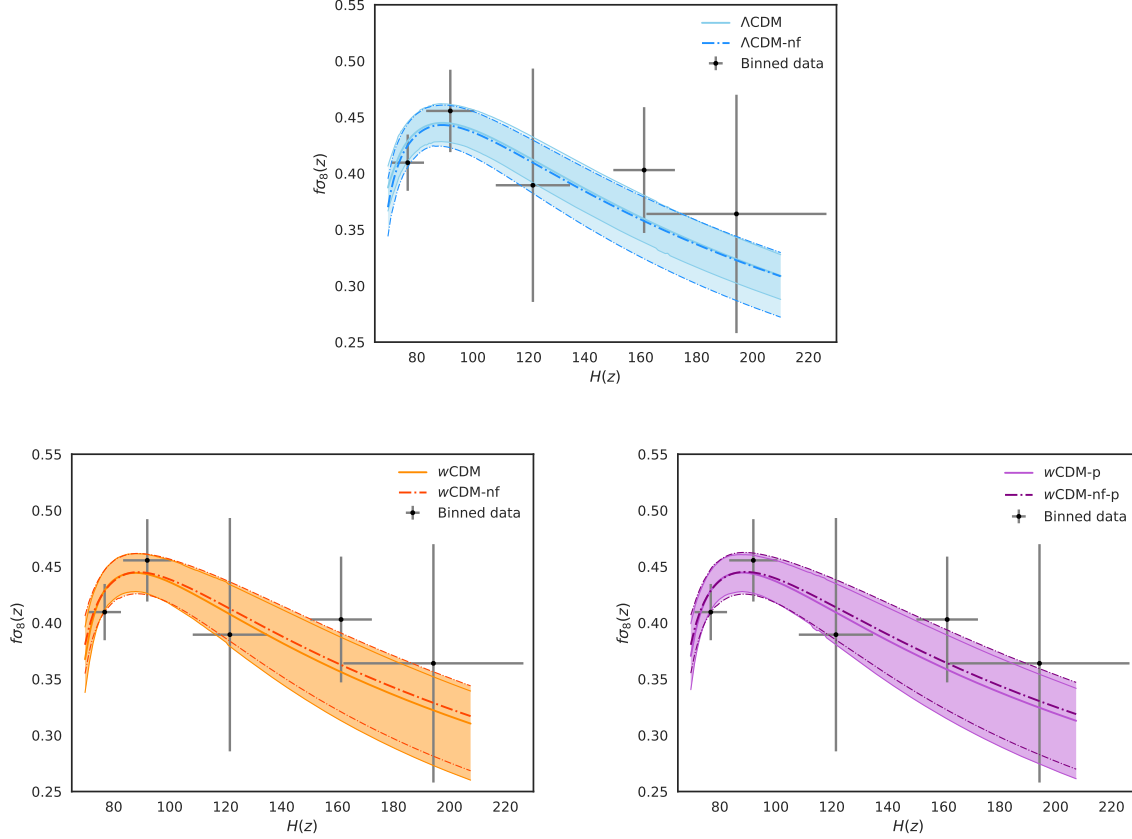
Figure 7.2: The conjoined plots of the cosmic growth $f\sigma_8(z)$ versus the cosmic expansion $H(z)$ for differents models described in the text: (upper panel) ΛCDM with ΛCDM-nf, (lower left panel) wCDM with wCDM-nf and (lower right panel) wCDM-p with wCDM-nf-p. Also the 1σ error regions (shaded areas) and the real binned data (gray points) are shown.

$\Omega_{\text{de},0}$, alters the results and the two shaded areas differ, specially at high redshift where the lower limit of the errors are larger for not flat ΛCDM: as a consequence the FoM and 3-FoM decrease of about 35% and 125%, respectively. The BIC and AIC$_c$ criteria used in this analysis increase of about 13% and 5% when the curvature parameter is considered, see Tab. 7.3.

As for the model comparison, the evidence gives inconclusive results, the AIC$_c$ criterion favors positively the flat ΛCDM over the non-flat ΛCDM model, the BIC criterion instead shows a strong evidence in favor of ΛCDM.

In Fig. 7.2 (lower left panel) are shown the reconstruction of the $H(z) - f\sigma_8(z)$ assuming flat and non-flat wCDM. These models have one parameter more with respect to the corresponding ΛCDM models discussed above. The addition of w as a free parameter increases the confidence regions substantially, as it can be seen from the figures and also reported in Tab. 7.3, where the FoM decreases compared to previous cases. Here the variation is due to the parameter itself rather than the addition of an extra parameter; in fact, if we consider the non-flat ΛCDM model, which has the same number of parameters as wCDM

model, the FoM reduces from 0.145 to 0.124 which corresponds to almost 15%. However, the 3-FoM manifests an opposite behavior, it increases to about 8%. The reason is that $f\sigma_8(z)$ is sensitive to the variation of the parameters almost at any redshift, whereas the Hubble parameter is more sensitive at high redshifts (fixing one value of H_0, the variation on $H(z)$ can only appear when the z is increased). For the non-flat ΛCDM model the area enclosed by $f\sigma_8(z)$ is smaller than the era enclosed for the wCDM model, hence giving a lager FoM. However, the maximum value of the Hubble parameter is larger for non-flat ΛCDM model, 210.06 against 204.89 for the wCDM model. This effect is taken into account in the 3-FoM, where the errors on $H(z)$ are considered. The two effects are counterbalanced, giving almost the same value in the 3-FoM.

The evidence is weakly in support of the non-flat ΛCDM model over wCDM and the same is found for the BIC and AIC_c criteria. Adding curvature to the wCDM model makes the FoM increase of about 1% meaning that the 1σ errors are almost the same, however, their best-fits differ. The 3-FoM decreases for the non-flat model showing that their errors are less constrained. The BIC and AIC_c supports the flat model but the evidence is inconclusive.

Model	$\Omega_{m,0}$	$\Omega_{de,0}$	w_0	w_a	c_s^2	H_0	σ_8
ΛCDM	$0.286^{+0.032}_{-0.038}$	$1-\Omega_{m,0}$	-	-	-	69.7 ± 2.3	0.779 ± 0.039
ΛCDM-nf	0.37 ± 0.16	$0.83^{+0.29}_{-0.24}$	-	-	-	70.4 ± 3.1	$0.762^{+0.044}_{-0.084}$
wCDM	$0.280^{+0.045}_{-0.039}$	$1-\Omega_{m,0}$	$-1.11^{+0.38}_{-0.30}$	-	-	$70.6^{+4.2}_{-4.7}$	$0.782^{+0.045}_{-0.11}$
wCDM-p	$0.278^{+0.044}_{-0.037}$	$1-\Omega_{m,0}$	$-1.09^{+0.38}_{-0.30}$	-	0.50 ± 0.29	70.5 ± 4.4	$0.788^{+0.045}_{-0.11}$
wCDM-nf	$0.34^{+0.18}_{-0.22}$	$0.86^{+0.34}_{-0.41}$	$-1.08^{+0.49}_{-0.18}$	-	-	69.7 ± 4.2	$0.790^{+0.045}_{-0.11}$
wCDM-nf-p	$0.34^{+0.18}_{-0.22}$	$0.85^{+0.34}_{-0.42}$	$-1.07^{+0.50}_{-0.16}$	-	0.51 ± 0.29	$69.6^{+3.7}_{-4.7}$	$0.795^{+0.048}_{-0.11}$
CPL	$0.294^{+0.047}_{-0.041}$	$1-\Omega_{m,0}$	-1.20 ± 0.34	$-0.50^{+0.99}_{-0.46}$	-	71.9 ± 4.5	$0.747^{+0.026}_{-0.099}$
CPL-p	$0.293^{+0.046}_{-0.041}$	$1-\Omega_{m,0}$	-1.17 ± 0.33	$-0.50^{+1.0}_{-0.49}$	0.50 ± 0.29	71.7 ± 4.4	$0.751^{+0.030}_{-0.10}$
CPL-nf	$0.27^{+0.12}_{-0.24}$	$0.72^{+0.21}_{-0.45}$	$-1.27^{+0.63}_{-0.27}$	$-0.42^{+1.0}_{-0.47}$	-	$70.5^{+4.0}_{-4.6}$	$0.778^{+0.048}_{-0.11}$
CPL-nf-p	$0.27^{+0.11}_{-0.26}$	$0.72^{+0.23}_{-0.45}$	$-1.28^{+0.65}_{-0.27}$	$-0.38^{+0.98}_{-0.45}$	0.50 ± 0.29	$70.4^{+3.9}_{-4.5}$	$0.778^{+0.054}_{-0.10}$

Table 7.4: Parameter constraints derived from Nested Sampling to each (non-analytical) model described in the text.

In Fig. 7.2 (lower right panel) we show the reconstruction of the $H(z) - f\sigma_8(z)$ assuming flat and non-flat wCDM with the further addition of perturbations in the dark energy sector parameterized with c_s^2 as an extra free parameter. If we compare the latest results with the former case we realize that the FoM does not change from wCDM to wCDM-p, whereas it decreases of about 2.4% from wCDM-nf to wCDM-nf-p . These negligible variations are repeated for the 3-FoM that does not change from wCDM-nf to wCDM-nf-p and it increases of about 7.1% from flat wCDM to wCDM-p. As expected, dark energy perturbations are weakly constrained with the data available (dark energy perturbations affect only the growth of matter). This is shown in Tab. 7.4 where the best fits of the models with and without dark energy perturbations are basically the same. This behavior is shown in all the criteria used in this work, except for the BIC criterion which indeed favors the model without dark energy perturbations. However, this is a pure mathematical effect as the BIC criterion always penalizes the model with extra parameters.

In Fig. 7.3 (left panel) are shown the reconstructions of the $H - f\sigma_8(z)$ assuming flat and non-flat CPL. If we look at Tab. 7.3, we realized that the FoM constrains better CPL over
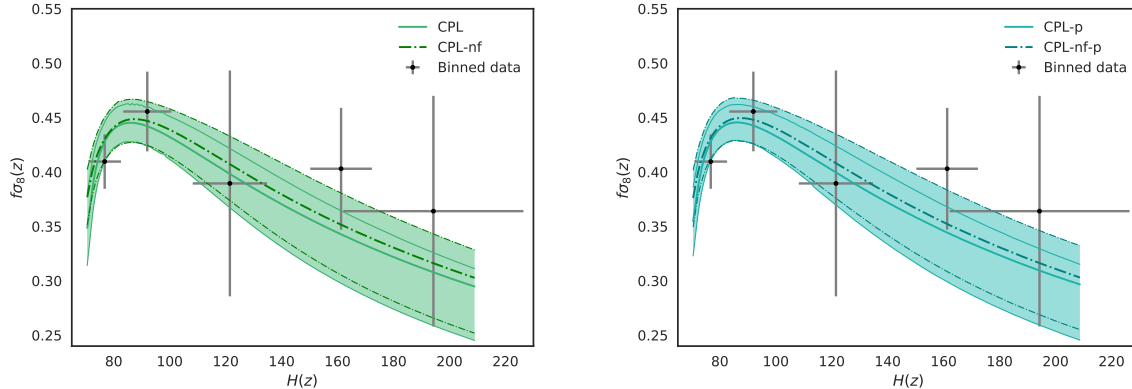
Figure 7.3: The conjoined plots of the cosmic growth $f\sigma_8(z)$ versus the cosmic expansion $H(z)$ for differents models described in the text: (left panel) CPL with CPL-nf and (right panel) CPL-p with CPL-p-nf. Also the 1σ error regions (shaded areas) and the real binned data (gray points) are shown.

wCDM which might sound peculiar because one would naively expect that a model with more parameters has larger 1σ errors. Here, the difference in the FoM comes from the asymmetric values of the errors on w_a; this asymmetry is due to the choice of the prior for w_a, for which we chose to bind it to w_0 in order to guarantee an accelerated expansion. This asymmetry led to a smaller area in the upper part, reducing the enclosed 1σ area of $f\sigma_8(z)$. The 3-FoM is more stable and this is again due to the value of the Hubble parameter at high redshifts: for the wCDM model $H(z=2) = 204.89$, whereas for CPL model is 208.97. This 2% difference is accounted in the final 3-FoM which decreases with respect to its companion. The evidence gives inconclusive results, manifesting the negligible effects of w_a on the two observables; the same conclusion is obtained with the AIC_c criterion. However, the BIC criterion strongly penalizes CPL just because of the extra parameter in the model.

The CPL, wCDM-nf, and wCDM-nf models have the same number of parameters, thus BIC and AIC_c criteria change less than 1% between them, showing again that they depend strongly on the number of parameters. The FoM and 3-FoM show that CPL is better constrained but, as mentioned, this is due to the priors on w_a used. The evidence weakly favors the wCDM-p model over CPL but it is inconclusive with respect of wCDM-nf.

The non-flat CPL have the same number of parameters as wCDM-nf-p, but the FoM shows that CPL is better constrained by the data whereas the 3-FoM does not change. The BIC and AIC_c change less than 1% and the evidence weakly favors wCDM-nf-p model.

In Fig. 7.3 (right panel) we show the reconstruction of the $H - f\sigma_8(z)$ assuming flat and non-flat CPL-p models. The behavior is similar to the previous case (CPL versus CPL-nf). By adding the curvature parameter the evidence is inconclusive and the other indicators favor the flat model because it has one parameter less. When we take into account dark energy perturbations into the CPL models, we obtain a similar behavior as seen for the wCDM models. Again, with the available data we are not able to constrain c_s^2, hence all the criteria are insensitive to the variation of the sound speed. The only exceptions are BIC and AIC_c

criteria, which penalize the addition of the sound speed into the analysis.

For completeness we also performed our analysis using the analytical solutions for the growth rate of matter, the models are ΛCDM, wCDM and wCDM with dark energy perturbations. The results are reported in the Appendix 7.8 and the results are shown in Tab. 7.6, whereas the best fit of these three models can be found in Tab. 7.7. All the three analytical models give results in excellent agreement with the full numerical analysis, demonstrating that the analytical solutions found in the literature are consistent.

Redshift bin	$H(z)$ [km s^{-1} Mpc^{-1}]	$f\sigma_8(z)$
$0 < z \leq 0.4$	76.8 ± 5.8	0.410 ± 0.025
$0.4 < z \leq 0.8$	92.0 ± 8.6	0.456 ± 0.037
$0.8 < z \leq 0.12$	121.5 ± 13.3	0.390 ± 0.104
$0.12 < z \leq 0.16$	161.2 ± 11.0	0.404 ± 0.056
$1.16 < z \leq 1.2$	194.2 ± 32.2	0.364 ± 0.106

Table 7.5: Binned measurements of $H(z)$ and $f\sigma_8(z)$ with equispaced redshifts points and its uncertainties. These are the gray points shown in Figs. 7.2 and 7.3.

7.7   Summary and Conclusion

In our work we implemented the conjoined $H(z) - f\sigma_8(z)$ method in order to test an entire family of ten dark energy models; we started with the simplest model, ΛCDM which is described by three parameters only, and we systematically increased the level of complexity of the model by adding extra parameters, being the non-flat CPL with dark energy perturbation the most complex model (with seven parameters).

For each model, we first found the best fit using MCMC analysis by combining the most recent cosmic chronometer and growth data available. Subsequently, we compared the dark energy models with five different statistical criteria, aiming at highlighting the potentiality and the weakness of each criterion.

As expected, we found that the evidence is the most accurate statistical test to compare different models as it takes into account the information of the entire likelihood of the parameters and it does not always penalize a model with extra parameters. The 3-FoM better characterizes the sensitivity of the parameters according to the data used. This criterion takes into account simultaneously the errors from both $f\sigma_8(z)$ and $H(z)$; in particular, we showed that the errors of the Hubble parameter increase with redshift and this has an important effect on the constraining power of the test. The FoM instead is limited only to $f\sigma_8(z)$, hence neglecting the information from $H(z)$, which might be crucial if the analysis is extended at high redshift. As a complementary test, we performed the same analysis in the same redshift range as in [98] and we found consistent results.

For the last two criteria, BIC and AIC$_c$, we showed that they always penalize the addition of extra parameters; in fact, if we consider the two extreme models, i.e. ΛCDM with only three

parameters and non-flat CPL with dark energy perturbations, which has seven parameters, we find that $\Delta \text{BIC} \sim 40$ manifesting a *very* strong evidence in favor of the ΛCDM model. Similarly, but less decisive is ΔAIC_c for which we find a value of ~ 10, which still favors strongly ΛCDM but more moderately than BIC.
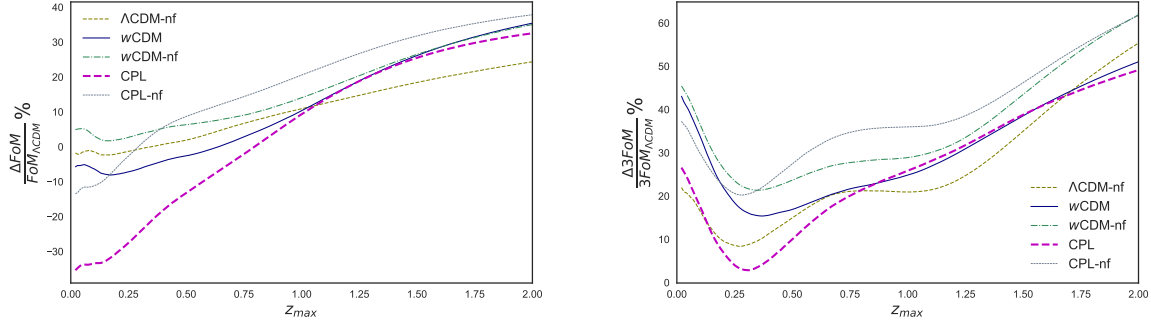

Figure 7.4: These figures show the percentage difference of FoM (left panel) or 3-FoM (right panel) between a model and ΛCDM. We only present the models without perturbations in the dark sector. Here, $\Delta \text{FoM} = \text{FoM}_{\Lambda\text{CDM}} - \text{FoM}_{\text{model}}$ and likewise for the 3-FoM.

To demonstrate the power of the 3-FoM, we compute the FoM and 3-FoM at different redshifts starting from $z = 0$ up to the z_{max}. These results are shown in Fig. 7.4 where we plotted the relative difference of the FoM (left panel) and the 3-FoM (right panel) for each model with respect to ΛCDM. It is interesting to notice that at low redshifts the FoM for wCDM, wCDM-nf, CPL, and CPL-nf is larger than ΛCDM, meaning that the former is better constrained than the latter. This effect is not manifested in the 3-FoM which is always larger for the ΛCDM model.

Appendix

7.8   Comparison with analytical solutions

Model	$\log(E)$	FoM	3FoM	BIC	AIC_c	H_{max}
ΛCDM-a	-22.07	0.192	0.027	51.34	34.01	202.05
wCDM-a	-23.28	0.125	0.014	59.32	36.39	204.80
wCDM-p-a	-23.23	0.126	0.014	67.22	38.79	204.66

Table 7.6: Results of the different methods for each analytic model. These are almost equal to their numerical versions.

The master equation for the matter density constrast in the small scale regime (i.e. $k \to \infty$) while not considering dark energy perturbations is obtained using only the matter component in the full set of differential equations Eqs. (7.1) - (7.4). After uncoupling them,

Model	$\Omega_{m,0}$	w_0	c_s^2	H_0	σ_8
ΛCDM-a	$0.286^{+0.033}_{-0.038}$	-	-	69.8 ± 2.4	0.780 ± 0.040
wCDM-a	$0.281^{+0.044}_{-0.039}$	$-1.10^{+0.36}_{-0.31}$	-	70.4 ± 4.4	$0.784^{+0.042}_{-0.11}$
wCDM-p-a	0.281 ± 0.044	$-1.10^{+0.36}_{-0.30}$	0.50 ± 0.29	$70.5^{+4.1}_{-4.7}$	$0.784^{+0.044}_{-0.11}$

Table 7.7: Parameter constraints derived from Nested Sampling to each analytical model described in the text.

the result is given by [69]

$$a^2\delta_m'' + (3 - \varepsilon(a))\, a\delta_m' - \frac{3}{2}\Omega_m(a)\delta_m(a) = 0 \ , \tag{7.35}$$

with $\varepsilon(a) = -\mathrm{d}\log H(a)/\mathrm{d}\log a$. As we are describing late time solutions we will always take the growing mode solution given by [69]

$$\delta(a) = a_2 F_1\left(\frac{w-1}{2w}, -\frac{1}{3w}, 1 - \frac{5}{6w}, 1 - \Omega_m^{-1}(a)\right) , \tag{7.36}$$

where we omitted the integration constant because it will cancel out when we evaluate $f(a)$. The result to ΛCDM is given by setting $w = -1$.

The study about perturbations in the dark energy sector is detailed in [119]. There, it is possible to find an exact solution when $c_s^2 = 0$ and an approximated one when $c_s^2 \neq 0$. Also, a joint solution is described in which case the density contrast is given by

$$\delta(a) = a_2 F_1\left(\frac{1}{4} - \frac{5}{12w} + B, \frac{1}{4} - \frac{5}{12w} - B, 1 - \frac{5}{6w}, 1 - \Omega_m^{-1}(a)\right) , \tag{7.37}$$

where B is used as B_{joint} in [119], which corresponds to:

$$B = \frac{1}{12w}\sqrt{(1 - 3w)^2 + 24\frac{1+w}{1 - 3w + \frac{2}{3}\frac{k^2}{H_0^2\Omega_{m,0}}c_s^2}} \ . \tag{7.38}$$

Utilizing these equations, it is possible to recover the exact solution when $c_s^2 = 0$.

Conclusion

In this thesis, we studied some applications of Bayesian statistics to large-scale structure-related observations, such as the growth rate of structure and the clustering of galaxies via the observed power spectrum. These quantities encapsulate the behavior of matter and galaxies through gravitation on the largest scales. While being in the era of precision cosmology, there is a special urge to analyze data as effectively as possible, to keep up with the technical improvements on observations. We must do this in order to draw trustworthy information, in hopes of achieving a satisfactory explanation to the mystery of the true nature of gravitation, the physical phenomenon that shapes the Universe.

To date, almost all cosmological data available points ΛCDM as the best candidate for explaining the Universe. It has, though, some problems like the fine-tuning problem of Λ or the coincidence problem of dark energy and matter densities, among others. On top of that, there are tensions between experiments on the values of its parameters, like H_0 [120] or σ_8 [94], or even complete cosmological probes against each other [121], which may or not be an issue of ΛCDM, as it could be a problem of underestimating error sources (for which Bayesian methods are tailored to the task of taking external nuisance phenomena into account) or unknown systematics. There are even new improvements in the statistical assumptions made at the beginning of the analyses [122].

The first of our projects explored an extension to the traditional Fisher Matrix forecasts for an LSST-like galaxy clustering survey. We went beyond the Fisher matrix framework by considering higher order derivatives on the Taylor expansion on the covariance matrix of the experiment. This extension, called the Derivative Approximation of LIkelihoods (DALI) is capable of capturing shapes other than Gaussians on the parameter confidence regions, unveiling parameter degeneracies that are ignored in a Fisher forecast. However, we faced numerical limitations when calculating accurate derivatives of the observed power spectrum which needs to be integrated, having to increase the accuracy of the Boltzmann solver and the step taken for the numerical derivatives respect to the parameters. Nevertheless, we concluded that DALI is better than the FM not only in its shape capturing aspect but also at estimating the size of the confidence regions (compared to MCMC samples), which can be even more important, as this translates into the error bars for the parameters. This behavior is more important for low redshift bins, where the Fisher regions are several times the size of the MCMC regions, even when considering physically-motivated flat priors.

Our second project took on another almost unexplored methodology: the Internal Robustness (iR) test, which exploits the question of whether a dataset might be split into (two)

subsets that obey different cosmological models, or different parameters of a given model. The method could also detect systematics or even new physics. The dataset used in the analysis was a compilation of growth rate measurements from different surveys, in the search of differences on the treatment of systematics of each experiment. Our work contemplated the exploration of all the possible subset combinations from the dataset (summing more than 2 million combinations), plus the comparison to mock catalogs, to check if the behavior of the dataset was in the expectation boundaries. We found that there were no outliers in the distribution of the Internal Robustness (iR-PDF), and the iR-PDF of the dataset is within the expected values compared to the mock catalogues, ensuring the iR of the dataset, increasing its fidelity, which could be important in future work that may involve the treatment of tensions on the σ_8 parameter.

Our third and final project involved comparing equation of state dark energy models from two datasets. The first dataset was the same growth rate compilation used in the second project whereas the second dataset was a compilation of direct Hubble measurements from *cosmic chronometers*. The objective of the work was to use five comparison methods, to see how they perform at ranking the models. We included the Bayesian evidence comparison as our cornerstone method, as it is the most accurate and well-cemented method. Thus, we used the evidence to gauge how well the other criteria performed. The second and third methods were the (corrected) Akaike Information Criteria (AIC_c) and the Bayesian Information Criterion (BIC), which rank models in terms of the minimum posterior probability and the number of data and parameters. The fourth method was a Figure of Merit (FoM) based upon recent findings stating that the conjoined $f\sigma_8$ versus H plot can be used to discriminate models, as the probe combination breaks any degeneracy, and the FoM consisted in the inverse of the area of the 68% confidence region of the plot. We noted that this method only contains the errors on the cosmic growth part, so we went further and developed an extension to this FoM that considers the inverse volume of a plot that also contains the errors of cosmic expansion, hence its name: the '3-FoM'. Our analysis resulted in the 3-FoM being the best comparison method among the four (not including the evidence, of course) and proved to be a more reliable tool that could be used for forecasts if one relaxes its definition in order to allow for approximations.

Bayesian methods in cosmology is a rather new field that is growing each day, closely tied to the new *big data* paradigm that is the theme for upcoming years of cosmological surveys. This implies that the frontiers of this field have a great amount of exploration and refining that needs to be achieved in order to match the level of sophistication of observational facilities. Especially, upcoming large-scale structure surveys will have the task of mapping billions of galaxies, and the statistical methodology has the responsibility of giving that data the best use, in hopes to deliver useful information that might lead to a new paradigm for the Universe.

Bibliography

[1] Roberto Trotta. Bayes in the sky: Bayesian inference and model selection in cosmology. *Contemp. Phys.*, 49:71–104, 2008.

[2] Xingang Chen, Cora Dvorkin, Zhiqi Huang, Mohammad Hossein Namjoo, and Licia Verde. The Future of Primordial Features with Large-Scale Structure Surveys. *JCAP*, 1611(11):014, 2016.

[3] N. Suzuki et al. The Hubble Space Telescope Cluster Supernova Survey: V. Improving the Dark Energy Constraints Above z>1 and Building an Early-Type-Hosted Supernova Sample. *Astrophys. J.*, 746:85, 2012.

[4] Albert Einstein. On the General Theory of Relativity. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1915:778–786, 1915. [Addendum: Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)1915,799(1915)].

[5] Slava G. Turyshev. Experimental Tests of General Relativity. *Ann. Rev. Nucl. Part. Sci.*, 58:207–248, 2008.

[6] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016.

[7] Edwin Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proc. Nat. Acad. Sci.*, 15:168–173, 1929.

[8] Alan Heavens, Yabebal Fantaye, Elena Sellentin, Hans Eggers, Zafiirah Hosenie, Steve Kroon, and Arrykrishna Mootoovaloo. No evidence for extensions to the standard cosmological model. *Phys. Rev. Lett.*, 119(10):101301, 2017.

[9] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. 2018.

[10] B. Ryden. *Introduction to cosmology*. Cambridge University Press, 2016.

[11] Luca Amendola and Shinji Tsujikawa. *Dark Energy*. Cambridge University Press, 2015.

[12] Antonino Del Popolo and Morgan Le Delliou. Small scale problems of the ΛCDM model: a short review. *Galaxies*, 5(1):17, 2017.

[13] Adam G. Riess et al. Observational evidence from supernovae for an accelerating

universe and a cosmological constant. *Astron. J.*, 116:1009–1038, 1998.

[14] S. Perlmutter et al. Measurements of Omega and Lambda from 42 high redshift super-novae. *Astrophys. J.*, 517:565–586, 1999.

[15] Steven Weinberg. The Cosmological constant problems. In *Sources and detection of dark matter and dark energy in the universe. Proceedings, 4th International Symposium, DM 2000, Marina del Rey, USA, February 23-25, 2000*, pages 18–26, 2000.

[16] Donald G. York et al. The Sloan Digital Sky Survey: Technical Summary. *Astron. J.*, 120:1579–1587, 2000.

[17] Alexandra Abate et al. Large Synoptic Survey Telescope: Dark Energy Science Collaboration. 2012.

[18] Luca Amendola et al. Cosmology and fundamental physics with the Euclid satellite. *Living Rev. Rel.*, 16:6, 2013.

[19] Domenico Sapone. Dark Energy in Practice. *Int. J. Mod. Phys.*, A25:5253–5331, 2010.

[20] A. Friedman. On the Curvature of space. *Z. Phys.*, 10:377–386, 1922. [Gen. Rel. Grav.31,1991(1999)].

[21] G. Lemaître. Le mouvement varié d'un solide d'apres la théorie de la relativité. *Annales Soc. Sci. Bruxelles A*, 47(4):103–109, 1927. [Publ. Lab. Astron. Geodesie Univ. Louvain4,239(1927)].

[22] H. P. Robertson. On the Foundations of Relativistic Cosmology. *Proceedings of the National Academy of Science*, 15:822–829, November 1929.

[23] A. G. Walker. On Milne's Theory of World-Structure. *Proceedings of the London Mathematical Society, (Series 2) volume 42, p. 90-127*, 42:90–127, 1937.

[24] Albert Einstein. The Field Equations of Gravitation. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1915:844–847, 1915.

[25] Adam G. Riess et al. A 2.4% Determination of the Local Value of the Hubble Constant. *Astrophys. J.*, 826(1):56, 2016.

[26] Michel Chevallier and David Polarski. Accelerating universes with scaling dark matter. *Int. J. Mod. Phys.*, D10:213–224, 2001.

[27] Eric V. Linder. Exploring the expansion history of the universe. *Phys. Rev. Lett.*, 90:091301, 2003.

[28] Ivaylo Zlatev, Li-Min Wang, and Paul J. Steinhardt. Quintessence, cosmic coincidence, and the cosmological constant. *Phys. Rev. Lett.*, 82:896–899, 1999.

[29] David W. Hogg. Distance measures in cosmology. 1999.

[30] Pierre Astier. The expansion of the universe observed with supernovae. *Rept. Prog. Phys.*, 75:116901, 2012.

[31] Julien Guy et al. SALT2: Using distant supernovae to improve the use of Type Ia supernovae as distance indicators. *Astron. Astrophys.*, 466:11–21, 2007.

[32] Hee-Jong Seo and Daniel J. Eisenstein. Probing dark energy with baryonic acoustic oscillations from future large galaxy redshift surveys. *Astrophys. J.*, 598:720–740, 2003.

[33] Pauline Zarrouk et al. The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: measurement of the growth rate of structure from the anisotropic correlation function between redshift 0.8 and 2.2. *Mon. Not. Roy. Astron. Soc.*, 477(2):1639–1663, 2018.

[34] Julien Lesgourgues. The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview. 2011.

[35] Diego Blas, Julien Lesgourgues, and Thomas Tram. The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes. *JCAP*, 1107:034, 2011.

[36] Antony Lewis, Anthony Challinor, and Anthony Lasenby. Efficient computation of CMB anisotropies in closed FRW models. *Astrophys. J.*, 538:473–476, 2000.

[37] Antony Lewis and Sarah Bridle. Cosmological parameters from CMB and other data: A Monte Carlo approach. *Phys. Rev.*, D66:103511, 2002.

[38] Chung-Pei Ma and Edmund Bertschinger. Cosmological perturbation theory in the synchronous and conformal Newtonian gauges. *Astrophys. J.*, 455:7–25, 1995.

[39] Elisabetta Majerotto, Domenico Sapone, and Björn Malte Schäfer. Combined constraints on deviations of dark energy from an ideal fluid from Euclid and Planck. *Mon. Not. Roy. Astron. Soc.*, 456(1):109–118, 2016.

[40] Dandan Wang et al. The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: Anisotropic Baryon Acoustic Oscillations measurements in Fourier-space with optimal redshift weights. *Mon. Not. Roy. Astron. Soc.*, 477(2):1528–1535, 2018.

[41] Licia Verde. Statistical methods in cosmology. *Lect. Notes Phys.*, 800:147–177, 2010.

[42] Steen Hannestad and Thomas Tram. Optimal prior for Bayesian inference in a constrained parameter space. 2017.

[43] S. I. Bitioukov and N. V. Krasnikov. The relation between frequentist confidence intervals and Bayesian credible intervals. 2012.

[44] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092, 1953.

[45] W. K. Hastings. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, 57:97–109, 1970.

[46] David W. Hogg and Daniel Foreman-Mackey. Data analysis recipes: Using Markov Chain Monte Carlo. *Astrophys. J. Suppl.*, 236(1):11, 2018.

[47] G. O. Roberts, A. Gelman, and W. R. Gilks. Weak convergence and optimal scaling of random walk metropolis algorithms. *Ann. Appl. Probab.*, 7(1):110–120, 02 1997.

[48] Andrew Gelman and Donald B. Rubin. Inference from Iterative Simulation Using Multiple Sequences. *Statist. Sci.*, 7:457–472, 1992.

[49] Elena Sellentin, Miguel Quartin, and Luca Amendola. Breaking the spell of Gaussianity: forecasting with higher order Fisher matrices. *Mon. Not. Roy. Astron. Soc.*, 441(2):1831–1840, 2014.

[50] John Skilling. Nested sampling for general bayesian computation. *Bayesian Anal.*, 1(4):833–859, 12 2006.

[51] R. Shaw, Michael Bridges, and M. P. Hobson. Clustered nested sampling: Efficient Bayesian inference for cosmology. *Mon. Not. Roy. Astron. Soc.*, 378:1365–1370, 2007.

[52] F. Feroz, M. P. Hobson, and M. Bridges. MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Mon. Not. Roy. Astron. Soc.*, 398:1601–1614, 2009.

[53] Max Tegmark. Measuring cosmological parameters with galaxy surveys. *Phys. Rev. Lett.*, 79:3806–3809, 1997.

[54] Elena Sellentin. A fast, always positive definite and normalizable approximation of non-Gaussian likelihoods. *Mon. Not. Roy. Astron. Soc.*, 453(1):893–898, 2015.

[55] Elena Sellentin and Björn Malte Schäfer. Non-Gaussian forecasts of weak lensing with and without priors. *Mon. Not. Roy. Astron. Soc.*, 456(2):1645–1653, 2016.

[56] Luca Amendola, Claudia Quercellini, and Emanuele Giallongo. Constraints on perfect fluid and scalar field dark energy models from future redshift surveys. *Mon. Not. Roy. Astron. Soc.*, 357:429–439, 2005.

[57] Alberto Bailoni, Alessio Spurio Mancini, and Luca Amendola. Improving Fisher matrix forecasts for galaxy surveys: window function, bin cross-correlation, and bin redshift uncertainty. *Mon. Not. Roy. Astron. Soc.*, 470(1):688–705, 2017.

[58] C. Alcock and B. Paczynski. An evolution free test for non-zero cosmological constant. *Nature*, 281:358–359, 1979.

[59] N. Kaiser. Clustering in real space and in redshift space. *Mon. Not. Roy. Astron. Soc.*, 227:1–27, 1987.

[60] Bryan Sagredo, Savvas Nesseris, and Domenico Sapone. The Internal Robustness of Growth Rate data. 2018.

[61] Luca Amendola, Valerio Marra, and Miguel Quartin. Internal Robustness: systematic search for systematic bias in SN Ia data. *Mon. Not. Roy. Astron. Soc.*, 430:1867–1879, 2013.

[62] Shinji Tsujikawa. Modified gravity models of dark energy. *Lect. Notes Phys.*, 800:99–145, 2010.

[63] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 594:A13, 2016.

[64] Martin Kunz and Domenico Sapone. Crossing the Phantom Divide. *Phys. Rev.*, D74:123503, 2006.

[65] S. Nesseris and Leandros Perivolaropoulos. Crossing the Phantom Divide: Theoretical Implications and Observational Status. *JCAP*, 0701:018, 2007.

[66] Paul A. Abell et al. LSST Science Book, Version 2.0. 2009.

[67] V. Silveira and I. Waga. Decaying Lambda cosmologies and power spectrum. *Phys. Rev.*, D50:4890–4894, 1994.

[68] Will J. Percival. Cosmological structure formation in a homogeneous dark energy background. *Astron. Astrophys.*, 443:819, 2005.

[69] Alicia Bueno Belloso, Juan Garcia-Bellido, and Domenico Sapone. A parametrization of the growth index of matter perturbations in various Dark Energy models and observational prospects using a Euclid-like survey. *JCAP*, 1110:010, 2011.

[70] Savvas Nesseris, George Pantazis, and Leandros Perivolaropoulos. Tension and constraints on modified gravity parametrizations of $G_{\mathrm{eff}}(z)$ from growth rate and Planck data. *Phys. Rev.*, D96(2):023542, 2017.

[71] Spyros Basilakos and Savvas Nesseris. Testing Einstein?s gravity and dark energy with growth of matter perturbations: Indications for new physics? *Phys. Rev.*, D94(12):123525, 2016.

[72] Laura Taddei, Matteo Martinelli, and Luca Amendola. Model-independent constraints on modified gravity from current data and from the Euclid and SKA future surveys. *JCAP*, 1612(12):032, 2016.

[73] Chris Blake et al. The WiggleZ Dark Energy Survey: Joint measurements of the expansion and growth history at z < 1. *Mon. Not. Roy. Astron. Soc.*, 425:405–414, 2012.

[74] Gong-Bo Zhao et al. The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: a tomographic measurement of cosmic

structure growth and expansion rate based on optimal redshift weights. 2018.

[75] Dragan Huterer, Daniel Shafer, Daniel Scolnic, and Fabian Schmidt. Testing LCDM at the lowest redshifts with SN Ia and galaxy velocities. 2016.

[76] Stephen J. Turnbull, Michael J. Hudson, Hume A. Feldman, Malcolm Hicken, Robert P. Kirshner, and Richard Watkins. Cosmic flows in the nearby universe from Type Ia Supernovae. *Mon. Not. Roy. Astron. Soc.*, 420:447–454, 2012.

[77] Michael J. Hudson and Stephen J. Turnbull. The growth rate of cosmic structure from peculiar velocities at low and high redshifts. *Astrophys. J.*, 751:L30, 2013.

[78] Marc Davis, Adi Nusser, Karen Masters, Christopher Springob, John P. Huchra, and Gerard Lemson. Local Gravity versus Local Velocity: Solutions for β and nonlinear bias. *Mon. Not. Roy. Astron. Soc.*, 413:2906, 2011.

[79] Martin Feix, Adi Nusser, and Enzo Branchini. Growth Rate of Cosmological Perturbations at z 0.1 from a New Observational Test. *Phys. Rev. Lett.*, 115(1):011301, 2015.

[80] Cullan Howlett, Ashley Ross, Lado Samushia, Will Percival, and Marc Manera. The clustering of the SDSS main galaxy sample. Mock galaxy catalogues and a measurement of the growth of structure from redshift space distortions at $z = 0.15$. *Mon. Not. Roy. Astron. Soc.*, 449(1):848–866, 2015.

[81] Yong-Seon Song and Will J. Percival. Reconstructing the history of structure formation using Redshift Distortions. *JCAP*, 0910:004, 2009.

[82] Chris Blake et al. Galaxy And Mass Assembly (GAMA): improved cosmic growth measurements using multiple tracers of large-scale structure. *Mon. Not. Roy. Astron. Soc.*, 436:3089, 2013.

[83] Lado Samushia, Will J. Percival, and Alvise Raccanelli. Interpreting large-scale redshift-space distortion measurements. *Mon. Not. Roy. Astron. Soc.*, 420:2102–2119, 2012.

[84] Ariel G. Sanchez et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological implications of the full shape of the clustering wedges in the data release 10 and 11 galaxy samples. *Mon. Not. Roy. Astron. Soc.*, 440(3):2692–2713, 2014.

[85] Chia-Hsun Chuang et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: single-probe measurements from CMASS anisotropic galaxy clustering. *Mon. Not. Roy. Astron. Soc.*, 461(4):3781–3793, 2016.

[86] A. Pezzotta et al. The VIMOS Public Extragalactic Redshift Survey (VIPERS): The growth of structures at $0.5 < z < 1.2$ from redshift-space distortions in the clustering of the PDR-2 final sample. 2016.

[87] Teppei Okumura et al. The Subaru FMOS galaxy redshift survey (FastSound). IV.

New constraint on gravity theory from redshift space distortions at $z \sim 1.4$. *Publ. Astron. Soc. Jap.*, 68(3, id. 38):24, 2016.

[88] Stefano Camera, Matteo Martinelli, and Daniele Bertacca. Easing Tensions with Quartessence. 2017.

[89] F. Kohlinger et al. KiDS-450: The tomographic weak lensing power spectrum and constraints on cosmological parameters. *Mon. Not. Roy. Astron. Soc.*, 471(4):4412–4435, 2017.

[90] T. M. C. Abbott et al. Dark Energy Survey Year 1 Results: Cosmological Constraints from Galaxy Clustering and Weak Lensing. 2017.

[91] Adria Gomez-Valent and Joan Sola. Relaxing the σ_8-tension through running vacuum in the Universe. *EPL*, 120(3):39001, 2017.

[92] Adria Gomez-Valent and Joan Sola. Density perturbations for running vacuum: a successful approach to structure formation and to the σ_8-tension. *Mon. Not. Roy. Astron. Soc.*, 478(1):126–145, 2018.

[93] Bruno J. Barros, Luca Amendola, Tiago Barreiro, and Nelson J. Nunes. Coupled quintessence with a ΛCDM background: removing the σ_8 tension. 2018.

[94] Gaetano Lambiase, Subhendra Mohanty, Ashish Narang, and Priyank Parashari. Testing Dark energy models in the light of σ_8 tension. 2018.

[95] Radouane Gannouji and David Polarski. On the consistency of the expansion with the perturbations. 2018.

[96] Bryan Sagredo, Javier Silva Lafaurie, and Domenico Sapone. Comparing Dark Energy models with Hubble versus Growth Rate data. 2018.

[97] Eric V. Linder. Cosmic Growth and Expansion Conjoined. *Astropart. Phys.*, 86:41–45, 2017.

[98] Spyros Basilakos and Savvas Nesseris. Conjoined constraints on modified gravity from the expansion history and cosmic growth. *Phys. Rev.*, D96(6):063517, 2017.

[99] Roberto Trotta. Applications of Bayesian model selection to cosmological parameters. *Mon. Not. Roy. Astron. Soc.*, 378:72–82, 2007.

[100] Gideon Schwarz. Estimating the Dimension of a Model. *Annals Statist.*, 6:461–464, 1978.

[101] Nariaki Sugiura. Further analysts of the data by akaike' s information criterion and the finite corrections. *Communications in Statistics - Theory and Methods*, 7(1):13–26, 1978.

[102] M. Betoule et al. Improved cosmological constraints from a joint analysis of the SDSS-II

and SNLS supernova samples. *Astron. Astrophys.*, 568:A22, 2014.

[103] Camille Bonvin, Ruth Durrer, and Martin Kunz. The dipole of the luminosity distance: a direct measure of H(z). *Phys. Rev. Lett.*, 96:191302, 2006.

[104] Martin Kunz and Domenico Sapone. Dark Energy versus Modified Gravity. *Phys. Rev. Lett.*, 98:121301, 2007.

[105] Michele Moresco and Federico Marulli. Cosmological constraints from a joint analysis of cosmic growth and expansion. *Mon. Not. Roy. Astron. Soc.*, 471(1):L82–L86, 2017.

[106] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, Dec 1974.

[107] Valerio Marra and Domenico Sapone. Null tests of the standard model using the linear model formalism. *Phys. Rev.*, D97(8):083510, 2018.

[108] Michele Moresco, Lucia Pozzetti, Andrea Cimatti, Raul Jimenez, Claudia Maraston, Licia Verde, Daniel Thomas, Annalisa Citro, Rita Tojeiro, and David Wilkinson. A 6% measurement of the Hubble parameter at $z \sim 0.45$: direct evidence of the epoch of cosmic re-acceleration. *JCAP*, 1605(05):014, 2016.

[109] Cong Zhang, Han Zhang, Shuo Yuan, Tong-Jie Zhang, and Yan-Chun Sun. Four new observational $H(z)$ data from luminous red galaxies in the Sloan Digital Sky Survey data release seven. *Res. Astron. Astrophys.*, 14(10):1221–1233, 2014.

[110] Joan Simon, Licia Verde, and Raul Jimenez. Constraints on the redshift dependence of the dark energy potential. *Phys. Rev.*, D71:123001, 2005.

[111] Daniel Stern, Raul Jimenez, Licia Verde, Marc Kamionkowski, and S. Adam Stanford. Cosmic Chronometers: Constraining the Equation of State of Dark Energy. I: H(z) Measurements. *JCAP*, 1002:008, 2010.

[112] 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.

[113] Michele Moresco. Raising the bar: new constraints on the Hubble parameter with cosmic chronometers at z = 2. *Mon. Not. Roy. Astron. Soc.*, 450(1):L16–L20, 2015.

[114] A. L. Ratsimbazafy, S. I. Loubser, S. M. Crawford, C. M. Cress, B. A. Bassett, R. C. Nichol, and P. Väisänen. Age-dating Luminous Red Galaxies observed with the Southern African Large Telescope. *Mon. Not. Roy. Astron. Soc.*, 467(3):3239–3254, 2017.

[115] Kyle Barbary. Nestle. `https://github.com/kbarbary/nestle`, 2015.

[116] Andrew R. Liddle. How many cosmological parameters? *Mon. Not. Roy. Astron. Soc.*, 351:L49–L53, 2004.

[117] Savvas Nesseris and Juan Garcia-Bellido. Is the Jeffreys' scale a reliable tool for

Bayesian model comparison in cosmology? *JCAP*, 1308:036, 2013.

[118] Judit Pérez-Romero and Savvas Nesseris. Cosmological constraints and comparison of viable $f(R)$ models. *Phys. Rev.*, D97(2):023525, 2018.

[119] Savvas Nesseris and Domenico Sapone. Accuracy of the growth index in the presence of dark energy perturbations. *Phys. Rev.*, D92(2):023013, 2015.

[120] David Camarena and Valerio Marra. The impact of the cosmic variance on H_0 on cosmological analyses. *Phys. Rev.*, D98:023537, 2018.

[121] Marco Raveri. Are cosmological data sets consistent with each other within the Λ cold dark matter model? *Phys. Rev.*, D93(4):043522, 2016.

[122] Elena Sellentin and Alan F. Heavens. On the insufficiency of arbitrarily precise covariance matrices: non-Gaussian weak lensing likelihoods. *Mon. Not. Roy. Astron. Soc.*, 473(2):2355–2363, 2018.