# Playing dice with the universe: Bayesian statistical analyses of cosmological models and new observables

Cañas Herrera, G.

# Chapter 1

# Introduction

Looking up at the night sky. A very small body movement that costs an insignificantly small number of calories. However, the consequences of bending the neck up 45 degrees are extraordinary. First, it triggers our innate human curiosity about the world we live in. Second, this curiosity activates our critical thinking by our wish to learn something new. All through history, physicists have tried to explain how nature works using this method of critical thinking, summarized in the scientific method. In the 20th century, a tiny neck movement translated into a series of remarkable discoveries in the field of Theoretical Physics, culminating in the work of A. Einstein about the Theory of General Relativity (Einstein, 1916). This milestone, together with discoveries in the atomic world, brought about the era of *Modern Physics*, establishing a revolution in our understanding of the universe.

During that century, humanity discovered that the universe is expanding (encoded in the Hubble parameter) (Hubble, 1929). Years later, the *Cosmic Microwave Background* radiation was predicted and detected (Penzias & Wilson, 1965). Furthermore, there appeared strong evidence that galaxies rotate faster in their outer regions pointing out the existence of an unknown type of matter called *Dark Matter* (Rubin, 1983), and ultimately, it was found out that the universe is currently undergoing an accelerated expansion thanks to the study of Supernovae type Ia observations (Schmidt et al., 1998).

These discoveries started the so-called *Modern Cosmology* era. Cosmologists aim to narrate the biography of the universe to understand its origin, structure and evolution, its properties, and, eventually, its final fate. Cosmological observations have increased our understanding of the universe's evolution from the time of the creation of the first nuclei up to the present. Yet, at earlier (and future) times the lack of cosmological information makes speculation dominate.

The ultimate goal of cosmologists is to find a cosmological model able to explain the current observational data, and so far, it seems that they have accomplished it. The *Standard Cosmological model*, also known as the ΛCDM model, establishes that our universe is mainly composed of two well-measured yet unknown components: a type of matter that is known to only interact through gravitation, Cold Dark Matter (CDM), and a substance responsible of the current accelerated expansion of the universe that can be modelled by a cosmological constant ($\Lambda$). The ordinary mat-

ter from which everything we observe is made of commonly called Baryonic Matter[1] (BM), only accounts for less than 5% of the total composition of the universe. Cosmologists have been able to test the robustness of the ΛCDM model mostly thanks to the high-accuracy measurements of the temperature anisotropies of the Cosmic Microwave Background (CMB) obtained by successful experiments such as the NASA WMAP mission (WMAP, 2003) and the ESA Planck mission (Planck Collaboration et al., 2014a).

The Standard Cosmological model relies on the assumption that the underlying distribution of the primordial density perturbations that seeded the early universe was Gaussian. These density fluctuations in the early universe are the origin of the current Large Scale Structure (LSS) of our universe that is observed today. The simplest model of inflation (an exponential expansion phase in the primordial universe) is compatible with this assumption, apart from explaining why the cosmological observations indicate that the universe is causally connected and flat.

Still, the ΛCDM model, though successful, fails to answer hot-burning questions in the field such as "what is the nature of Λ?" or "why is there a tension between measurements of the Hubble parameter using different observational probes?". For this reason, theoretical cosmologists focus on developing modifications of the Standard Cosmological model, and test them against astrophysical data to check whether alternative scenarios can provide a better explanation of the observations. Furthermore, cosmologists also work on constructing theoretical predictions of new observables that can be used to test a given model beyond ΛCDM. The key discipline used so far to compare theory with cosmological data is *Statistics*. In the last 20 years, not only the amount of cosmological and astrophysical data has increased significantly, but also our statistical analysis tools have evolved dramatically (Trotta, 2017).

In the coming decades, a substantial change in the type of cosmological observations used for statistical analysis will take place. Although the era of precision Cosmology based on the observations of the CMB has extensively advanced our understanding of the universe, it is close to an end[2]. In the next decade, the most precise constraints on the parameters of the ΛCDM model (or extensions of this model) will come from the information embedded in the large scale structure (LSS) of the universe, with several cosmological experiments and missions planned. The ESA mission *Euclid* (Laureijs et al., 2011), whose launch is expected in 2023, will precisely map the universe's structure by studying two observables: the Weak Lensing effect (WL), which studies the apparent change of the shapes of galaxies) and Galaxy Clustering (GC), which focuses on how well the distribution of galaxies in the universe can trace the underlying matter distribution.

Even though the goal of the *Euclid* mission is to shine some "light" on the origin and nature of Λ and Dark Matter, its data can also be used to explore the origin of

[1]Cosmologists usually misuse this term, as when they talk about "baryonic" matter, they also refer to leptons.

[2]The CMB still offers an incredible source of information encoded in other observables such as spectral distortions or the polarization signal. There are still efforts in proposing, designing, pursuing and realizing new experiments based on observations of the CMB. See for instance (Suzuki et al., 2018).

the primordial density fluctuations that seeded the very early universe and evolved up to the structure visible today. Therefore, the next generation of LSS surveys will have the power to indirectly constrain different inflationary models by studying the underlying spatial distribution of the cosmic web.

Furthermore, LSS surveys, together with observations of Gravitational-Wave (GW) events (Black Hole or Neutron Star mergers), lay the foundation for multimessenger Cosmology: cross-correlating the information contained in GW merger events and the LSS information of our universe opens a new window for research in Cosmology, in which the universe is observed not only through the electromagnetic spectrum but also through gravitational signals and the combination between these two.

In this self-contained chapter, we review the most important concepts of the Standard Cosmological Model following the reviews by (Baumann, 2009), (Carroll & Ostlie, 2014), (Dodelson, 2003) and (Mukhanov, 2005).

## 1.1 Evolution of a homogeneous and isotropic universe

Cosmological observations at sufficiently large scales ($> 100$ Mpc) support the idea that the universe is *isotropic* on large scales. If we also assume that we are not a special observer in the universe, we infer *homogeneity* of space. These two characteristics are the foundations of the *Cosmological Principle*. In the language of General Relativity, such a universe is described by the Friedmann-Lemaitre-Robertson-Walker (FLRW) metric (Friedmann, 1924; Lemaître, 1931; Robertson, 1935; Walker, 1937):

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a^2(t)\gamma_{ij}dx^i dx^j, \tag{1.1}$$

where:

- $\mu, \nu, i, j$ take the values $\{\mu, \nu\} = \{0, 1, 2, 3\}$ and $\{i, j\} = \{1, 2, 3\}$.
- $a(t)$ is the scale factor, which accounts for the relative size of space-like hypersurfaces at a given time.
- $x$ is the comoving coordinate, which defines a fixed set of points on a coordinate grid that grows with the expansion of the universe given by scale factor $a(t)$.
- $t$ is the cosmic time, which is the time measured by comoving observers (observers who move with the space expansion determined by the rate of change of the scale factor $a$).
- $\gamma_{ij}$ is the tensor corresponding to the spatial part of the metric, which in spherical coordinates $(r, \theta, \phi)$ and together with the differentials $dx^i dx^j$ takes the form:

  $$\gamma_{ij}dx^i dx^j = \frac{dr^2}{1 - \kappa r^2} + r^2 d\phi^2 + r^2 \sin\theta^2 d\theta^2, \tag{1.2}$$

  where $\kappa$ is the intrinsic curvature of 3-surfaces, and represents a *flat* ($\kappa = 0$), *positively curved* ($\kappa = +1$), or *negatively curved* ($\kappa = -1$) spatial slices.

For convenience, we define the *conformal time* $\tau$:

$$d\tau = \frac{dt}{a} \rightarrow ds^2 = a^2(\tau)\left[-d\tau^2 + \gamma_{ij}dx^i dx^j\right]. \tag{1.3}$$

Light travels along null-geodesics ($ds^2 = 0$), and in the transformed FLRW metric (1.3), its propagation is the same as in Minkowski space when $\kappa = 0$. The scale factor $a(t)$ encodes the time evolution of the universe. Usually, it is normalised so that today $t_0$ it takes the value $a(t_0) = 1$. Due to the expansion of the universe, photons travelling from their place of emission towards us are redshifted. This redshift $z$ is due to the stretching of electromagnetic waves along their path to us and it is quantified as

$$z + 1 = \frac{1}{a(t)}. \tag{1.4}$$

The FLRW metric is the solution to *Einstein Equations*:

$$G_{\mu\nu} = \frac{8\pi G}{c^2}T_{\mu\nu}, \tag{1.5}$$

where $G_{\mu\nu}$ is the Einstein tensor and $T_{\mu\nu}$ is the energy-momentum tensor of the matter components. If we choose to describe the matter present in the universe as a perfect fluid in the rest frame defined by its energy density $\rho$ and pressure $p$, the tensor $T_{\mu\nu}$ takes the form:

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

where $g_{\mu\nu}$ is the metric of equation (1.1) and $u_\mu$ is the four velocity of a perfect fluid. Solving equation (1.5) using natural units $c = \hbar = 1$, and defining the reduced Planck mass as $M_P = \sqrt{\frac{1}{8\pi G}}$, we obtain the well-known First and Second[3] *Friedmann* equations respectively (Baumann, 2009),

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{\rho}{3M_P^2} - \frac{\kappa}{a^2}, \tag{1.7}$$

$$\dot{H} + H^2 \equiv \frac{\ddot{a}}{a} = -\frac{1}{6M_P^2}(\rho + 3p). \tag{1.8}$$

where the derivative with respect to the cosmic time $t$ is denoted with the dot ($\dot{} = d/dt$) and $H$ is the *Hubble* parameter, a crucial function in Cosmology defined as:

$$H(t) = \frac{\dot{a}(t)}{a(t)}. \tag{1.9}$$

Combining both equations (1.7) and (1.8) (or alternatively, calculating directly $\nabla_\mu T^{\mu\nu} = 0$) we obtain the continuity equation

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

[3]The Second Friedmann equation is usually called the acceleration equation.

The equation of state of a perfect fluid is often given in terms of the dimensionless state parameter, $w$, as $p = w\rho$. Solving equation (1.10) for each species $i$ as a function of the scale factor $a(t)$,

$$\rho_i = \rho_{i,0}\,(a)^{-3(1+w_i)} \rightarrow \begin{cases} \rho_m \propto a^{-3} & (w=0) \quad \text{matter (Cold Dark Matter and baryons)} \\ \rho_r \propto a^{-4} & \left(w=\frac{1}{3}\right) \quad \text{radiation} \end{cases}, \tag{1.11}$$

where the sub-index 0 indicates a reference value, which we have taken at the present time. There is an extra component that must be taken into account to explain current cosmological observations: the cosmological constant $\Lambda$. This constant was first introduced in the Einstein Equations (1.5) as a term $\Lambda g_{\mu\nu}$ to obtain static solutions of the Friedmann equations, and was later reintroduced when the High-Z Supernova Search Team discovered that the expansion of the universe is accelerating thanks to the study of supernovae type Ia observations (Riess et al., 1998). Given equations (1.8) and (1.11), we observe that a late time acceleration of the expansion of the universe occurs when the universe starts to be dominated by an energy component with an equation of state $w \leq -1/3$; that is, a substance that violates the strong energy condition that is given by $w \geq -1/3$.

Why is this acceleration taking place? In quantum field theory there exists the prediction of an accelerated expansion of the universe due to the non-zero quantum mechanical vacuum energy. This vacuum energy behaves exactly like a cosmological constant with an equation of state $w_\Lambda = -1$, which results in a energy density

$$\rho_\Lambda(t) = \rho_\Lambda(t_0). \tag{1.12}$$

This constant energy density is described by the energy-momentum tensor

$$T^\Lambda_{\mu\nu} = -M_P^2 \Lambda g_{\mu\nu}. \tag{1.13}$$

Unfortunately, the theoretical predictions of the estimated value of the vacuum energy density do not agree with the cosmological observations. Still, current observations can be explained by a cosmological constant $\Lambda$ behaving as a fluid-like energy component of the universe responsible of the accelerated late-time expansion. This approach receives the name of dynamical Dark Energy (DE), which characterizes the fluid by an equation of state $w_{\rm DE} = -1$. To infer information from the observational data, a phenomenological parametrization is often used, denominated CPL (Chevallier & Polarski, 2001)(Linder, 2003) defined as,

$$w_{\rm DE}(z) = w_0 + w_a(1 - a(z)). \tag{1.14}$$

We aim to provide DE or modified gravity (MG) models that are able to explain the current observational data to obtain some insight into the nature and origin of the current accelerated expansion. To find a balance between observations and plausible theories, it is interesting to study the Large Scale Structure of the universe from the predictions of the $\Lambda$CDM model in terms of the phenomenological functions[4]

[4]We are using the subscript "MG" to differentiate the symbol $\mu$ from other quantities that will be introduced later in this chapter.

$\Sigma_{MG}$ and $\mu_{MG}$, which parametrize modifications of the perturbed Einstein's equations relating the matter density contrast to the lensing of light and the Newtonian potential (Frusciante & Perenon, 2020). The function $\Sigma_{\rm MG}$ will be later used in chapter 6.

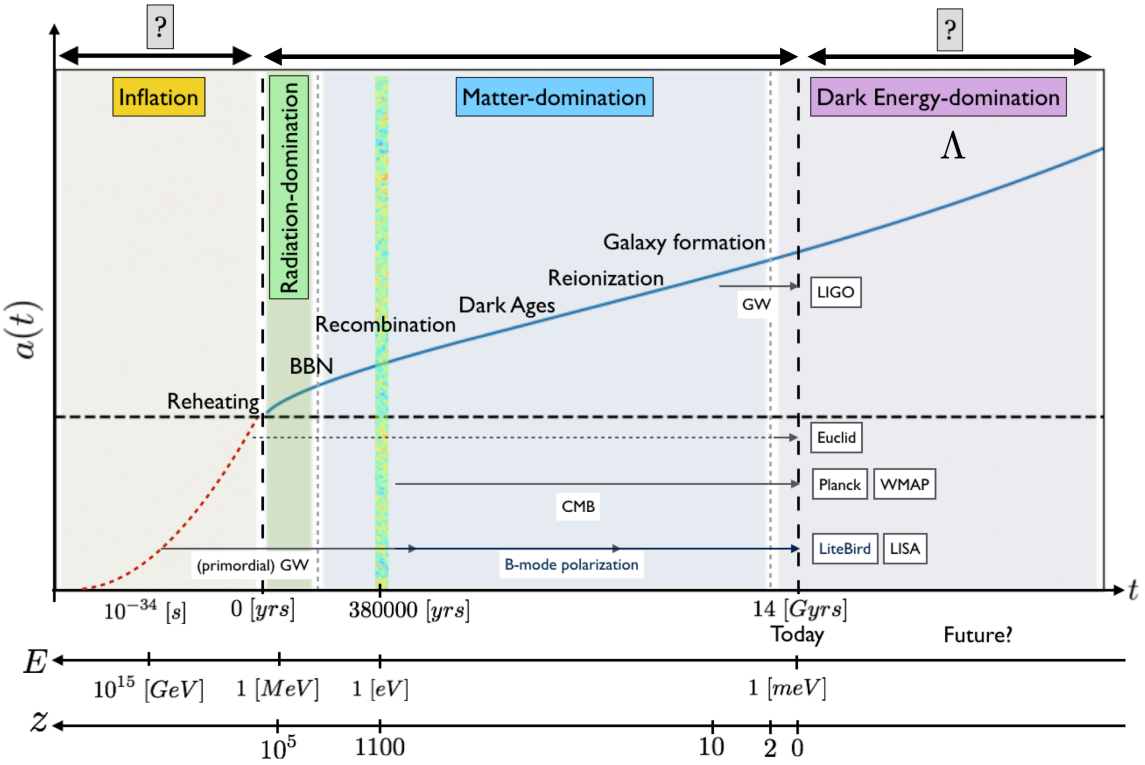


Figure 1.1: Evolution of the scale factor $a(t)$ as a function of the cosmic time $t$, redshift $z$ and energy scale $E$. The main key events in the history of the universe are pointed out during the universe's timeline. We highlight speculative epochs of the universe: at earlier times, inflation and at later future times, Dark Energy-domination, where the current accelerated expansion of the universe is modelled in term of the cosmological constant $\Lambda$. Numerical integration of $a(t)$ has been done with (*Calculadora Cosmologica: code that solves Friedman Equations numerically given values for components of universe*, 2015). The time evolution of the scale factor during radiation-domination scales as $a(t) \propto t^{1/2}$ and during matter-domination as $a(t) \propto t^{2/3}$. Some cosmological space missions as well as experiments placed on Earth are included only for representative purposes. All acronyms can be found in the Glossary.

Given (1.11), we can determine the evolution of the scale factor, $a(t)$, as a function of time in the presence of these 3 different species. Defining the critical density $\rho_c \equiv 3M_p^2 H_0$ and the density parameter[5] $\Omega_{i,0} = \rho_i/\rho_c$ ($H_0$ is the Hubble parameter

[5]We are denoting the density parameters at present time with the sub-index 0. However, this sub-index is most of the times dropped, and we will follow this notation in the rest of the thesis.

evaluated today), we can re-write Friedmann equation (1.7) as,

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = H_0^2 \left[\Omega_{m,0}\left(\frac{1}{a}\right)^3 + \Omega_{r,0}\left(\frac{1}{a}\right)^4 - \frac{\kappa}{H_0^2}\left(\frac{1}{a}\right)^2 + \Omega_{\Lambda,0}\right]. \tag{1.15}$$

Given that observational data, in particular, very precise measurements of the Cosmic Microwave Background (CMB) radiation, support that the geometry of our universe is *flat* ($\kappa = 0$) (Planck Collaboration et al., 2014a), we can solve the differential equation of $a(t)$ as a function of time $t$ ,

$$\dot{a} \sim a^{\frac{1}{2}(1+3w)} \rightarrow a(t) = \begin{cases} t^{2/(3+3w)} & \text{matter and radiation} \\ e^{Ht} & \text{Cosmological constant} \end{cases}. \tag{1.16}$$

The re-written Friedmann equation (1.15) shows how the three cosmological species evolve differently with respect to the scale factor, and how these species also scale in time according to equation (1.16). Knowing the values of the density parameters today ($\Omega_m, \Omega_r$ and $\Omega_\Lambda$), we can account for the expansion of the universe throughout its evolution. This difference in the scaling found in equation (1.11) shows how the universe progressed through different epochs, starting with a Radiation-domination era followed by Matter and, ultimately, by Dark Energy-domination (see Figure 1.1, where the evolution of the scale factor $a(t)$ as a function of time $t$ is depicted).

Given Friedmann equations and equation (1.16), we can now explain the thermal history of the universe. At its very early stages, the universe behaved like a hot plasma made of relativistic particles forming a thermal batch. As a consequence of the universe's expansion, the universe's temperature cooled down, allowing the decoupling of various particles from the thermal bath. At a temperature $T \approx 10^{15}\ K$, it became more energetically favourable for quarks to exist in bound states such as protons and neutrons instead of being in a quark-gluon plasma. Later, it is believed that neutrinos decoupled from the cosmic thermal bath around $T \approx 10^{10}\ K$. Below that temperature, high-energy photons (gamma-rays) can no longer produce electron-positron pairs, and the populations of both electrons and positrons started to decrease steeply by particle-antiparticle annihilation.

When the universe was approximately 3 minutes old, it had reached the necessary low temperature to allow several light nuclei (Hydrogen, Helium and some Lithium) to form through a process called Big Bang Nucleosynthesis (BBN). This mechanism was responsible for the transition between a radiation-dominated universe and a matter-dominated universe ($z \approx 3600$). The predictions made in the context of BBN (such as the primordial abundances of those nuclei, which can be calculated by studying the corresponding cross-sections) are in agreement with the current experimental evidence (Cyburt, Fields, & Olive, 2003). The primordial abundance predictions are one of the two main successes of the Big Bang Theory.

At a temperature $T \approx 1$ eV (around $z \approx 1100$) the universe entered the so-called Recombination epoch. During this time, light nuclei began to bind with electrons to form neutral atoms. These bindings induced a drop in the number density of free electrons. Hence, the decoupling of photons from matter took place, as photons

could no longer scatter off of electrons, which is called *recombination*. These photons could finally travel freely through space, with an almost isotropic distribution. This primordial radiation is called the Cosmic Microwave Background (CMB). Its experimental detection by Penzias and Wilson (Penzias & Wilson, 1965) represents the second main success of the Big Bang Theory. More details about the modelling of the CMB observables are given in subsection 1.5.2.

Finally, after recombination, only the CMB radiation is emitted, introducing an epoch called *dark ages*. At around $z \approx 20$, the Large Scale Structure (LSS) of the universe started to form, where Dark Matter (DM) is thought to have a very important role. It is assumed that DM is composed of an "exotic" yet-unknown particle[6] that decoupled at earlier stages from the thermal bath and started to collapse gravitationally into halos. Thus, after recombination, stars began to form when Baryonic Matter (BM) collapsed to the centre of these pre-existing DM halos. Afterwards, due to gravitation, galaxies were formed. These objects emit enough light to re-ionize the medium by ripping electrons from the neutral atoms. At lower redshifts $z \approx 6$, clusters of galaxies began to form thanks to gravity, building the well-known web-like structure of the universe. At approximately $z \approx 0.55$, the contribution of DE is approximately equal to the contribution of matter, causing the universe to enter into the Dark Energy-domination era.

## 1.2 A problem of initial conditions and inflation

Despite cosmological observations having accounted for experimental evidence of the BBN as well as the CMB radiation, the detection of the CMB radiation also brought some of the biggest puzzles that Modern Cosmology had to face. First of all, it confirmed that our universe is currently dominated by Dark Energy and that we need close to 30% of Dark Matter to explain the observations of the angular power spectrum of the CMB temperature anisotropies. Furthermore, it showed the necessity of fine-tuning the initial conditions of the Big Bang. The fine-tuning of the initial conditions of the universe is usually shown in terms of the *horizon problem* and the *flatness problem*. To introduce these problems, we calculate the maximum comoving distance photons can travel in space-time from emission ($t_e$) to a detecting observer ($t_d$):

$$\chi_P = \int_{t_e}^{t_d} \frac{dt'}{a(t')} = \int_{a(t_d)}^{a(t_e)} \frac{1}{aH} d\ln a, \tag{1.17}$$

where the factor $1/aH$ is denominated the *comoving Hubble radius*[7]. The first problem is that not-causally connected patches in the CMB may seem to be in thermal equilibrium. Solving the integral in equation (1.17) for the different epochs in the universe, we observe that $\chi_P$ is an increasing function during Matter and Radiation-domination. If we calculate its size at the moment of the last scattering during the

[6]There are other alternative theories to DM that aim to explain the current behaviour observed in the observational data. See for instance (Verlinde, 2017)

[7]Cosmologists misuse this term often calling to the *comoving Hubble radius* with the word *horizon*.

CMB radiation emission, we discover that the CMB map is composed of at least $\sim 10^4$ causally-disconnected patches despite showing the same temperature up to $10^{-5}$ K.

On the other hand, the second problem exposes that, given today's observed values for the different components of the universe and the fact that our universe is *flat* ($\kappa = 0$), a *flatter* universe was required in the past due to the dependence of density with time. Rewriting equation (1.15) by absorbing the time dependencies into the density parameters $\Omega_i$ we see that

$$1 - \sum_{i\in\{m,\Lambda,r\}} \Omega_i(a) = \frac{-\kappa}{(aH)^2}. \tag{1.18}$$

According to the calculations, we need to fine-tune, at least, at the moment of BBN $\sum_{i\in\{m,\Lambda,r\}} \Omega_i(a_{\rm BBN}) - 1 \approx 10^{-16}$ to obtain a flat universe $\kappa = 0$ at present time.

An elegant solution to solve these two problems is a decreasing comoving Hubble radius in a very early stage of our universe. This accelerated expansion period in the early history of the universe is called inflation. Given Friedmann equations (1.7, 1.8):

$$\frac{d}{dt}\left(\frac{1}{aH}\right) < 0 \quad \longleftrightarrow \quad \frac{d^2a}{dt^2} > 0 \quad \longleftrightarrow \quad \rho + 3p < 0 \quad \longleftrightarrow \quad \omega < -\frac{1}{3}. \tag{1.19}$$

Re-writing the first term of equation (1.19), we obtain

$$\epsilon_1 \equiv -\frac{\dot{H}}{H^2} = -\frac{d\ln H}{dN} < 1, \tag{1.20}$$

where $N$ is the *number of e-folds* (number of expansion times of the universe $a = \exp(N(t))$ so that $dN = Hdt$), $\epsilon_1$ is usually called the *first slow-roll (kinematic) parameter*, and to ensure accelerated expansion $0 < \epsilon_1 < 1$. In the limit when $\epsilon_1 \to 0$, the Hubble parameter becomes constant and therefore, the space-time becomes de Sitter space. However, inflation needs to finish and last long enough to solve the *horizon problem* (approximately between 50 and 60 e-folds). This condition is studied by introducing a new parameter, $\epsilon_2$, which relates the relative change of $\epsilon_1$ during one e-fold as follows

$$\epsilon_2 \equiv -\frac{d\ln\epsilon}{dN} = \frac{\dot{\epsilon}}{H\epsilon}, \tag{1.21}$$

and is known as the *second slow-roll kinematic parameter*. The condition $|\epsilon_2| < 1$ ensures that the change of $\epsilon_1$ is small and consequently, inflation can last. A shrinking Hubble radius (1.19) also provides a natural mechanism to explain the generation of the initial perturbations of the distribution of matter in the early universe that evolved up to the current Large Scale Structure visible today. To explain the mechanism, we use the Simple Harmonic Oscillator (SHO) as an example. Let's model the behaviour of small perturbations as damped SHOs due to the expanding space-time. Before inflation, all small-scale perturbations are inside the Hubble radius. When the comoving Hubble radius shrinks, the Hubble friction term starts to dominate and the scales become over-damped. If the scale is larger than the Hubble radius, they cannot longer move becoming *frozen*. After inflation, when the Hubble radius grows again, the friction term does not longer dominate the behaviour of the SHO, allowing the scales to *move* again. In the next section, we will introduce this mechanism using the simplest setting of single-field slow-roll inflation.

## 1.3 Canonical single-field slow-roll inflation

Inflation needed to stop at some point because we know, from cosmological observations, that the radiation and matter-domination epochs took place. For inflation to end, the space-time had to deviate from a perfect de Sitter space. Still, de Sitter space-times remains a good approximation for $\epsilon_1 \ll 1$. Together with $\epsilon_2 \ll 1$, they are the slow-roll conditions. We can discuss how microscopic physics can fulfil the slow-roll conditions by defining a new substance that violates the strong energy condition, similarly to the case of the cosmological constant. For simplicity, let us model this substance as a single scalar field $\phi$ that is homogeneous, so that $\phi(t, \mathbf{x}) = \phi(t)$. To describe the dynamics of the system, we introduce the second-order action for the field $\phi$ with a canonical kinetic term and a potential $V(\phi)$, minimally coupled to gravity, given by

$$S_2 = \int d^4x\sqrt{-g}\left[\frac{M_P^2}{2}R - \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right], \tag{1.22}$$

where $R$ is the Ricci scalar curvature of the space-time and the first and second terms of the action describe the coupling of the *inflaton* $\phi$ to gravity. All this setting is called canonical single-field slow-roll inflation. Calculating $T_{\mu\nu}$ from the action (1.22) and comparing to the expression for a perfect fluid (1.6), we obtain the corresponding expressions for the density, $\rho_\phi$, the pressure, $p_\phi$, and the homogeneous field $\phi$

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V, \qquad p_\phi = \frac{1}{2}\dot{\phi}^2 - V. \tag{1.23}$$

Introducing expressions (1.23) in (1.7) and (1.10), we can obtain the corresponding to the continuity equation as well as the Friedmann equations in terms of $V$ and $\phi$ respectively:

$$\ddot{\phi} + 3H\dot{\phi} + \frac{dV}{d\phi} = 0, \qquad H^2 = \frac{1}{3M_P^2}\left(\frac{1}{2}\dot{\phi}^2 + V\right), \qquad \dot{H} = \frac{1}{M_P^2}\left(-\frac{1}{2}\dot{\phi}^2\right). \tag{1.24}$$

The equations in (1.24) can be used to express the slow-roll parameters $\epsilon_1$ and $\epsilon_2$ in terms of the potential $V$ and the field $\phi$. To fulfil the slow-roll conditions, the kinetic energy of the inflaton field has to be negligible with respect to the potential energy ($\dot{\phi} \ll V$), which makes the Hubble parameter to be nearly constant $\dot{H}/H \ll 1$, allowing an almost-exponential expansion with $a(t) \propto \exp Ht$. Also, the acceleration of the field has to be very small ($\ddot{\phi} \ll H\dot{\phi}$), ensuring inflation to last long enough. Moreover, when the slow-roll conditions are fulfilled, we can quickly see from equations (1.23) that the equation of state of this substance should be $w_\phi \approx -1$, as in the cosmological constant case. The slow-roll parameters can be also expressed in terms of the potential $V$:

$$\epsilon_V \equiv \frac{M_P^2}{2}\left(\frac{\partial_\phi V}{V}\right) \approx \epsilon_1 \qquad\qquad |\eta_V| \equiv \frac{M_P^2}{2}\frac{\partial_{\phi\phi}V}{V} \approx 2\epsilon_1 - \frac{1}{2}\epsilon_2, \tag{1.25}$$

and the slow-roll conditions imply that $\epsilon_V, |\eta_V| \ll 1$. The potential slow-roll parameters are commonly used in the literature and can be also related to the kinetic slow-roll parameters $\epsilon_1$ and $\epsilon_2$.

### 1.3.1 Primordial perturbations

Apart from solving ad-hoc the fine-tuning of initial conditions, inflation explains the origin of the primordial seeds that evolved to the current structure observed in the universe. The origin of these primordial density perturbations lies in the quantum fluctuations of the inflaton field with respect to the homogeneous background and, also, in how the evolution of these perturbations $k$ behaves with respect to the comoving Hubble radius $aH^{-1}$. The overview is as follows (see Figure 1.2). When the universe underwent inflation, quantum fluctuations were stretched to a very large scale. When the physical wavelength[8] of this mode is larger than the scale of the Hubble radius, this mode is called *super-Hubble* ($k^{-1} \gg (aH)^{-1}$). These perturbations that exited the Hubble radius during this extreme expansion remained frozen. At later stages of the universe, when the comoving Hubble radius increased again during radiation and matter-domination (see equations (1.17) and (1.16)), these frozen modes re-entered the Hubble radius again, becoming *sub-Hubble* ($k^{-1} \ll (aH)^{-1}$) and forming the initial density fluctuations of the universe.

Following the discussion presented in (Baumann, 2009; Achúcarro & Welling, 2015), where the original references can be also found, we can illustrate this behaviour mathematically using the canonical single-field slow-roll inflation scenario as an example. For that, we use a perturbed expression of the FLRW metric in (1.1), $g_{\mu\nu}(t,\mathbf{x}) \equiv \bar{g}_{\mu\nu} + \delta g_{\mu\nu}(t,\mathbf{x})$, and of the field $\phi(t,\mathbf{x}) \equiv \phi(t) + \delta\phi(t,\mathbf{x})$. To parametrize $\delta g_{\mu\nu}$, a particular useful choice of coordinates is the *comoving gauge*,

$$\delta\phi = 0, \qquad \delta g_{ij}(t,\mathbf{x}) = a^2\left[(1-2\mathcal{R}(t,\mathbf{x}))\delta_{ij} + h_{ij}(t,\mathbf{x})\right], \qquad \partial_i h_{ij} = 0, \tag{1.26}$$

where $\mathcal{R}$ denotes the scalar metric fluctuation (comoving curvature perturbation) and $h_{ij}$ is the tensor part (gravitational waves) of the perturbed metric. In this gauge, the inflation field $\phi$ is unperturbed (the density of the fluid is constant) and the scalar degrees of freedom are parametrized by $\mathcal{R}$. Substituting equation (1.26) into the expression of the action for a single-field (1.22), and expanding in powers of $\mathcal{R}$, we find the quadratic action for the scalar metric fluctuation,

$$S_2 = \frac{1}{2}\int d^4x a^3 \frac{\dot{\phi}^2}{H^2}\left[\dot{\mathcal{R}}^2 - \frac{(\partial_i\mathcal{R})^2}{a^2}\right]. \tag{1.27}$$

Our goal is to derive the equation of motion for $\mathcal{R}$, show that it has a Simple Harmonic Oscillator (SHO) form, and promote the classical field $\mathcal{R}$ into a quantum operator. In order to solve the equations of motion derived from (1.27), the *Mukhanov-Sasaki variable* $v$ is introduced (Mukhanov, 1988; Sasaki, 1983). It is defined as

$$v \equiv z\mathcal{R} \qquad\qquad z^2 \equiv 2a^2\epsilon. \tag{1.28}$$

Using the conformal time $\tau$, defined in equation (1.3), and the Mukhanov-Sasaki variable $v$, we obtain the full quadratic action in perturbations with canonical kinetic

[8] A physical wavelength of a mode is proportional to $k^{-1}$. Larger physical wavelengths imply smaller values of $k$.

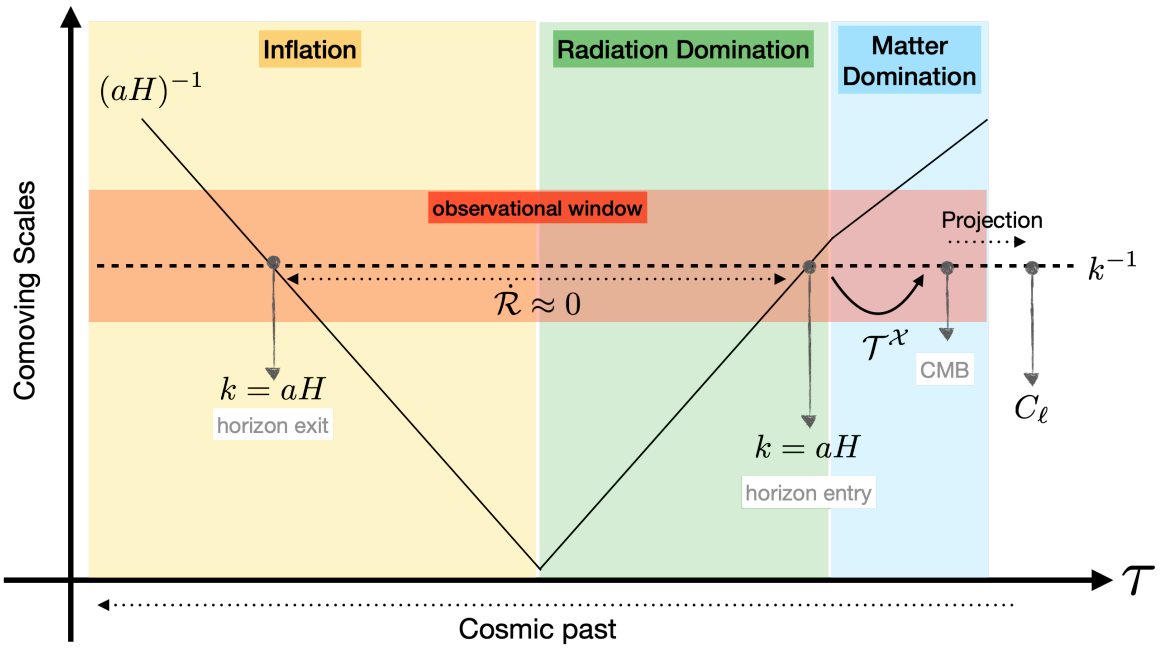


Figure 1.2: Graphical representation of the history of a particular comoving scale $k^{-1}$ (dashed black line). This comoving scale is a sub-Hubble scale ($k^{-1} \ll (aH)^{-1}$) until it leaves the comoving Hubble radius $1/aH$ due to inflation, where it remains constant and super-Hubble ($k^{-1} \gg (aH)^{-1}$). It re-enters during the radiation-domination epoch, becoming sub-Hubble again ($k^{-1} \ll (aH)^{-1}$). The evolution of $k^{-1}$ after re-entering is modelled according to the dynamics of the radiation and matter-domination epochs, and is encoded in the transfer function $\mathcal{T}^{\mathcal{X}}$, which is a mathematical tool that explains the physical phenomena in the period between inflation, radiation domination, matter domination, and the emission of the Cosmic Microwave Background. We can study the distribution of density perturbations in the angular power spectra of anisotropies of the Cosmic Microwave Background (CMB) $C_\ell$. Adapted from (Baumann, 2009).

terms:

$$S_2 = \frac{1}{2}\int d\tau d^3x \left[(v')^2 - (\partial_i v)^2 + \frac{z''}{z}v^2\right], \tag{1.29}$$

where the derivative with respect to the conformal time $\tau$ is denoted as $df/d\tau = f'$. This action demonstrates that the *Mukhanov-Sasaki variable* $v$ is the one we should use as the canonical quantization variable. In Fourier space, we obtain the equation of motion for $v$:

$$v_{\mathbf{k}}'' + \left(k^2 - \frac{z''}{z}\right) v_{\mathbf{k}} = 0. \tag{1.30}$$

where $v_{\mathbf{k}}$ is the amplitude of the Fourier mode and depends only on the magnitude of $k$. Equation (1.30) resembles the equation of a Simple Harmonic Oscillator (SHO) with a mass depending on the conformal time $m^2(\tau) \approx -z''/z$. Therefore, the quantization of the field $v$ is obtained analogously to the treatment of the quantum SHO. We promote $v$ and its corresponding conjugate momentum $v'$ to a quantum operator $\hat{v}$,

$$\hat{v}_{\mathbf{k}} = v_{\mathbf{k}}\hat{a}_{\mathbf{k}} + v^*_{\mathbf{k}}\hat{a}^\dagger_{\mathbf{k}}, \tag{1.31}$$

where $\hat{a}_{\mathbf{k}}$ and $\hat{a}^\dagger_{\mathbf{k}}$ are the creation and annihilation operators, which satisfy the canonical commutation relation

$$[\hat{a}_{\mathbf{k}}, \hat{a}^\dagger_{\mathbf{k}'}] = (2\pi)^3\delta(\mathbf{k} - \mathbf{k}'). \tag{1.32}$$

Solving equation (1.30) is complicated as $z$ depends on the background dynamics. However, some insight can be obtained if we study certain limits and constraints. However, an important constraint can be imposed at the earliest stage of our universe when $\tau \to -\infty$, which implies that all comoving scales were within the Hubble horizon (as $\tau \propto -(aH)^{-1}$, given equation (1.17)). In this limit, we choose the vacuum state for the fluctuation,

$$\hat{a}_{\mathbf{k}} |0\rangle = 0, \tag{1.33}$$

meaning that every mode $k$ is assumed to have started its evolution in the vacuum state, such that there was no particle production. According to this, modes with high $k$ do no feel the curvature of space-time, and the expectation value of the Hamiltonian in the minimal energy state corresponds to the one in flat (Minkowski) space. This boundary condition is known as the *Bunch-Davies* vacuum.

In the *quasi-De Sitter* regime, where $H$ and $\epsilon$ are approximately constant, the full solution to equation (1.30) is,

$$v_{\mathbf{k}}(\tau) = \frac{e^{-ik\tau}}{\sqrt{2k}}\left(1 - \frac{i}{k\tau}\right). \tag{1.34}$$

Now, we can trace the history of a mode $k$ mathematically by studying the asymptotic limits of the solution (1.34) in the quasi-De Sitter case ($H, \epsilon \approx$ constant):

- $|k\tau| \gg 1 \leftrightarrow k \gg aH \leftrightarrow k^{-1} \ll (aH)^{-1}$: the scale $k$ is sub-Hubble, and the solution of $v_{\mathbf{k}}(\tau)$ is dominated by the oscillating exponential part $e^{-ik\tau}$. This is just the result of the above-mentioned Bunch-Davis vacuum.

- $|k\tau| \ll 1 \leftrightarrow k \ll aH \leftrightarrow k^{-1} \gg (aH)^{-1}$: the scale $k$ is super-Hubble, and the dominant contribution to the solution is the divergent factor $1/\tau$. This means that, in this range, equation (1.34) is,

$$\lim_{\tau \to 0-} v_{\mathbf{k}}(\tau) = \frac{1}{\sqrt{2}k^{3/2}\tau} \propto \frac{aH}{k^{3/2}} \tag{1.35}$$

  where we have used that $\tau \sim (aH)^{-1}$. Recovering the definition of $v$ from equation (1.28), we find that the curvature perturbation $\mathcal{R}$ is constant and

shows that super-Hubble scales remain frozen when they leave the horizon:

$$\mathcal{R} = \frac{v}{z} = \frac{v}{a\sqrt{2\epsilon}} \propto \frac{H}{k^{2/3}\sqrt{\epsilon}} \rightarrow \lim_{\tau \to 0^-} \dot{\mathcal{R}} = 0, \tag{1.36}$$

where this expression for $\mathcal{R}$ should be evaluated for each mode at the Hubble crossing radius $k = aH$.

In conclusion, during inflation, perturbation modes exit the horizon, becoming super-Hubble scales. Quantum fluctuations in the inflaton field lead to some parts of the universe being stretched for a longer period of time than others, which means that some sub-Hubble scales exit the Hubble radius due to the accelerated expansion of the universe later than other scales. All scales remained frozen until they re-enter at later stages of the universe when the comoving Hubble radius increases again. The before mentioned local time delay of the Hubble exit of different scales translates into density fluctuations, which follow their evolution during radiation-matter domination. These local differences in density become the seeds of all the current structures of the universe.

### 1.3.2 Primordial power spectrum

We can study the statistical probability distribution of the primordial fluctuations by using the power spectrum, $P_{\mathcal{R}}(k)$, which is the Fourier transform of two-point correlation function $\langle \mathcal{R}(\mathbf{k}_1)\mathcal{R}(\mathbf{k}_2)\rangle$:

$$\langle \mathcal{R}(\mathbf{k}_1)\mathcal{R}(\mathbf{k}_2)\rangle \equiv (2\pi)^3 \delta^3(\mathbf{k}_1 - \mathbf{k}_2) P_{\mathcal{R}}(k), \tag{1.37}$$

where $\delta^3$ is the three dimensional Dirac delta and $P_{\mathcal{R}}(k)$ is only dependent on the magnitude of the momentum scale $k = |\mathbf{k}|$, and not on the direction, due to isotropy. The primordial power spectrum is usually redefined as a dimensionless quantity, $\mathcal{P}_{\mathcal{R}}(k)$, as

$$\mathcal{P}_{\mathcal{R}}(k) = \frac{k^3}{2\pi^2} P_{\mathcal{R}}(k). \tag{1.38}$$

Cosmological observations support statistically an almost scale invariant power spectrum, and this is why a phenomenological parametrization of $\mathcal{P}_{\mathcal{R}}(k)$ is broadly used. Using a Taylor-expansion up to the first order, the expression is

$$\log \frac{\mathcal{P}_{\mathcal{R}}(k)}{\mathcal{P}_{\mathcal{R}}(k_*)} \approx (n_s - 1)\left[\log k - \log k_*\right] + ... \rightarrow \mathcal{P}_{\mathcal{R}}(k) = A_s \left(\frac{k}{k_*}\right)^{n_s - 1}, \tag{1.39}$$

where $k_*$ is a pivot scale, $A_s$ is the scalar amplitude (the value of the spectrum at the pivot scale) and $n_s$ is called the *spectral index*. The spectral index quantifies the scale dependence of $\mathcal{P}_{\mathcal{R}}(k)$. As mentioned above, current observational data are consistent with a power spectrum which is *nearly scale-invariant*. One can write the spectral index $n_s$ and the scalar amplitude $A_s$ as a function of the slow-roll parameters $\epsilon_1$ and $\epsilon_2$ for canonical single-field slow-roll inflation:

$$n_s = 1 - 2\epsilon_1 - \epsilon_2. \tag{1.40}$$

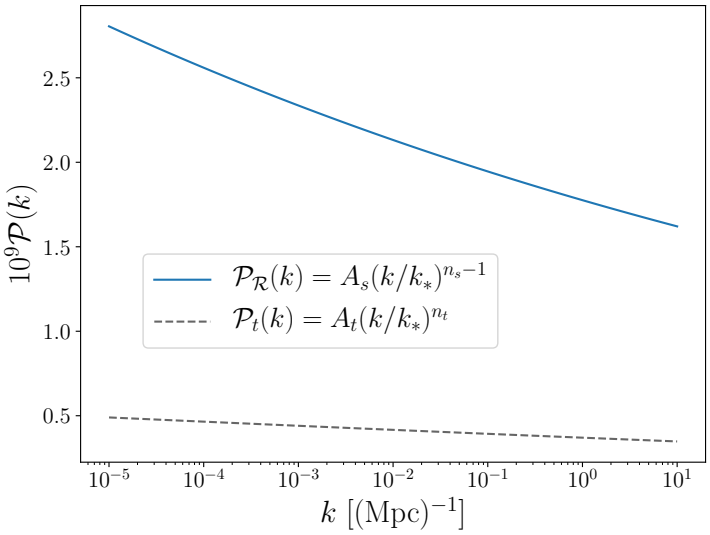


Figure 1.3: Tensor $\mathcal{P}_t(k)$ and scalar $\mathcal{P}_\mathcal{R}(k)$ primordial power spectra for the best fit values of ΛCDM taking the tensor-to-scalar ratio $r = 0.2$ for illustration purposes. Plot generated with `CAMB`.

$$A_s = \frac{H^2}{8\pi^2\epsilon_1}. \tag{1.41}$$

In the single-field slow-roll scenario, $\epsilon_1$ and $\epsilon_2$ are taken to be small, which indicates that $n_s$ should slightly deviate from one, and this is why, by construction, single-field slow-roll inflation is consistent with current observations. If we assume that the distribution of the density perturbations is Gaussian, the power spectrum contains all information to characterize it. This time-dependence of the Hubble parameter (during inflation, the Hubble parameter decreases as $\epsilon$ increases) induces a slight increment of $\mathcal{P}_\mathcal{R}(k)$ for low values of $k$, producing a *red-tilted* primordial power spectrum. Moreover, the parametrization (1.39) is also used for the perturbations of the metric (tensor modes),

$$\mathcal{P}_t(k) = A_t \left(\frac{k}{k_*}\right)^{n_t}, \tag{1.42}$$

where $A_t$ is the tensor amplitude and $n_t$ is the *tensor tilt*. The ratio between both, the scalar and tensor amplitudes, is called the *tensor-to-scalar ratio*:

$$r = \frac{A_t}{A_s}. \tag{1.43}$$

### 1.3.3 Beyond canonical single-field slow-roll inflation

As seen in the previous section, the simplest inflationary scenario (the canonical single-field slow-roll case) gives by construction a nearly scale-invariant power

spectrum, which is the one supported by cosmological observations (Bennett et al., 2013)(Planck Collaboration et al., 2014b)(T. M. C. Abbott et al., 2018). If we are interested in studying deviations from this power-law primordial power spectrum (1.39), any alternative proposed approach has to predict a spectral tilt $n_s$ smaller than one so that the power spectrum is still red-tilted.

In theoretical physics, it is customary to use a bottom-up approach to build an Effective Field Theory (EFT) to describe low-energy physical phenomena in a model-independent way. The idea behind EFT is that in nature there is often separation of scales: many aspects of physics at a given energy or length scale do not depend on detailed information about physical processes on much smaller length scales or much larger energies ("UV-scales"). EFT is a powerful framework to study the low-energy physics (with some limitations) by parametrizing the "unknowns" at UV-scales. In the case of inflation, we can have a single degree of freedom responsible for driving the perturbations and the background evolution, as in the canonical single field case we just discussed. However, in other scenarios, a single degree of freedom is responsible for the dynamics of the perturbations whereas some unknown effects on the background physics are encoded in other primordial functions. In this last case, the effective single field action for the perturbations up to second order, neglecting higher-order slow-roll corrections ($\sim \mathcal{O}(\epsilon^2)$) is given by,

$$S_2 = \int d^4x a^3 M_P^2 \epsilon_1 \left[ \frac{\dot{\mathcal{R}}^2}{c_s^2} - \frac{(\partial_i \mathcal{R})^2}{a^2} \right], \tag{1.44}$$

where $c_s(t)$ is the curvature perturbation's speed of sound and encodes part of this unknown physics. This action has a single degree of freedom and reduces to the action (1.27) when $c_s = 1$. This action is constructed following a bottom-up approach by (Cheung, Fitzpatrick, Kaplan, Senatore, & Creminelli, 2008), where the action is written in terms of the Goldstone boson of time diffeomorphisms $\pi(t, \mathbf{x})$, which is related to the curvature perturbation $\mathcal{R}$ through the relation $\mathcal{R} = -H\pi$ and assuming $\dot{\pi}^3$ to be small and approximately constant. On the other hand, this action can be also obtained in a multifield setting by integrating out the heavy fields, as in reference (Achúcarro, Gong, Hardeman, Palma, & Patil, 2012). This action has been used in chapter 2 to study localized deviations from the power-law (1.39) in the primordial power spectrum, denominated in the literature as *features*.

## 1.4 The Large Scale Structure of the universe

The Friedmann-Lemaitre-Robertson-Walker (FLRW) metric provides a good description of the universe at large scales (those larger than 100 Mpc). Nevertheless, on smaller scales, the universe is no longer isotropic and homogeneous, and it begins to show a web of clustered matter, known as the Large Scale Structure (LSS) of the universe. The description of the physics taking place at smaller scales can be done by studying the perturbations around the FLRW cosmological background, assuming

that the perturbations remain small. Scales that can be treated linearly in perturbation theory are called *linear scales*. A detailed explanation of the theory of linear order perturbations can be found in (Dodelson, 2003).

We are interested in the time evolution of cosmological perturbations, which can be divided into three different epochs. At early times, all of the perturbation modes are outside the horizon and follow the same evolution. At intermediate times, the modes' wavelengths are within the horizon and the universe evolves from radiation to matter domination. The matter-radiation equality and the stage of the horizon crossing play an important role in the evolution of the modes. At late times, all the modes follow the same evolution again. We mainly observe the distribution of matter at late epochs, which corresponds to this latest stage.

Following the same perturbation theory approach explained above, let us start by defining the dimensionless matter density contrast $\delta_{\mathrm{m}}$ in terms of the matter density $\rho_{\mathrm{m}}$ and the spatially constant density $\bar{\rho}_{\mathrm{m}}$:

$$\delta_{\mathrm{m}} = \frac{\delta\rho_{\mathrm{m}}}{\bar{\rho}_{\mathrm{m}}} = \frac{\rho_{\mathrm{m}} - \bar{\rho}_{\mathrm{m}}}{\bar{\rho}_{\mathrm{m}}}. \tag{1.45}$$

Poisson's equation, which is obtained by the theory of linear perturbations on the FLRW metric, relates the gravitational potential in Fourier space $\Phi(k, a)$ and the matter density contrast as:

$$\Phi = \frac{4\pi G\rho_{\mathrm{m}}a^2\delta_{\mathrm{m}}}{k^2} \tag{1.46}$$

We are mostly focused on studying the Fourier transform of $\delta_{\mathrm{m}}$ because at linear scales, the different Fourier modes $\mathbf{k}$ are independent. Moreover, for an isotropic Gaussian field, the Fourier-transformed quantity is described by the power spectrum. In particular, the Fourier transformed matter density contrast $\delta_{\mathrm{m}}(\mathbf{k})$ can be described in terms of the matter power spectrum $P_{\delta\delta}(k, z)$ as

$$\langle\delta_{\mathrm{m}}(\mathbf{k}_1)\delta_{\mathrm{m}}(\mathbf{k}_2)\rangle = (2\pi)^3 P_{\delta\delta}(k, z)\delta^3(\mathbf{k}_1 - \mathbf{k}_2) \tag{1.47}$$

where $\delta^3(\mathbf{k}_1-\mathbf{k}_2)$ is the three-dimensional Dirac delta function, and assuming isotropy, the vector-dependency on $\mathbf{k}$ can be dropped. The matter power spectrum can be related to the primordial power spectrum $\mathcal{P}_{\mathcal{R}}(k)$ of curvature perturbations via the matter transfer function $\mathcal{T}_m$ through

$$P_{\delta\delta}(k, z) = 2\pi^2\mathcal{T}_m^2(k, z)\mathcal{P}_{\mathcal{R}}(k), \tag{1.48}$$

where the transfer function $\mathcal{T}_m$ is introduced to describe the combined effect of physical processes that affect the growth of our perturbations such as acoustic oscillations (see subsection 1.5.2), silk damping, radiation drag, free-streaming damping, among others. The transfer function $\mathcal{T}_m$ is usually computed by the numerical Boltzmann solvers like `CAMB` (Lewis, Challinor, & Lasenby, 2000; Howlett, Lewis, Hall, & Challinor, 2012) and `CLASS` (Essinger-Hileman et al., 2014). See Figure 1.4 for a visualization of the matter power spectrum. Combining the Boltzmann equations governing

the Dark Matter evolution and the Poisson equation described in equation (1.46) we can derive a second-order differential equation for the matter density perturbation $\delta_{\rm m}$ in Fourier Space:

$$\delta''_{\rm m}(z,k) + \left[\frac{H'(z)}{H(z)} - \frac{1}{1+z}\right]\delta'_{\rm m}(z,k) - \frac{3}{2}\frac{\Omega_{\rm m}(z)}{(1+z)^2}\delta_{\rm m}(z,k) = 0\,, \tag{1.49}$$

where, here, the prime denotes derivatives with respect to $z$, and the time-redshift evolution of $\Omega_{\rm m}(z)$ is given by equation (1.15). The second term of this equation can be seen as a friction term that shows how the growth of structures can be affected by the expansion of the universe. The last term is related to the enhancement of matter densities due to gravity. Note that this equation is only valid in the limit when $\delta_{\rm m} \ll 1$ (terms of order $\delta^2$ are ignored).

The solution $\delta_{\rm m}(z,k)$ of equation (1.49) is scale-independent at later times; see e.g. (Dodelson, 2003). This scale-independence motivates the introduction of the growth factor $D(z)$ through

$$\delta_{\rm m}(z,k) = \delta_{\rm m}(z_\star,k)\frac{D(z)}{D(z_\star)}\,, \tag{1.50}$$

where $z_\star$ is an arbitrary reference redshift in the early stages of the matter-dominated era. Both $\delta_{\rm m}(z,k)$ and $D(z)$ can be used to define another important quantity, i.e. the growth rate $f(z)$:

$$f(z,k) = -\frac{{\rm d}\ln\delta_{\rm m}(z,\vec{k})}{{\rm d}\ln(1+z)}. \tag{1.51}$$

When $\delta_{\rm m}(z,k) \equiv \delta_{\rm m}(z)$ applies, the growth rate is itself a solution of a first-order differential equation:

$$f'(z) - \frac{f(z)^2}{1+z} - \left[\frac{2}{1+z} - \frac{H'(z)}{H(z)}\right]f(z) + \frac{3}{2}\frac{\Omega_{\rm m}(z)}{1+z} = 0\,, \tag{1.52}$$

with initial condition $f(z=z_\star)=1$. In the Large Scale Structure literature, the power spectrum of matter fluctuations $P_{\delta\delta}$ is usually normalised at present times by requiring that the r.m.s. variance on a sphere of radius $R_8 = 8\,h^{-1}\,$Mpc is equal to a normalisation factor dubbed $\sigma_8$, namely

$$\sigma_8^2 \equiv \frac{1}{2\pi^2}\int {\rm d}k\, P_{\delta\delta}(k,z=0)\,|W_{\rm TH}(kR_8)|^2\,k^2, \tag{1.53}$$

with $W_{\rm TH}(x) = 3(\sin x - x\cos x)/x^3$ a top-hat filter in Fourier space. It is also customary to introduce the parameter $S_8$ defined as,

$$S_8 = \sigma_8\sqrt{\frac{\Omega_{\rm m}}{0.3}}. \tag{1.54}$$

### 1.4.1 The non-linear regime

As mentioned above, equation (1.49) holds when $\delta \ll 1$. If we go to much smaller scales the perturbations become more important and the linear theory breaks down, which means that the growth of smaller massive perturbations cannot any longer be tracked with this equation. This breakdown corresponds approximately to a scale today of around $k \approx 0.1$ Mpc$^{-1}$ and it is denominated the *non-linear regime*.

Assuming the Standard Cosmological Model, structures are considered to be formed hierarchically. This means that small structures are formed first around the primordial over-densities of the matter density field. Due to gravity, these structures collapse into dark matter halos. Larger structures form later by the accretion of mass nearby or by the merging of different halos. Beyond the scales of a galactic halo or inside it, it is very difficult to study the evolution of structures.

There exist different theoretical approaches to extend the model predictions to smaller scales; for instance, the standard perturbation theory (SPT), see the review by (Bernardeau, Colombi, Gaztañaga, & Scoccimarro, 2002). However, we need to rely on prescriptions obtained from the fitting of cosmological simulations to go to smaller scales. A widely-known used prescription to model the non-linear part of the matter power spectrum $P_{\delta\delta}$ is `halofit` (R. E. Smith et al., 2003), where the authors obtained the fitting prescription from a library of N-body cosmological simulations (see Figure 1.4 for the `halofit` correction to $P_{\delta\delta}$). Although the code `halofit` is able to predict correctly the non-linear matter power spectrum up to very small scales, the fitting prescription was obtained from Cold Dark Matter simulations. This means that `halofit` may provide biased results for the non-linear modelling if massive neutrinos are taken into account or the non-linear modelling requires to go to even smaller scales where the impact of baryonic interactions is non-negligible. Because of this reason, other corrections are needed, the most famous being the Bird and Takahashi corrections (Bird, Viel, & Haehnelt, 2012) (Takahashi, Sato, Nishimichi, Taruya, & Oguri, 2012).

Recently, other tools have been developed to predict the non-linear corrections to the matter power spectrum. For example, the code `hmcode` (Mead et al., 2016) introduced physically-motivated free parameters for the halo model formalism that were fitted using high-resolution N-body simulations for a variety of cosmological models. Moreover, another mechanism to make predictions for the non-linear regime is based on *emulators* (see for instance (Heitmann, Lawrence, Kwan, Habib, & Higdon, 2014)). The idea behind emulators is based on running high-resolution simulations at key points in the cosmological parameter space so that the full parameter space can be recovered evenly. Then, an emulator is constructed to interpolate between the values of the grid so that it gives predictions for any set of parameters within the space and corrections for the matter power spectrum are computed.

## 1.5 Observations

The number of precise and new observational probes used to constrain our understanding of the underlying physics that explain the universe has increased significantly in the last decades. In this section, we detail the most important ones. Some of these observational probes rely on definitions of *distances* in our universe. Cosmologists work mainly with three different types of distances:

- **Comoving distance $\chi$:** measures the distance travelled by a light ray between an object at redshift $z$ and us in a coordinate system that expands with the scale factor $a(t)$. The comoving distance has already been introduced in equation (1.17), and rewriting it in terms of the Hubble parameter $H(z)$ we obtain:

$$\chi = \int_0^z \frac{dz\prime}{H(z\prime)} \tag{1.55}$$

  where $z$ is the redshift and $z = 0$ is usually taken at present time.

- **Angular diameter distance $D_{\rm A}$:** measures the distance to a distant object of proper size $l$ subtended by an angle $\theta$ as

$$D_{\rm A} = \frac{l}{\theta}. \tag{1.56}$$

  For flat cosmologies (i.e: $\Lambda$CDM), relating the angle $\theta$ to the comoving distance $\chi$, we arrive at the redshift $z$ dependent expression:

$$D_{\rm A} = \frac{\chi}{1+z}. \tag{1.57}$$

- **Luminosity distance $D_{\rm L}$:** measures the distance to a source by means of studying the flux $F$ of an object with known luminosity $L$:

$$D_{\rm L}^2 = \frac{L}{4\pi F} \tag{1.58}$$

  For flat cosmologies (i.e: $\Lambda$CDM), we can relate $D_{\rm L}$ with the angular diameter distance $D_{\rm A}$:

$$D_{\rm L} = (1+z)^2 D_{\rm A}. \tag{1.59}$$

### 1.5.1 Supernovae Type Ia (SNIa)

Some of the best cosmological constraints so far have been obtained using observations of Supernovae Type Ia (SNIa): a type of supernovae that does not have any presence of hydrogen but does show strong Si II absorption lines in their spectra. This type of supernovae is believed to originate from the thermonuclear disruption of carbon-oxygen white dwarfs. Due to their characteristics (they are very bright, and can be spotted at high redshift), they are used as distance indicators in Cosmology.

In particular, the main observable used in SNIa observations is the distance modulus[9] $\mu_{\rm D}(z)$:

$$\mu_{\rm D} = 5\log_{10}\left(\frac{H_0}{c}D_{\rm L}\right) \tag{1.60}$$

where $D_{\rm L}$ is the luminosity distance defined above. Note that the speed of light factor $c$ was recovered in this equation.

### 1.5.2 CMB angular power spectrum

As we mentioned in the last section, the inflationary hypothesis is not only used to explain the fine-tuning of the initial conditions issues but it provides an elegant mechanism of production of the density fluctuations that seeded the early universe. These primordial density fluctuations have been seen in the temperature anisotropies $\Delta T/T_0$ of the Cosmic Microwave Background (CMB). The CMB is a relic radiation in our universe. After Big Bang nucleosynthesis, the universe was filled with a very dense baryon plasma, where baryons tend to cluster together due to gravitation. However, the pressure created by photons was so large that it could stop the clustering. We say that the baryonic plasma was suffering *acoustic oscillations*. When the universe was approximately 380000 years old (at a redshift $z_{\rm dec} \approx 1090$), it cooled down to a temperature of approximately 1 eV, allowing light nuclei begin to bind with electrons to form neutral atoms. These bindings induced a drop in the number density of free electrons. Hence, the decoupling of photons from matter took place, as photons could no longer scatter with electrons. This is the so-called *recombination* epoch. These photons could finally travel freely through space, with an almost isotropic distribution. The relic temperature of this radiation is now $T_0 \approx 2.7$ K and makes up the CMB. The acoustic oscillations got imprinted in a characteristic scale $r_s$ defined as

$$r_s(z_{\rm dec}) = \int_{z_{\rm dec}}^{\infty} d\tilde{z}\frac{c_s(\tilde{z})}{H(\tilde{z})}, \tag{1.61}$$

where $c_s$ is the sound speed of the photon-baryon fluid.

The temperature anisotropies $\Delta T/T_0$ can be related to the primordial power spectrum $\mathcal{P}_{\mathcal{R}}(k)$ through the angular power spectrum $C_\ell^{TT}$:

$$C_\ell^{TT} = \frac{2}{\pi}\int k^2 dk \mathcal{P}_{\mathcal{R}}(k)\mathcal{T}_\ell^T(k)\mathcal{T}_\ell^T(k), \tag{1.62}$$

where $\mathcal{T}_\ell^T(k)$ is the corresponding transfer function. The angular power spectrum of temperature anisotropies $C_\ell^{TT}$ can be seen in Figure 1.4. In general, the CMB background provides more information than the one encoded in the temperature anisotropies. In fact, in the CMB map, we can also study the polarization modes $E$ and $B$ of photons. Thus, the general expression for the angular power spectra of the CMB is

$$C_\ell^{XY} = \frac{2}{\pi}\int k^2 dk P(k)\mathcal{T}_\ell^X(k)\mathcal{T}_\ell^Y(k), \tag{1.63}$$

[9]We have decided to use the subscript $D$ for the distance modulus $\mu$ to differentiate this variable from the cosine of the angle in the expression for the galaxy power spectra in subsection 1.5.3.

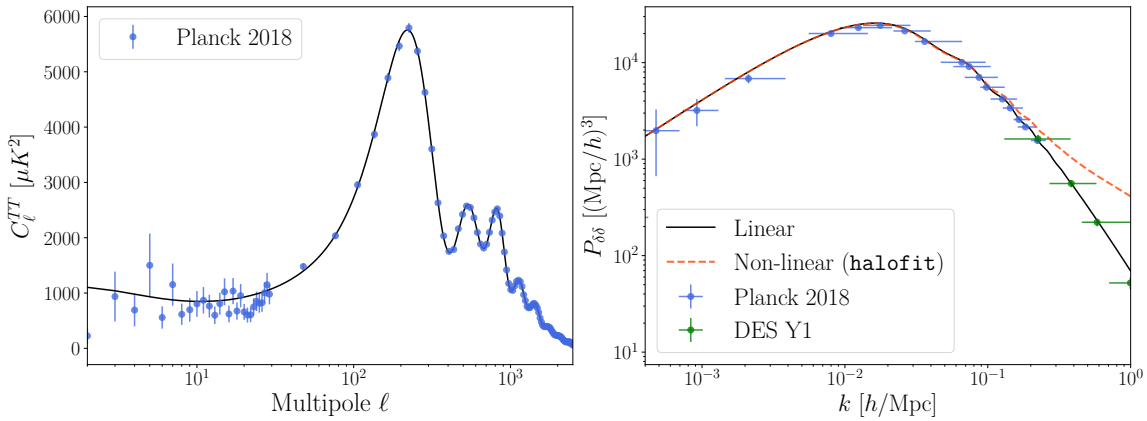


Figure 1.4: **Left:** CMB angular power spectrum of temperature anisotropies $C_\ell^{\mathrm{TT}}$ with the experimental data of Planck 2018. **Right:** matter power spectra $P_{\delta\delta}$ (linear and non-linear) with Planck 2018 and DES Y1 data. The plot has been generated with `CAMB` and `matplotlib` using the available data provided by (Chabanier et al., 2019) and (Planck Collaboration, Aghanim, Akrami, Ashdown, et al., 2020b). Solid line represents the best theoretical fit according to Planck 2018 data using the linear prediction for the matter power spectrum. The amplitude of the non-linear prediction (dashed red line), using `halofit`, has been increased for illustration purposes.

where $X$ and $Y$ may refer either to the temperature T or the polarization modes $E$ and $B$. Furthermore, $\mathcal{T}_\ell^X$ are the transfer functions, which are usually written as a line-of-sight integration in conformal time that contains the source factor $S^X(k,\tau)$ and the geometric projection $P_\ell^X(k|\tau_0-\tau|)$ based on Bessel functions (see Figure 1.2):

$$\mathcal{T}_\ell^X(k) = \int_0^{\tau_0} d\tau S^X(k,\tau) P_\ell^X(k|\tau_0-\tau|). \tag{1.64}$$

The computation of the transfer functions is usually done numerically. There are several codes that are designed for this goal. The most well-known ones are `CLASS` (Essinger-Hileman et al., 2014) and `CAMB` (Lewis et al., 2000; Howlett et al., 2012). The Planck mission (Planck Collaboration, Aghanim, Akrami, Ashdown, et al., 2020b) has measured successfully $C_\ell^{TT}$, the cross-correlation spectrum $C_\ell^{TE}$ and $C_\ell^{EE}$.

### 1.5.3 Galaxy Clustering

We have seen in section 1.4 that we can model the distribution of matter in our universe by the matter power spectrum $P_{\delta\delta}$. The key question is how we can infer from observations the matter distribution by looking at the bright distant objects present in the night sky. Galaxy Clustering (GC) describes how well the distribution of galaxies in the universe traces the underlying matter distribution. The theoretical prediction for the positions of galaxies in the universe depends on the cosmological model, so measuring the position from galaxies is a powerful proof to constrain cosmological parameters. For that, we need a robust modelling for the relation between the galaxy

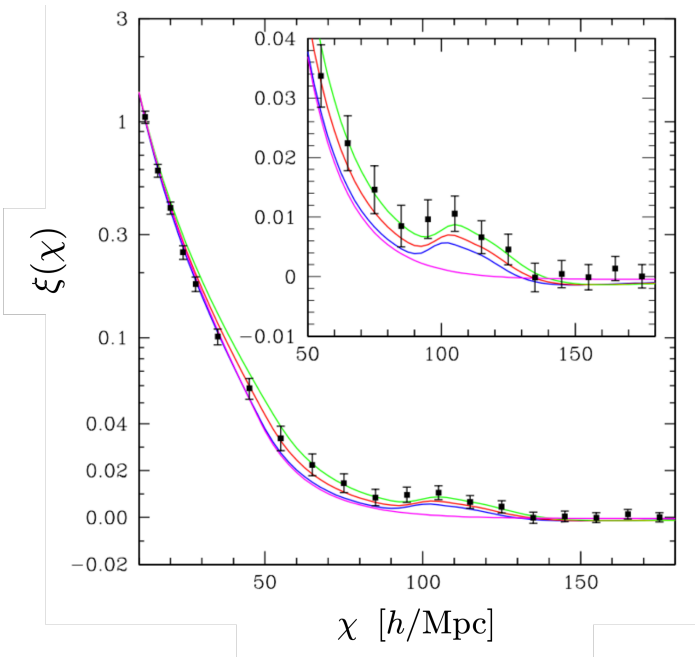


Figure 1.5: Redshift-space correlation function compared to different cosmological models. The data points are taken from the Sloan Digital Sky Survey (SDSS). The magenta line shows the prediction of ΛCDM without baryons ($\Omega_b = 0$), which demonstrates that the BAO peak is statistically significant. The plot is original from (Eisenstein et al., 2005).

and the matter distributions. This relation is encoded in the so-called *galaxy bias* $b$. One of the simplest models relating the matter and galaxy densities is given by a constant factor (Peacock & Dodds, 1994):

$$\delta_g = b\delta_m. \tag{1.65}$$

This model is known to be very simplistic but is still used for forecasting. In chapter 6 different bias models are shown.

When galaxy clustering is used as an observational probe to constrain cosmological models, other observational probes can be taken into account to increase in accuracy. One of these probes is the Baryonic Acoustic Oscillations (BAO), which do not only affect photons (as we have seen in the case of the CMB in subsection 1.5.2) but also baryons. The characteristic radius scale formed when the waves froze during electron-photon decoupling (at around $z_{\text{drag}} \approx 1020$ during the so-called drag epoch) is imprinted as an excess of power on the distribution of baryons. Since baryons and dark matter interact gravitationally, the latter would tend to cluster on this particular scale. Consequently, galaxies likely form in the higher density regions corresponding to the BAO scale $r_{\text{s}}(z_{\text{drag}})$ (see Figure 1.5).

Another probe is Redshift Space Distortions (RSD). When we measure galaxy redshifts, we need to take into account both the peculiar velocities of galaxies[10] and the Hubble flow. If only the Hubble recession is considered when we convert redshift measurements to distances, we will recover a distorted field showing RSD. We can write down the relation between the redshift space galaxy power spectrum $P_{\rm gg}$ and the real-space matter power spectrum $P_{\delta\delta}$ as (Kaiser, 1987)(Hamilton, 1998)

$$P_{\rm gg}(k, \mu_k, z) = P_{\delta\delta}(k)(b + b_\nu f(z)\mu_k^2)^2 \tag{1.66}$$

where $b$ is the bias introduced in equation (1.65), $b_\nu$ is the bias between galaxy and matter velocity distributions (which is usually neglected), $f$ is the growth rate introduced in section 1.4 and $\mu_k$ is the cosine of the angle to the line-of-sight. The RSD also carry cosmological information in the growth rate $f$.

Finally, redshift surveys such as the future *Euclid* mission (see chapter 6) will measure the positions of galaxies and their redshift. Still, to compare this information to theoretical predictions based on cosmological models, we need to translate the position and redshift of galaxies to comoving coordinates. To do this process, we need to assume a fiducial cosmological model. If the used fiducial model does not agree with the real cosmological model of our universe, some extra distortions similar to the RSD (Ballinger, Peacock, & Heavens, 1996) will be introduced. This is called the Alcock-Pacyznski (AP) effect. This effect is taken into account when the recipe of the *Euclid* mission observables is introduced in chapter 6.

### 1.5.4 Weak Lensing

Another approach to studying the formation of structures and constraining the underlying cosmological model is not to look only at the position of galaxies in the sky but also at their shapes. One of the most fascinating predictions of Einstein's Theory of General Relativity[11] (GR) is that the light from distant objects will be deflected if there is a massive concentration along its way, inducing, for instance, distortions in the shape of galaxies. This effect is called Gravitational Lensing (GL) and we can use it to infer the mass distribution in the universe. There are different regimes of gravitational lensing. When the deflection of light coming from the source is large, we are in the so-called strong lensing limit, where even multiple images of the same object or giant arcs are seen on the sky (see Figure 1.6). Another example is when we study the light coming from background distant galaxies. The magnitude of the distortions is smaller than in the strong lensing regime, where the effect induced in the shapes of galaxies is very subtle. However, we can statistically average all these small distortions to infer the mass distribution in the LSS of the universe (Kaiser, 1998). We call this limit the Weak Lensing (WL) regime, and we can relate the distortion of the shapes of galaxies with the underlying mass power spectrum.

[10]The gravitational force exerted on galaxies by their neighbouring mass can induce peculiar velocities deviating from the Hubble flow.

[11]In fact, General Relativity was widely accepted after the detection of the deflection of a starlight ray during a solar eclipse, whose magnitude was well in agreement with that predicted by Einstein's theory (Dyson, Eddington, & Davidson, 1920).

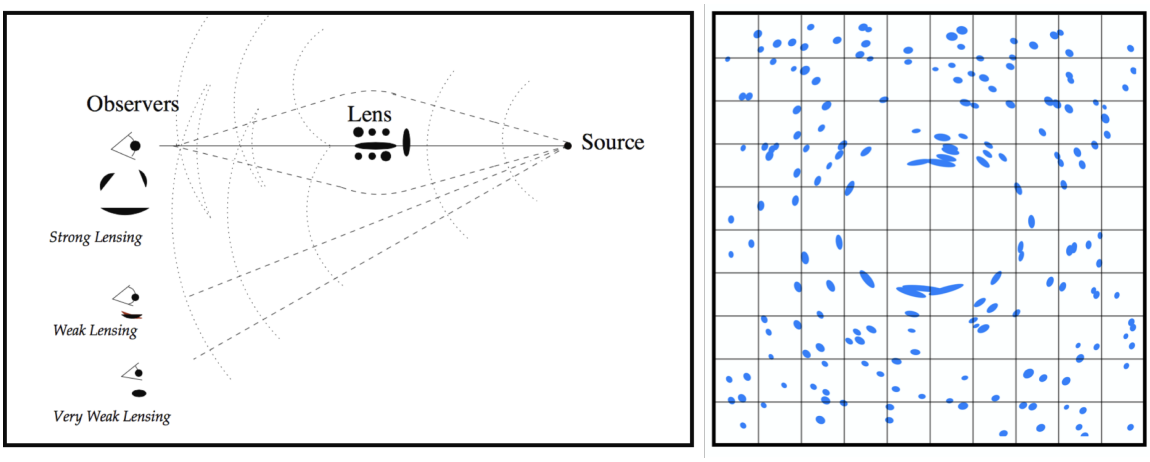


Figure 1.6: **Left**: schematic illustration of the light wavefront and the different regimes of gravitational lensing depending on the position of the observers. Deflection arcs and double images, for instance, fall in the strong lensing limit (Courbin et al., 2002). **Right**: illustration of the distortions caused by strong gravitational lensing (centre of the image) and by weak lensing (the shape of the galaxies shows small variations as we move away from the centre of the image) (Peirone, 2020).

Mathematically, the distortion of light that induces the apparent change in the intrinsic ellipticities of galaxies is described using GR. Deriving the geodesic equation for the transverse motion of photons travelling from the source towards the observer and integrating twice, we can obtain the position of the imaged shape as a function of the source position:

$$\frac{\partial\theta^i_{\text{source}}}{\partial\theta^j} - \delta^{\text{K}}_{ij} \equiv \int_0^{\chi_{\text{source}}} d\tilde{\chi}\left(1 - \frac{\tilde{\chi}}{\chi_{\text{source}}}\right)\tilde{\chi}(\Phi_{,ij}), \tag{1.67}$$

where $\chi_{\text{source}}$ is the comoving distance to the source, $\theta^j$ is the angle under the given light ray between the source and the observer, $\theta^i_{\text{source}}$ is the unlensed angle, $i, j$ represent the direction on the sky, $\delta^{\text{K}}_{ij}$ is the Kronecker delta and $\Phi$ is the gravitational potential introduced in equation (1.46). We usually describe the change between the source position and the observed one using the two-dimensional symmetric matrix:

$$\frac{\partial\theta^i_{\text{source}}}{\partial\theta^j} - \delta_{ij} = \begin{pmatrix} 1 - \kappa_{\text{GL}} - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa_{\text{GL}} + \gamma_1 \end{pmatrix} \tag{1.68}$$

where $\kappa_{\text{GL}}$ is denominated *convergence*[12], which describes the magnification of an image, and $\gamma_1, \gamma_2$ are the two components of the so-called *shear*, which describe the distortions and is the most important quantity for weak lensing studies.

Therefore, weak lensing makes circular images look like elliptical, and this effect can be studied statistically to obtain valuable cosmological information using current

[12]We have added the subscript GL to differentiate this $\kappa$ from the one appearing in the FLRW metric.

and future galaxy surveys. However, the high-quality of incoming surveys also requires accurate modelling of the systematic effects of weak lensing measurements. In this respect, it is crucial to model the ellipticity of the galaxy shapes robustly. Furthermore, the Intrinsic Alignment (IA) of galaxies must also be taken into account so that the weak lensing signal does not become contaminated. During their formation and evolution, galaxies tend to align with the matter distribution and neighbouring galaxies due to tidal forces induced by the surrounding large scale structure. These intrinsic alignments, if not taken into account, can mimic the weak lensing signal and produce a systematic signal larger than the experimental errors. In chapter 6, the modelling for the *Euclid* weak lensing observable, including IA corrections, is explained in detail.

## 1.6 The Standard Cosmological model

Current cosmological observations described in section 1.5 support statistically the Standard Cosmological Model, also known as the Concordance model or the ΛCDM model. This model concludes that we live in a spatially flat universe, which contains less than 5% of Baryonic Matter (BM) and approximately 26% of Cold Dark Matter (CDM). The other main contribution comes from Dark Energy (DE), which behaves like vacuum energy and is modelled with the cosmological constant Λ. The percentage of radiation is negligible today. The ΛCDM model assumes that the underlying distribution of the primordial density perturbations is Gaussian, and that the corresponding primordial power spectrum of curvature perturbations is almost-scale invariant.

This model is able to explain most of the current cosmological observations with only six free parameters, whose values have been obtained with high accuracy (see Table 1.1). The procedure and techniques to obtain the values of these parameters using cosmological data coming from the CMB and LSS are explained in detail in section 1.7.

In the last years, data analysis of the ΛCDM model using different observational probes has given values for the ΛCDM parameters that are in tension[13]. One example is the case is the value of the Hubble constant $H_0$, whose value obtained by direct measurements from SNIa is significantly higher than the indirect value obtained using CMB data (Riess et al., 1998). In the regime of LSS, other tensions are found. This is the case for the amplitude $\sigma_8$, whose value obtained from CMB measurements is in tension with the one obtained from cluster abundance. This tension in $\sigma_8$ induces another tension of $3\sigma$ on the value of $S_8$ measured from WL when compared to the CMB Planck results. A detailed review about the current state of cosmological tensions can be found in reference (Abdalla et al., 2022).

Different explanations aim to bring some light on these tensions. Some scientists strongly believe that the tensions may hint at new physics beyond the Concordance

[13]A parameter value given two data sets is referred to be in tension when their mean values and confidence intervals overlap partially or slightly.

| Symbol | Description | Value |
|---|---|---|
| $\Omega_b h^2$ | Density parameter of Baryonic Matter | $0.02242 \pm 0.00014$ |
| $\Omega_c h^2$ | Density parameter of Cold Dark Matter | $0.11933 \pm 0.00091$ |
| $100\theta_{MC}$ | Measure of the sound horizon at last scattering. It can be used to infer a value for $H_0$ | $1.04101 \pm 0.00029$ |
| $\tau_{reion}$ | Thomson scattering optical depth due to reionization | $0.0561 \pm 0.0071$ |
| $\ln(10^{10} A_s)$ | Scalar amplitude of the primordial power spectrum | $3.047 \pm 0.014$ |
| $n_s$ | Spectral index of the primordial power spectrum | $0.9665 \pm 0.0038$ |

Table 1.1: Mean values and 68% intervals for the base-ΛCDM model parameters from Planck CMB power spectra, in combination with CMB lensing and BAO, obtained by the Planck Collaboration. The decimal notation used in this table agrees with that of (Planck Collaboration, Aghanim, Akrami, Ashdown, et al., 2020b). The parameter $h$ represents the reduced Hubble constant defined as $h \equiv H_0/(100\,\mathrm{km\,s^{-1}Mpc^{-1}})$. The scale amplitude $A_s$ is fitted using a redefinition of it as $\ln(10^{10} A_s)$ by the Planck collaboration. The inflationary parameters $n_s$ and $A_s$ were evaluated at the pivot scale $k_* = 0.005$ Mpc$^{-1}$.

model. Others support the assumption that the data analysis is contaminated by observational systematical errors that were beyond the control of the analysis methodology. In any case, the effort of the cosmological community is undeniable, either by proposing extensions of the ΛCDM model that could explain the tensions, or by further reviewing any possible systematics introduced on the measurements or on the data analysis.

## 1.7 Data Analysis in Cosmology

Statistics are crucial to evaluate the fit of cosmological observations to any theoretical model. The starting point is to define the *scientific question*. For example, you may be interested in finding the parameters that fit best a chosen theoretical model, or you may want to discover if a model, given some parameters, is statistically preferred than another one. These questions in Cosmology are usually answered using a Bayesian Statistical framework. This section aims to be a practical guide on how Bayesian statistics are currently used in the context of Cosmology and represents the main methodology followed in the scientific content of this thesis.

### 1.7.1 Bayesian Statistics

Bayesian statistics assume that the parameters of interest, given that the model parameters are treated as random variables, have probability distributions (Ivezic, Connolly, VanderPlas, & Gray, 2014). Bayesian statistical analyses rely on *Bayes' Theorem*, which provides the probability distribution of the parameters $\theta$ given a model $M$ and the observed data $\vec{d}$. This probability distribution, $P(\theta|\vec{d}, M)$, called *posterior distribution*, is defined as

$$P(\theta|\vec{d}, M) = \frac{\mathcal{L}(\vec{d}|\theta, M)\Pi(\theta|M)}{\mathcal{Z}(\vec{d}|M)}. \tag{1.69}$$

In equation (1.69), we find several probability distributions defined below:

- $\mathcal{L}(d|\theta, M)$ is the so-called *likelihood*, which gives the probability of observing the data $d$ given the model $M$ and the parameter values $\theta$. It measures the compatibility of the data with the hypothesis.
- $\Pi(\theta|M)$ is the *prior distribution*, which is the probability distribution of the parameters $\theta$ taking into account all available external information. For instance, this information may come from previously collected data, limits imposed by theory, results of previous experiments, etc. This prior should not take into account the actual data $d$.
- $\mathcal{Z}(d|M)$ is the *evidence* (also called *marginal likelihood*), which gives the probability of observing the data given the external information and the chosen model.

#### Likelihood

The likelihood $\mathcal{L}$ is a probability distribution that determines how well a model $M$ with a given set of parameter values $\theta$ matches the given experimental data $d$. The likelihood considers the assumptions of the model as well. In Cosmology, we are especially interested in testing whether a cosmological model $M$, such as $\Lambda$CDM or extensions of it, with its parameter values $\theta$, can fit current data. Therefore, a relation between the experimental data, $d$, and the theoretical predictions for cosmological observables given by the model is needed. In many scenarios, the likelihood $\mathcal{L}$ is modeled as a multivariate Gaussian distribution function of dimension $i$:

$$\mathcal{L} = \frac{1}{(2\pi)^i\sqrt{|C|}} e^{-\frac{1}{2}(\vec{d}-\vec{T}(\theta))^t C^{-1}(\vec{d}-\vec{T}(\theta))}, \tag{1.70}$$

where $\vec{d}$ is the data vector, $\vec{T}(\theta)$ is the theory vector, $C$ is the covariance matrix of the data $\vec{d}$, $i$ is the number of variables of the data set $\vec{d}$ and $|C|$ is the determinant of the covariance matrix. This approach of using a multivariate Gaussian distribution to model the likelihood is applied in CMB experiments, for instance, or in the *Euclid* mission (see chapter 6, in particular section 6.5).

### Prior

The already known information about the parameters of interest is summarized in a probability distribution denominated prior. The information coming from previous experiments can be used as prior distributions. Moreover, it is customary in Cosmology to use Gaussian priors for the parameters $\theta$ if the results from previous experiments show that the posterior distributions of those parameters $\theta$ resemble a Gaussian distribution. In general, to define a prior for the set of parameters $\theta$, we need to set the probability distribution and the hyper-parameters[14] of the prior distribution.

In our research, in some cases, we restrict the hyper-parameters of the prior distribution according to the theory. In these cases, for each $\theta$, we assign a uniform distribution, which is defined within the limits allowed by the theory, to ensure that the probability area in the parameters space is equal for all the parameter values. These priors are sometimes called uninformative priors, although there is a notorious debate in the field of Statistics about whether uniform probability distributions are indeed the less informative distributions. In other situations, we define more complex prior distributions, according to the theory, so that we have an efficient sampling (for instance, a Beta function, described in chapter 2).

We distinguish between two types of prior distributions: the conjugate and non-conjugate priors. In the case of using a conjugate prior, the posterior distribution is in the same probability distribution family as the prior distribution (for example, if the prior is a Gausiaan distribution and the likelihood is defined as in equation (1.70), the posterior distribution is also Gaussian). Otherwise, for non-conjugate priors, the posterior distribution is not part of the same family as the prior and this usually increases the difficulty of the sampling procedure (see subsection 1.7.2).

### Bayesian Statistics vs. Frequentist Statistics

A different approach to statistical problems is the *Frequentist approach*. This approach is very common in particle physics and other fields of science. Working with an example, this approach is, in general, based on defining a *null hypothesis* such as "the Primordial Power Spectrum does not deviate from an almost-scale invariant power law", and an *alternative hypothesis*, for instance, "the Primordial Power Spectrum has some features described by the model $M$" (Chluba, Hamann, & Patil, 2015). Based on the results, the null hypothesis is either rejected or retained.

In the Frequentist approach, the most important output is not a probability density distribution of the parameters, but a numerical value called *statistic*, which is used to determine whether the null hypothesis can be rejected or not. Associated with the statistic, the assumed underlying probability distribution, and the degrees of freedom we have a $p$-value (Robert, 2011). This value is defined as the probability of obtaining a statistic value equal to or more extreme than the observed one under

[14]We call hyper-parameters the parameters of the prior distributions. Examples of hyper-parameters are, for instance, the bounds of a uniform distribution. More examples of hyper-parameters can be seen in chapter 2

the assumption that the null hypothesis is correct. If the associated $p$-value is smaller than the probability of rejecting the null hypothesis when it is true[15], $\alpha$, we reject the null hypothesis in favour of the alternative one. In general, $\alpha$ is set to 0.05.

There is a notorious debate regarding the choice of a Bayesian or Frequentist approach to tackle scientific problems. Both of them offer pros and cons: many researchers criticise the use of a prior in Bayesian Statistics; in particular, whether the selection of the prior distribution may be misleading. On the other hand, Frequentist statistics is considered less flexible. Besides, the publication bias has induced a tendency on obtaining $p$-values smaller than 0.05 (Viechtbauer, 2007). In some exceptional cases, there has been a malpractice called *p-hacking*, which consists on changing models and hypothesis on the fly during research and/or modifying the data sets to obtain a $p$-value below 0.05 (Reyes, Dieste, C., & Juristo, 2020; Head, Holman, Lanfear, Kahn, & Jennions, 2015). In any case, the use of Bayesian statistics in infering the parameters distributions of a cosmological model or in the research of extended cosmological models is more popular, due to the demanding numerical simulations and other procedures required to analyse the data in the Frequentist approach (Schafer & Stark, 2003). One of these requirements is the large number of experiments (or repetitions of measurements), such as in particle physics, where a collision can be repeated numerous times. Unfortunately, in Cosmology, we cannot perform several experiments for the same research question and this is one of the main reasons behind the application of Bayesian statistical analysis (Schafer, 2015).

### 1.7.2 Parameter Inference

When only interested in the estimation of the values of the set of parameters $\theta$ that best fit a model $M$, the normalization $\mathcal{Z}(d|M)$ of the Bayes' Theorem (Equation 1.69) is usually omitted, because it provides only a re-scaling of the normalization of the distribution. Thus

$$P(\theta|d, M) \propto \mathcal{L}(d|\theta, M)\Pi(\theta|M). \tag{1.71}$$

To infer the best value of the parameters $\theta$ we need to explore the parameter space and test if the model fits the data well for a wide range of parameter values within this space. In Cosmology, we usually deal with a very large number of parameters with non-conjugate prior distributions, which complicate the analytical evaluation of the posterior distribution $P(\theta|d, M)$ (Equation 1.71). Therefore, we need to use numerical tools to sample computationally the posterior distribution. Usually, this procedure is done by evaluating possible parameter values using a grid. These tools are based on a random sample drawn from the real posterior distribution. The most common ones are *Monte-Carlo Markov Chain* (MCMC) methods (see (Eckhardt, 1987) for a review). The name Monte Carlo simply means obtaining a representation of the distribution by sampling it randomly, while the Markov Chain improves the sampling efficiency.

[15]This is also known as type-I error.

A Markov Chain is defined by a series of random variables where the probability of the outcome of a random variable in the current step depends only on the outcome of the random variable in the previous step (Markov, 2006). We are interested in two main properties of the Markov Chains. First, we need the chain to be *stationary*: the distribution should not depend on time (meaning, it should not depend on the number of steps). Second, we need the chain to reach a state where the next elements of the chain are picked from the high-density regions of the posterior distribution (that is, the elements are selected from the region close to the likelihood maximum[16]).

In the simplest scenario for an MCMC, we compute the set of parameter values iteratively. At each iteration, we check whether the new set of values fits better than the previous one; this means that the set of values sampled from the posterior distribution is closer to the maximum of the likelihood. This step is done by checking the values given a selection criterion. Different sampling algorithms use different selection criterion. If the selection criterion is fulfilled, the algorithm selects the new set of values. Otherwise, it keeps the previous one. At the end of each iteration, the set of parameter values is saved. The ordered set of saved iterations forms what we call a *chain*. The density of points in the chain gives the posterior distributions of the parameters.

Once the posterior distribution is sampled, we usually marginalise it to show the results of a multi-parameter fit. Marginalization involves projecting the distributions to all other dimensions, and it is usually achieved by summing or integrating over the unwanted distributions of the parameters as

$$P_{\Theta_1}(\theta_1) = \int_{\theta_2} P_{\Theta_1,\Theta_2}(\theta_1, \theta_2) d\theta_2 \tag{1.72}$$

where $\Theta_1$ are the set of random variables with values $\theta_1$, which are the variables that we want to keep, and $\Theta_2$ are the set of random variables with values $\theta_2$ that we aim to get integrated (or "marginalized" over). In Cosmology, it is customary to use already available codes to post-process the MCMC chains, being one of the most popular codes `GetDist` (Lewis, 2019).

There are several sampling algorithms for the MCMC process. In this thesis, we focus on Metropolis-Hastings (MH) and Nested Sampling algorithms. In Cosmology, there are several Bayesian Analysis Framework codes designed to carry out statistical analyses of parameter inference of cosmological models against data. The most famous ones are `MontePython` (Audren, Lesgourgues, Benabed, & Prunet, 2013), `CosmoMC` (Lewis & Bridle, 2002) (and its new `python` version `Cobaya` (Torrado & Lewis, 2021, 2019)) and `CosmoSIS` (Zuntz et al., 2015). These frameworks usually contain a so-called *theory code* (for instance `CAMB`, `CLASS` or both) to provide theoretical predictions of the cosmological observables, different data sets and several *samplers* (such as the nested samplers `PolyChord` (W. J. Handley, Hobson, & Lasenby, 2015a, 2015b) or `MultiNest` (Feroz & Hobson, 2008; Feroz, Hobson, & Bridges, 2009; Feroz, Hobson, Cameron, & Pettitt, 2019), and Metropolis-Hastings MCMC samplers).

[16]Maximize the likelihood $\mathcal{L}$ is equivalent to minimize the $-\log \mathcal{L}$.

### Metropolis-Hastings algorithm

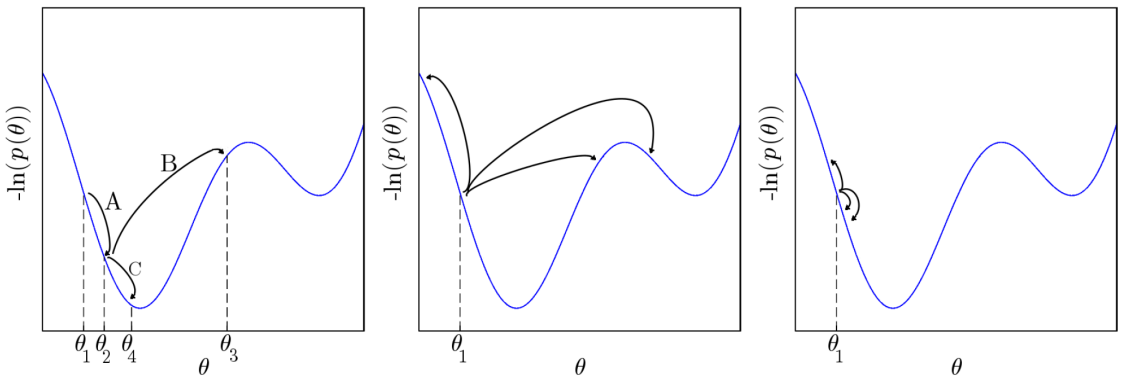


Figure 1.7: **Left**: an example of a Markov Chain constructed by the Metropolis-Hastings algorithm. It starts at 1, then 2 is proposed and accepted (step A), 3 is proposed and refused (step B), and finally, 4 is proposed and accepted (step C). The resulting chain is $(\theta_1, \theta_2, \theta_2, \theta_4)$. **Central**: an example of what happens when the proposal step is too broad: the different proposed new set of values are further away from the minimum of the log-posterior, and therefore, they get rejected. As a result, the chain lacks mobility because all the proposals are unlikely and it may get stuck. **Right**: an example of what happens with a proposal step that is too narrow: all the proposed new set of values are accepted, making the chain to move towards the minimum of the log-posterior but it samples the parameter space very slowly and the convergence takes longer. Figure from (Leclercq et al., 2014).

This algorithm (Hastings, 1970) is based on a particular choice of the selection criterion to decide whether to keep the set of values of the parameter space during the current step $\theta'$, or to come back to the previous step $\theta$. This criterion is based on calculating a ratio, called *acceptance ratio*, $a$, defined as[17]

$$a = \frac{p(\theta')q(\theta|\theta')}{p(\theta)q(\theta'|\theta)}, \tag{1.73}$$

where $p(\theta')$ is the posterior distribution and, the proposal distribution $q(\theta'|\theta)$, is the conditional probability of proposing a new $\theta'$ given the previous $\theta$ values. In general, $q$ is proposed to be a Gaussian distribution. At each step, a sample of $\theta'$ from $q(\theta'|\theta)$, and a random number $v$ from a uniform distribution between 0 and 1, are drawn. If $a \geq v$, we accept $\theta'$, and it becomes a new state of the chain. Otherwise, we reject $\theta'$ and the new state of the chain is $\theta$ again. This is shown in figure 1.7. In this figure, it can be observed that each accepted step only depends on the previous one and that, eventually, it is independent of the number of steps, making the chain a Markov Chain. To speed up the algorithm, a covariance matrix $C$ of previous sampling runs can be used in the likelihood calculation so that the algorithm samples

[17]Do not confuse with the scale factor $a(t)$ used in previous sections.

more efficiently the parameter space with prior information obtained from another statistical study.

Still, the major disadvantage of this algorithm is that the sampled sets of parameters are correlated because we only use the previous sample to obtain the current one. This means that, if we start the algorithm with a value within the parameter space that has a low probability, the chain will not reflect the underlying distribution very well until a higher likelihood region is reached. Therefore, a small percentage of the initial values of the chain are usually discarded (the so-called *burn-in* phase), in order to achieve the stationary property. Sometimes, the chains may also get stuck in a lower probability region and not be able to escape from it to further explore the parameter space. This is why, when this algorithm is used, several chains with different initial values are usually run. In any case, this algorithm also suffers in multi-modal posterior distributions where the chain can get stuck in relative maxima of the likelihood (see for instance multi-modal posterior distribution at Chapter 2). A similar problem will arise if there are high correlations or degeneracies among the parameters (showing correlated peaks in the posterior and "*banana*" contour plots).

The Metropolis-Hastings algorithm allows us to construct a Markov Chain containing as many values as parameters we are trying to infer per step. Once that the chain is stationary, we need to decide if we can stop the sampling. Therefore, a question arises: *how do we decide this?* A simple answer is not available, as no test can affirm whether a chain has converged or not. Still, some convergence diagnostics may point out some necessary conditions for a chain to show some convergence, although they are not sufficient conditions. Some diagnostics include:

- Individual segments of the chain show similar results, as long as the chain is much longer (sufficient large number of steps in the chain) than any obvious correlation.
- A reasonable number of accepted proposed steps. If the acceptance is too high, it may point out that the chain is slowly converging; however, if the acceptance is too small, it may indicate that the chain is locally stuck. An optimal acceptance rate for a Metropolis-Hastings MCMC is considered to be around 25-30% (Gelman, Gilks, & Roberts, 1997).
- As a result of running several chains at the same time with different starting points to avoid them getting stuck, we can apply a different stopping criterion: the *Gelman-Rubin* statistical criterion R (Gelman & Rubin, 1992). This test consists of comparing the variance within individual chains to the variance between chains. Their ratio, R, should be close to 1 when the chains have converged. In Cosmology, the usual convergence ratio is $R-1<0.01$.

**Nested Sampling**

To circumvent the issues presented by the Metropolis-Hastings algorithm, an alternative way of sampling the posterior was developed by J. Skilling (Skilling, 2006):

the *Nested Sampling* algorithm. This algorithm is based on calculating the evidence $Z$ (marginalized likelihood). From the Bayes' theorem (1.69):

$$Z(d|M) \times P(\theta|d, M) = \mathcal{L}(d|\theta, M) \times \Pi(\theta|M) \rightarrow \tag{1.74}$$

$$\int Z(\theta) \times p(\theta) d\theta = \int \mathcal{L}(\theta) \times \Pi(\theta) d\theta \quad \rightarrow \quad Z = \int \mathcal{L} dX. \tag{1.75}$$

where we have defined $dX$ as $dX = \Pi(\theta)d\theta$ and the prior and posterior are normalised to the unit total. From this equation (1.75), we observe that $Z$, which is the normalization in equation (1.69), is defined as the area below the curve $\mathcal{L}dX$ (see right panel of Figure 1.8). Therefore, the algorithm developed by Skilling set up a genuine way to estimate $Z$ doing a numerical integration, whereas the sampling of the posterior $p$ is a by-product of this calculation.

In 1-dimension, the algorithm works as follows: in the first step, we generate a starting set $S_0$ of $n$ samples uniformly distributed over the space and allowed by the prior $\Pi$. Next, we delete the lowest likelihood sample $\mathcal{L}_0$ in $S_0$, and replace it with a new uniform sample with a higher likelihood $\mathcal{L}_1 > \mathcal{L}_0$ (this is called a hard constraint on likelihood value), moving to step $S_1$.

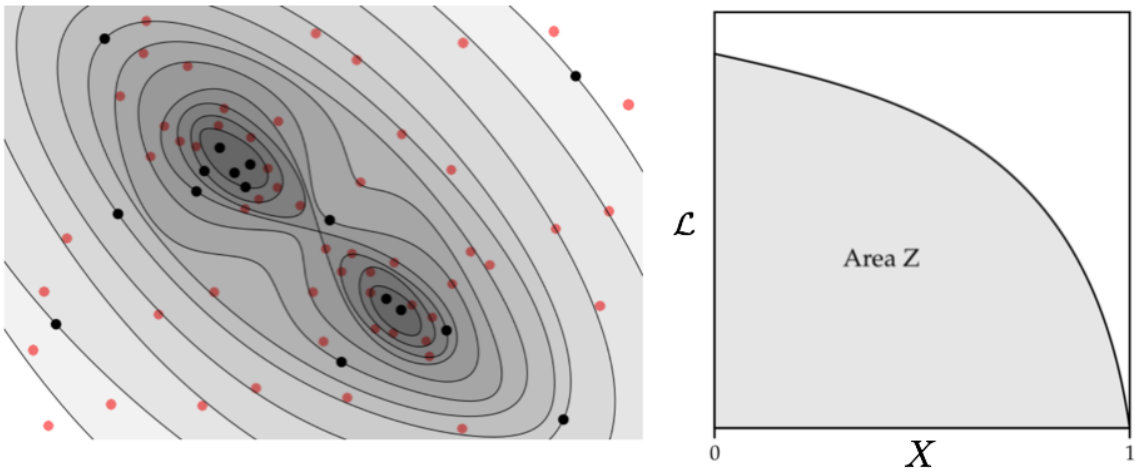


Figure 1.8: *Left*: Graphical interpretation of the Nested Sampling algorithm by looking at the parameter space: the contours specify the areas of the search after each step, where the point with the lowest likelihood was discarded. In every step, the contours are reduced, moving closer to the regions of higher likelihood and finding different likelihood clusters (separated contour regions with high likelihood). This figure is the result of an animation from *W. Handley* lecture at (Lesgourgues, 2018). *Right*: Likelihood function $\mathcal{L}$ as a function of the parameter $X = \int d\theta \Pi(\theta)$, showing how the area below the curve is the evidence $Z$. Adapted from (Skilling, 2006).

The live-evidence, $Z_{live}$, is related to the live points $X_{live}$ and their corresponding likelihood $\mathcal{L}_{live}$. The live points $X_{live}$ are the survivor points of the n+1 sample $S_{n+1}$ (see left panel of figure 1.8, black points). The live evidence is approximately equal

to $Z_{live} \approx \langle \mathcal{L}_{live} \rangle X_{live}$ (the value of the likelihood is similar when we are reaching convergence, so we can approximate the integral as the average value of the likelihood of every point times the number of points). The last process is repeated until the live evidence, $Z_{live}$ is a small fraction of the total evidence $Z$. The algorithm needs a precise value for this fraction, given by the user, to determine if convergence is achieved or not. The value of this fraction is what we call the "stopping criterion". The set $S$ of $n$ samples is constantly updated after every step and the set of dead points (the erased points with low likelihood) with an appropriate weighting factor are the posterior samples (see left panel of figure 1.8, red points). This procedure can be generalized to multiple dimensions (Skilling, 2006). The points chosen randomly from the region $\mathcal{L}(X)$ are representative of the posterior, and therefore, the samples of the posterior distribution $p$ are obtained by assigning weights $w_j$ to the discarded points so that:

$$p(\theta) \approx \frac{w_j \mathcal{L}(\theta)}{Z}. \tag{1.76}$$

### 1.7.3 Model Comparison

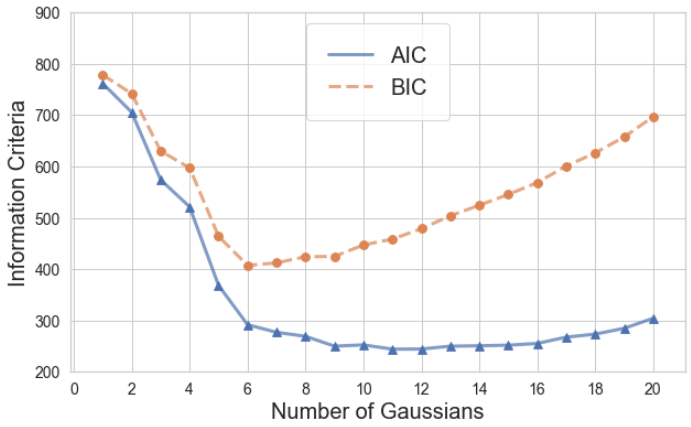


Figure 1.9: Bayesian Information Criteria: proof-of-concept. The data set `make_moons` from the `python` package `sklearn` has been fitted using the Gaussian Mixture model, also available in `sklearn`. The horizontal axis represents the number of Gaussian functions used to fit the data (more Gaussian functions represent an increment in the number of parameters $k$). The vertical axis shows both the *Akaike Information Criterion* (AIC) and the *Bayesian Information Criterion* (BIC). This example shows the asymptotic behaviour of the AIC (blue triangles) and the growing tendency for the BIC (orange circles), after some local minima. We conclude that the most statistical convenient number of Gaussian functions is 6 (according to the BIC) and 9 (according to the AIC, although after 6 the gain is not really significant) because, after those local minima, both criteria remain either stable (AIC) or increase (BIC). Example inspired by the documentation of `sklearn` (*Scikit-learn 1.1.1 documentation: Gaussian Mixture Model Selection*, 2022).

In many different scenarios, we are interested in knowing whether a model $M_1$ is statistically favoured with respect to a different model $M_2$. In Bayesian statistics, this is usually computed by calculating the *Bayes Factor*, which is inferred from applying model comparison to Bayes' theorem (Skilling, 2006):

$$B = \frac{P(M_1|d)}{P(M_2|d)} = \frac{Z(d|M_1)\Pi(M_1)}{Z(d|M_2)\Pi(M_2)}, \tag{1.77}$$

where $P(M|d)$ are the posterior probability distributions integrated over all the parameter space $\Theta$. If our sampling algorithm allows us to calculate evidences $Z$, for example with the Nested Sampling algorithm, we can compute $B$ directly by taking the ratio of the evidences and priors (in some circumstances when the priors are uniform, the ratio of priors $\Pi(M)$ is assumed to be close to unity). If $B$ is larger than one, this means that the data support the model $M_1$ statistically better in comparison to model $M_2$.

The numerical integration to obtain $P(M|d)$ may be infeasible, because, in general, we might need a large number of samples. For this reason, alternatives to the Bayes Factor are sometimes used, which are defined in terms of the likelihood $\mathcal{L}$ and the complexity of the model, and are called *information criteria*. In general, these methods do not simply choose one model versus a different one by analysing which likelihood is the highest, because a model with more parameters mostly leads to a higher likelihood. In most cases, the information criteria penalize a model according to its number of parameters $k$. The most popular information criteria are the *Akaike Information Criterion* (AIC) (Akaike, 1974) and the *Bayesian Information Criterion* (BIC) (Stoica & Selen, 2004). The AIC and BIC are defined respectively as

$$AIC = -2\log\mathcal{L} + 2k, \tag{1.78}$$

$$BIC = -2\log\mathcal{L} + k\ln N, \tag{1.79}$$

where $N$ is the number of data points used in the fit. In principle, the model with lower BIC or AIC is favoured. It is worth mentioning that the BIC and AIC tend to show an asymptotic behaviour beyond a certain number of parameters (see figure 1.9), implying that no gaining is obtained when more parameters are included in the model. The quantity $-2\log\mathcal{L}$ is known as the *deviance*[18].

## 1.8 About this thesis

Since the last decade, Modern Cosmology is increasingly becoming heavily-based on data science, in particular, on Bayesian Statistics. This is the primary methodology and common motto of all the chapters of this thesis. The thesis is divided into three different parts:

[18]In Cosmology, we denote the deviance $-2\log\mathcal{L}$ with the symbol $\chi^2$. This is a language misuse, probably because the deviance follow a $\chi^2$ probability distribution.

- The first part is focused on data science and inflation, pursuing to constrain inflationary models using advanced inference techniques and forecasting tools. **Chapter 2** shows the first-ever results of the reconstruction of the inflaton's speed of sound using the latest Cosmic Microwave Background (CMB) data from Planck 2018 and modern algorithms (Gaussian Processes). **Chapter 3** is dedicated to the forecast of a particular class of single-field inflation models, known as $\alpha$-attractor, for a future CMB stage-IV experiment using a model-dependent alternative approach for the sampling of the inflationary parameters based on current constraints obtained by CMB and LSS data. These two chapters are based on references (Cañas-Herrera, Torrado, & Achúcarro, 2021) and (Cañas-Herrera & Renzi, 2021), respectively.

- The second part of the thesis is dedicated to the novel interest in the cross-correlations of Gravitational-Wave (GW) physics and Large Scale Structure observables, concretely, Galaxy Clustering. In particular, we study how we could exploit the information contained in these new observables by forecasting their behaviour and possible detection using future experimental set-ups and statistics. **Chapter 4** focuses on the study of unresolved GW events that form the Astrophysical Gravitational-Wave background, and how we can use the cross-correlation of the anisotropies of that background with Galaxy Clustering to extract both astrophysical and cosmological information. This work is based on (Cañas-Herrera, Contigiani, & Vardanyan, 2020). On the other hand, in **Chapter 5** we investigate how machine learning techniques can be used to reconstruct the propagation of tensor perturbations by combining the spatial correlation between resolved GW mergers and galaxies. This chapter is based on (Cañas-Herrera, Contigiani, & Vardanyan, 2021).

- Finally, the third part of this thesis is dedicated to the *Euclid* mission and the current work developed by the author of this thesis as a member of the *Euclid Consortium*. **Chapter 6** focuses on a crucial data science tool for the mission: the code *Cosmological Likelihood for Observables in Euclid* (`CLOE`). In this chapter, the implemented cosmological recipe, as well as the description of the *Euclid* likelihood, are depicted. The results concerning the structure of `CLOE` and its performance to constrain cosmological parameters are a novel preview of the content of the *Key Project papers* of the *Euclid* Science Working Group Inter-Science Taskforce Likelihood (IST:L).