Search for features due to a reduction in the inflaton's speed of sound in the CMB

Thesis

submitted in partial fulfillment of the
requirements for the degree of

Master Of Science
in
Physics

Author :	Guadalupe Cañas Herrera
Student ID :	
Supervisor :	Ana Achúcarro
Co-supervisor :	Matteo Martinelli
2^{nd} corrector :	Alessandra Silvestri

Leiden, The Netherlands, Tuesday 17$^{\text{th}}$ July, 2018

Search for features due to a reduction in the inflaton's speed of sound in the CMB

Guadalupe Cañas Herrera

Huygens-Kamerlingh Onnes Laboratory, Leiden University
P.O. Box 9500, 2300 RA Leiden, The Netherlands

Tuesday 17$^{\text{th}}$ July, 2018

Abstract

The inflationary hypothesis was introduced as a solution to the fine-tuning issue in the initial conditions of the Big Bang theory. In this Master's research project, we introduce our work in the search for features in the Cosmic Microwave Background (CMB) power spectra that could result from reductions in the speed of sound of the inflaton. We study these features in the context of an effective single field theory of a multiple field scenario, due to the fact that a single field inflation approach is favoured by the current cosmological data, especially by the CMB. First, we present a brief review of the current cosmological model, the ΛCDM model, inflation and the possible extensions. Secondly, we review the physics of the CMB, the main theoretical cosmological codes and the needed data analysis tools from the point of view of Bayesian statistics. Finally, we update our current search for features using Planck 2015 temperature and polarization data introducing new parametrizations for the reduction of the speed of sound. In this search, we have recovered only some previous found modes, indicating the dependency of our results with respect to the parametrization we were using. For this reason, we have pointed out the necessity of reconstructing the reduction of the speed of sound, showing some preliminary results when Gaussian Processes are used as the reconstruction technique.

To my parents and Juanchito.
- Everyday, my love for you grows faster,
than any inflationary period.

Table of contents

Nomenclature

Acronyms / Abbreviations

ADM Arnowitt-Deser-Misner

CMB Cosmic Microwave Background

(C)DM (Cold) Dark Matter

BBN Big Bang Nucleosynthesis

BICEP Background Imaging of Cosmic Extragalactic Polarization

BM Baryonic Matter

CAMB Code for Anisotropies in the Microwave Background

CLASS Cosmic Linear Anisotropy Solving System

CMB Cosmic Microwave Background

EFT Effective Field Theory

FLRW Friedmann-Lemaitre-Robertson-Walker

GPs Gaussian Processes

GW Gravitational Waves

LIGO Laser Interferometer Gravitational-Wave Observatory

LISA Laser Interferometer Space Antenna

LSS Large Scale Structure

MCMC Monte-Carlo Markov Chain

ML Machine Learning

PPS Primordial Power Spectrum

SHO Simple Harmonic Oscillator

SRFT Slow-Roll Fourier Transform

WMAP Wilkinson Microwave Anisotropy Probe

Chapter 1

Preface

"Cosmologists are (masters) of
knowledge within ignorance"

Vicente Atal [17]

For many years, cosmologists have tried to narrate the biography of the universe with the purpose of establishing its structure, properties, evolution along time, and fate. Thanks to cosmological observations, we may have some confidence in its evolution from the time of creation of the first nuclei up to the present. Yet, in earlier times, we lack cosmological information and speculation dominates.

Still, a substantial development in modern cosmology has been made in the latest years, improving the understanding of our universe. Cosmologists have worked hard to find a cosmological model that is able to explain the current data. This theory is the ΛCDM model. According to this model and observations, our universe has mainly two components: Dark Energy (Λ) and Dark Matter (DM). Apart from the remarkable fact that we *know* the "ingredient recipe" of our universe, we also agree when we state that these two main constituents are *unknown*.

Furthermore, this cosmological model is based on the hypothetical inflation paradigm. In order to explain the observations, we postulate that the universe, at the very early stages of its evolution, expanded almost exponentially in a very short period of time. This is the key to explain why the universe seems (and is) causally connected and flat, apart from giving a mechanism to seed the inhomogeneities of our universe, which are the origin of its Large Scale Structure (LSS).

But, can we support inflation experimentally? At the moment, the cosmological data favour a model with only one field responsible for driving inflation. However, most unification theories predict the existence of more degrees of freedom. For this reason, we could attempt to use an Effective Field Theory (EFT) approach. This approach can build bridges to conciliate both points of view.

The main aim of this thesis is to analyse the hypothetical imprint that the reduction of the speed of sound of the inflaton can leave in the sky when it is supported by a heavy extra field. For that, we have developed a modification of the ΛCDM model and we have tested it against the Cosmic Microwave Background (CMB) data in order to find constraints on the primordial power spectrum.

This thesis is organized as follows. In chapter 2, basic theoretical concepts of the thermodynamical history of the universe and the currently accepted cosmological model are outlined. Moreover, also in this chapter, we argue why we need inflation: we review the major fine-tuning problems in the initial condition of the Big Bang theory, and how inflation is a solution to them. Furthermore, in chapter 3, extensions of the simplest model of inflation from an EFT point of view are outlined, focusing on the possible observational signatures. In chapter 4, we describe the main analysis tools required to pursue the goal of this research. For that, we start by explaining the information encoded in the CMB and how to use it to study its agreement with our cosmological model. Furthermore, an introduction to Bayesian Statistics, data manipulation and Machine Learning (ML) is also contained in this chapter. In chapter 5, we show our main results and the current state of art, explaining how we hare partially recovered some promising results from previous research of features in the primordial power spectrum. Finally, we expose our conclusions and future plans in chapter 6.

In order to keep this thesis to manageable proportions, three appendices have been also included. The first two detail the computational work carried out parallel to the search for features. In appendix A, we present the new cosmological code COBAYA, whose performance has been tested in this project. In appendix B, we describe the relevant modification we have coded in the cosmological code CAMB, which is planned to be included in the new public version release in the near future. The last appendix, C, describes how the effective single field action used to predict the features in the correlation functions is obtained from an inflationary theory where a heavy field supports the inflaton.

This research project shows not only the already accomplished results, but our current work in the search for features. This work is aimed to be expanded in the incoming years. Therefore, up to my mind, this thesis not only represents the final stages of my Master's education but also the beginning of the next researching steps as a PhD student.

Chapter 2

Introduction to Standard Cosmology

> "One may wonder: What came before? If space-time did not exist then, how could everything appear from nothing?... Explaining this initial singularity—where and when it all began—still remains the most intractable problem of modern cosmology"
>
> *Andrei Linde* [5]

In this self-contained chapter, a review of the most important concepts of the standard model of cosmology can be found. The main ideas and concepts are based on [19] and [22], and this is why, for the sake of brevity, most of the detailed calculations are not shown. In this thesis, we are using natural units $c = \hbar = 1$, and the reduced Planck mass, M_P, is defined as $M_P = \sqrt{\frac{1}{8\pi G}}$. The derivative with respect to the cosmic time t is denoted as $df/dt = \dot{f}$, and with respect to the conformal time τ, $df/d\tau = f'$.

2.1 Evolution of a homogeneous and isotropic Universe

Cosmological observations support the idea that the Universe is mostly *isotropic* on large scales. If we also assume that we are not a special observer in the Universe, *homogeneity* of space is inferred. 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, defined as

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a^2(t)\gamma_{ij}dx^i dx^j, \tag{2.1}$$

where:

- μ, ν, i, j are the several indices that take the values $\{\mu, \nu\} = \{0, 1, 2, 3\}$ and $\{i, j\} = \{1, 2, 3\}$.
- $a(t)$ is the so-called 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 grid points on a coordinate grid that grows with the expansion of the universe according to the scale factor a.

- 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).
- γ_{ij} is the tensor corresponding to the spatial part, which using spherical coordinates (r, θ, ϕ) and together with the differentials $dx^i dx^j$ takes the form:

$$\gamma_{ij} dx^i dx^j = \frac{dr^2}{1 - \kappa r^2} + r^2 d\phi^2 + r^2 \sin\theta^2 d\theta^2, \tag{2.2}$$

  where κ is the intrinsic curvature of 3-surfaces, and represents a *flat* ($\kappa = 0$), *positively curved* ($\kappa = +1$), or *negatively curved* ($\kappa = -1$) spatial slice. Also, we usually call the angular part $d\Omega$:

$$d\Omega^2 \equiv r^2 d\phi^2 + r^2 \sin\theta^2 d\theta^2. \tag{2.3}$$

For convenience, we often perform a *conformal* transformation of the metric in equation 2.1 using the scale factor as a common factor,

$$d\tau = \frac{dt}{a} \rightarrow ds^2 = a^2(\tau) \left[-d\tau^2 + \gamma_{ij} dx^i dx^j \right], \tag{2.4}$$

where τ is usually called *conformal time*. Light travels along null-geodesics ($ds^2 = 0$), and according to the transformed FLRW metric (2.4), its propagation is the same as in Minkowski space. The conformal time, although it simplifies the equations, does not measure the proper time corresponding to any particular observer in the universe.

Therefore, as $a(t)$ is the factor that defines the evolution of the universe at large physical distances, understanding the history of the universe is *analogous* to determining the dependence of the scale factor on function of time. To this aim, it is necessary to solve the *Einstein Equations*. In these equations, an expression for the *Energy-Momentum* Tensor, $T_{\mu\nu}$, is needed. It corresponds to the *matter* side of the equations. The cosmological principle requires that $T_{\mu\nu}$ takes the form of a perfect fluid, which in comoving coordinates is defined as,

$$T_{\mu\nu} = diag(\rho, p, g_{ii}), \tag{2.5}$$

where ρ is the density, p the pressure, and g is the metric of equation (2.1). Calculating the Einstein Equations with the above given expression for $T_{\mu\nu}$, we obtain the well-known First and Second[1] *Friedmann* equations respectively,

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{\rho}{3M_P^2} - \frac{\kappa}{a^2}, \tag{2.6}$$

$$\dot{H} + H^2 \equiv \frac{\ddot{a}}{a} = -\frac{1}{6M_P^2}(\rho + 3p). \tag{2.7}$$

[1]The Second Friedmann equation is usually called acceleration equation.

where H is the *Hubble* parameter. Combining both equations (2.6) and (2.7) we obtain the continuity equation equivalent to $\nabla_\mu T^{\mu\nu} = 0$ as,

$$\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + p). \tag{2.8}$$

Usually in cosmology, the equation of state of a perfect fluid is defined by the dimensionless state parameter, w, which is the ration between the pressure p and the density ρ,

$$p = w\rho \qquad \text{with} \qquad w = \begin{cases} 0 & \text{Matter (m)} \\ \frac{1}{3} & \text{Radiation (r)} \\ -1 & \text{Cosmological Constant } (\Lambda) \end{cases} . \tag{2.9}$$

Introducing equation (2.9) into equation (2.8), we can derive the evolution of the density of each species i as a function of the scale factor $a(t)$,

$$\rho_i = \rho_{i,0}\left(\frac{a}{a_0}\right)^{-3(1+w_i)} \tag{2.10}$$

where the sub-index 0 indicates a reference value, which is usually taken as the one at present time, and for convenience, we normalize the scale factor $a_0 = a(t_0) = 1$. With equation (2.10), we can determine the evolution of the scale factor, a, on function of time in the presence of these 3 different species: relativistic particles, such as photons (Radiation), baryonic and dark matter (Matter), and cosmological constant. Defining the critical density[2] $\rho_c \equiv 3M_p^2 H_0$ and the density parameter $\Omega_{i,0} = \rho_i/\rho_c$ (H_0 is the Hubble parameter evaluated today), we can re-write Friedman equation (2.6) as,

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = H_0^2\left[\Omega_m\left(\frac{1}{a}\right)^3 + \Omega_r\left(\frac{1}{a}\right)^4 - \frac{\kappa}{H_0^2}\left(\frac{1}{a}\right)^2 + \Omega_\Lambda\right], \tag{2.11}$$

and solving the differential equation of $a(t)$ as a function of time t for a flat geometry ($\kappa = 0$) we obtain,

$$\dot{a} \sim a^{\frac{1}{2}(1+3w)} \rightarrow a(t) = \begin{cases} t^{2/(3+3w)} & \text{Matter and Radiation} \\ e^{Ht} & \text{Dark energy} \end{cases} . \tag{2.12}$$

[2]The definition of critical density responds to historical reasons. It is the density of a universe with flat curvature.

2.2 Conventional Big Bang Theory

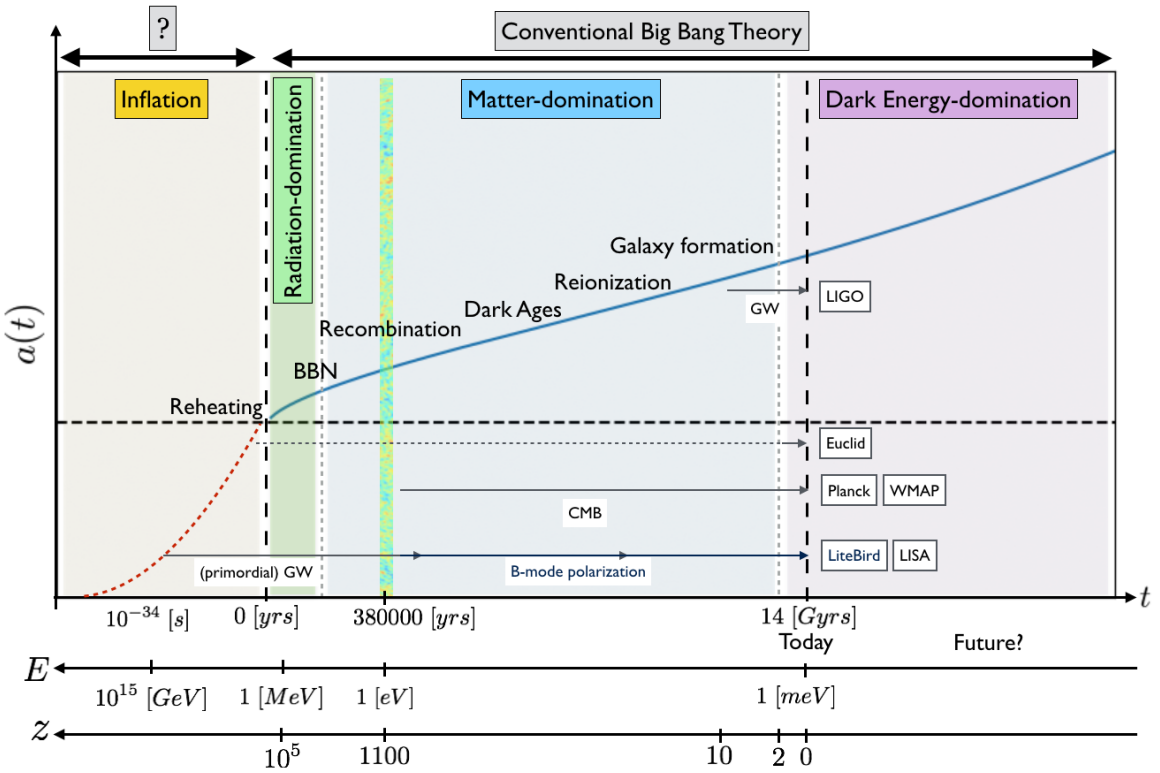


Fig. 2.1. Schematic Timeline of the Universe (evolution of the scale factor $a(t)$), highlighting key events in the history of the Universe, as a function of the cosmic time t, redshift z and energy scales E. A division is made to remark speculative epochs of the Universe (such as inflation), and the experimentally tested phases of the Conventional Big Bang Theory. All acronyms can be found in the Glossary. Numerical integration of $a(t)$ has been done with [2].

The re-written Friedmann equation (2.11) 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 (2.12) (see Figure 2.1). Knowing the values of the density parameters today (Ω_m, Ω_r and Ω_Λ), we can account for the expansion of the universe throughout its evolution. This difference in the scaling found in equation (2.10) shows how the universe progressed through different epochs:

$$\text{Radiation} \rightarrow \text{Matter} \rightarrow \text{Dark Energy} \tag{2.13}$$

If the Universe is currently expanding, going backwards in time means that the Universe was smaller, reaching a higher temperature and energy density. This thermal history is precisely what the conventional *Big Bang* theory aims to explain.

At very early stages, the universe behaved like a hot quantum gas, a thermal bath, formed by relativistic particles. It followed the laws of particle physics, which are able to explain the physics behind non-equilibrium processes. As a consequence of this expansion, the universe became cooler, allowing the decoupling of diverse 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, neutrinos decoupled from the cosmic thermal bath around $T \approx 10^{10}\ K$. Below that temperature, high-energetic photons (gamma-rays) are not energetic enough to produce electron-positron pairs, and the populations of both electrons and positrons started to decrease steeply by particle-antiparticle annihilation.

After the first 3 minutes, the early universe had reached the necessary low temperature to allow several light elements (Hydrogen, Helium and some Lithium) to be formed 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. 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 [26]. The primordial abundances predictions are one of the two main successes of the conventional Big Bang Theory.

When the universe was approximately 380000 years old, 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 with electrons. These photons could finally travel freely through space, with an almost isotropic distribution. This primordial radiation of photons is called the Cosmic Microwave Background (CMB). Its experimental detection by Penzias and Wilson [49] represents the second main success of the Big Bang Theory.

Up to this stage, we have been referring continuously only to Baryonic Matter (BM), which forms gas and dust. However, we know now that the universe is also composed by a different type of matter, the Dark Matter (DM), which only interacts gravitationally. It is assumed that DM decoupled at a very early moment in the history of the universe and started to collapse gravitationally into halos. Thus, after recombination, stars began to form when BM collapsed to the centre of the pre-existing DM halos. The first stars induced a period of re-ionization of the medium. Afterwards, due to gravitation, galaxies and clusters appeared, building what we call the Large Scale Structure (LSS) of the universe: a web-like structure where clusters are found in filaments.

Finally, according to the current SuperNovae (SN) data [56], it has been measured $\Omega_m \approx 0.3$ today. Thus, it has been concluded that the universe is currently dominated by Dark Energy (DE). This domination epoch started approximately 4 billions of years ago [19]. In cosmology, it is convenient to use as a time variable the so-called redshift. The cosmological redshift parameter, z, measures how stretched is the wavelength of a

photon due to the expansion of the universe, defined as

$$z + 1 = \frac{1}{a(t)} \tag{2.14}$$

Using equation (2.14), Dark Energy began to dominated around redshift $z \approx 2$.

2.3 Initial condition problems

As mentioned before, cosmological observations have accounted for experimental evidence of the BBN as well as the CMB radiation. Nevertheless, the detection of the CMB radiation also brought the biggest puzzles that modern cosmologists have to face. First of all, it confirmed that our universe is currently dominated by Dark Energy. Furthermore, it showed the necessity of fine-tuning the initial conditions of the Big Bang. In this section, we point out the two major issues related to this fine-tuning problem.

2.3.1 Horizon Problem

The CMB temperature has a blackbody spectrum. It shows an almost perfect isotropy where temperature anisotropies ΔT are very small, $\Delta T \sim 10^{-5}\ K$. This proves that at the moment the CMB radiation was emitted, the universe was in thermal equilibrium. At this temperature, Thomson scattering is the dominant type of interaction among photons. Therefore, photons could have reached equilibrium by means of Thomson scattering. However, we know that this was not the case.

Let us introduce the concept of *horizon*. It is the physical magnitude that gives the maximum distance that a photon can travel in our universe according General Relativity. Re-writing the metric given in equations (2.1) and (2.2) in terms of the denominated radial coordinate $d\chi = 1/\sqrt{1 - \kappa r^2}$, we find,

$$ds^2 = -dt^2 + a^2(t)[d\chi^2 + f(\chi)d\Omega^2], \tag{2.15}$$

where $f(\chi)$ is the corresponding function which accounts for the transformation of γ_{ij} due to this change. We define the comoving particle horizon, χ_P, as the distance a photon can travel radially ($d\Omega = 0$) between an initial time 0 and a time t,

$$\chi_P = \int_0^t \frac{dt'}{a(t')} = \int_0^a \frac{1}{aH} d\ln a, \tag{2.16}$$

where the factor $1/aH$ is denominated the *comoving Hubble radius*. Solving the integral (2.16) using the equation of state (2.9), we can find that the comoving particle

horizon for different epochs is,

$$\chi_P = \frac{2}{1+3w}\frac{1}{aH} = \begin{cases} a^{1/2} & \text{Matter-domination} \\ a & \text{Radiation-domination} \end{cases}. \tag{2.17}$$

Therefore, χ_P has always increased its size during Matter and Radiation domination. If we calculate its size at the moment of last scattering, we discover that the CMB radiation map (see Figure 2.2), is composed of $\sim 10^4$ disconnected patches [19]. Then, the following question arises: *how is it possible we observe photons not causally connected sharing the same temperature, if they could not interact via Thomson Scattering?* This is the **horizon problem**.

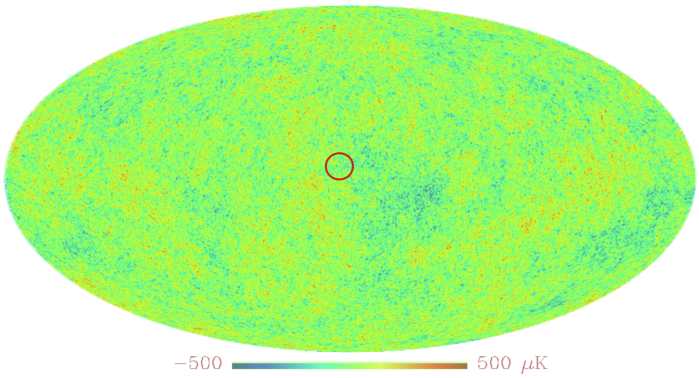


Fig. 2.2. Reconstruction of the CMB map using Planck and WMAP data showing temperature anisotropies [3], superimposed to the map with no anisotropies at $T \sim 2.7$ K. The small circumference is a representation of the comoving horizon χ, which indicates the maximum distance two photons can be apart from each other in order to be in causal contact.

2.3.2 Flatness Problem

Re-arranging equation (2.11) and evaluating it today, we obtain,

$$\sum_{i\in\{m,\Lambda,r\}} \Omega_i - 1 = \Omega_m + \Omega_\Lambda + \Omega_r - 1 = \frac{\kappa}{a^2 H_0^2}. \tag{2.18}$$

The analysis of the CMB radiation also gives us constraints on the values of the density parameters for different species. It shows that $\sum_i \Omega_i \approx 1$, and therefore, $\frac{\kappa}{a^2 H_0^2} \to 0$, meaning that our universe is flat today.

Having a flat universe today is not a problem *per se*. The problem is that, because we know how the density scales in time, in order to obtain a flat universe today, we needed to have a *flatter* universe in the past. We can estimate the fine-tuning needed in order to reach the current situation of flatness. In fact, according to the calculations,

we need to fine-tune, at least, at the moment of BBN:

$$|\sum_{i\in\{m,\Lambda,r\}} \Omega_i(a_{BBN}) - 1| \sim 10^{-16}. \tag{2.19}$$

This need of fine-tuning in the components of the universe is called the **flatness problem**.

2.4 Introduction to Inflation

As seen before, the conventional Big Bang theory cannot give us an answer to the previously mentioned problems, in spite of predicting successfully many other observations. Moreover, the Big Bang theory cannot give an explanation to the primordial inhomogeneities, which are traced in the CMB radiation as its tiny temperature anisotropies.

2.4.1 Inflation as a solution to the initial conditions problem

The goal to address is the following: *can we extend the conventional Big Bang Theory to solve the fine-tuning problems?* Observing both equations (2.16) and (2.18) (the latter evaluated at any arbitrary value of the scale factor), we see that both expressions depend on the comoving Hubble radius, $1/aH$. Moreover, from these two equations, we deduce that a decreasing comoving Hubble radius in a very early stage of our universe could solve both issues:

$$\frac{d}{dt}\left(\frac{1}{aH}\right) < 0. \tag{2.20}$$

The period in the history of our universe in which it underwent an accelerated expansion is called *inflation*.

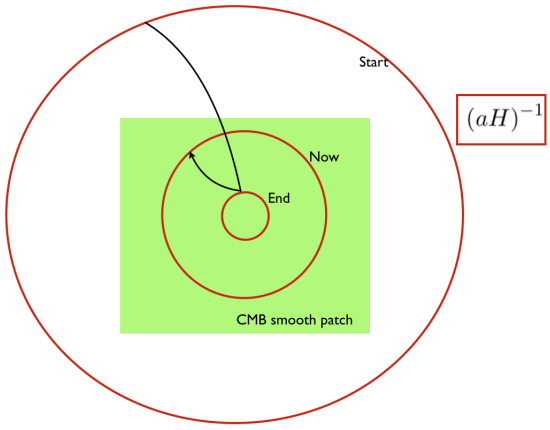


Fig. 2.3. Schematic representation of the solution to the horizon problem by means of introducing a decreasing comoving Hubble radius $1/aH$. Adapted from [19].

Let us review both the horizon and flatness problem individually. The key to understanding how inflation solves the horizon problem is related with the definition of the comoving Hubble radius, $1/aH$. The comoving Hubble radius points out the maximum distance two particles can be separated *today* to remain in causal contact. If the comoving Hubble radius was larger than expected early on, that would imply that χ_P is currently larger than $1/aH$. Therefore, and although the particles cannot communicate today, they could have been causally connected at earlier stages of the universe.

Using equations (2.6), (2.7) and (2.9), we observe how a decreasing comoving Hubble radius, $1/aH$, is equivalent to defining a period of accelerated expansion or a period where a fluid violating the strong energy condition dominates (negative pressure):

$$\frac{d}{dt}\left(\frac{1}{aH}\right) < 0 \quad \longleftrightarrow \quad \frac{d^2a}{dt^2} > 0 \quad \longleftrightarrow \quad \rho + 3p < 0 \quad \longleftrightarrow \quad \omega < -\frac{1}{3}. \tag{2.21}$$

Re-writing the first term of equation (2.21), we obtain

$$\epsilon \equiv -\frac{\dot{H}}{H^2} = -\frac{d\ln H}{dN} < 1, \tag{2.22}$$

where N is the *number of e-folds* (number of expansion times of the universe so that $dN = Hdt$), ϵ is usually called the *first slow-roll kinematic parameter*, which can be seen as a requirement for the Hubble parameter to change slowly (adiabatic theorem). If $\epsilon = 0$, it implies that H is constant and that inflation would be dominated by a never-ending exponential growth of the scale factor as described in equation (2.12) for $\omega = -1$. This is called *de Sitter* expansion. However, and in order to agree with cosmological observations, inflation has to end, consequently $0 < \epsilon < 1$.

Moreover, if we want inflation to solve the horizon and flatness problems, it also needs to last an enough period of time. For this reason, a second parameter is introduced. This parameter, η, studies the relative change of ϵ during one e-fold as follows

$$\eta \equiv -\frac{d\ln\epsilon}{dN} = -\frac{\dot{\epsilon}}{H\epsilon}, \tag{2.23}$$

and is known as the *second slow-roll kinematic parameter*. We require $0 < \eta < 1$ for inflation to last. To have slow-roll inflation, we require that $\{\epsilon, \eta\} \ll 1$.

2.4.2 Single Field inflation

We have concluded that we need a new substance, which violates the strong energy condition, in order to have inflation. For simplicity, let us model this substance as a single scalar field ϕ that is homogeneous, so that $\phi(t, \mathbf{x}) = \phi(t)$. To describe the dynamics of the system, the field ϕ follows the action,

$$S = \int d^4x\sqrt{-g}\left[\frac{M_P^2}{2}R - \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right], \tag{2.24}$$

where the first term of the action describes the coupling of the *inflaton* ϕ to gravity and $V(\phi)$ is the potential energy density that follows ϕ. Calculating $T_{\mu\nu}$ from the action (2.24) and comparing to the expression for a perfect fluid (2.5), we obtain the corresponding expressions for the density, ρ_ϕ, and the pressure, p_ϕ,

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V, \qquad p_\phi = \frac{1}{2}\dot{\phi}^2 - V. \tag{2.25}$$

Introducing expressions (2.25) in (2.6) and (2.8), we can obtain the analogous to the continuity equation as well as the Friedmann equations in terms of V and ϕ respectively:

$$\ddot{\phi} + 3H\dot{\phi} + \frac{dV}{d\phi} = 0, \qquad H^2 = \frac{1}{3M_P}\left(\frac{1}{2}\dot{\phi}^2 + V\right), \qquad \dot{H} = \frac{1}{M_P}\left(-\frac{1}{2}\dot{\phi}^2\right). \tag{2.26}$$

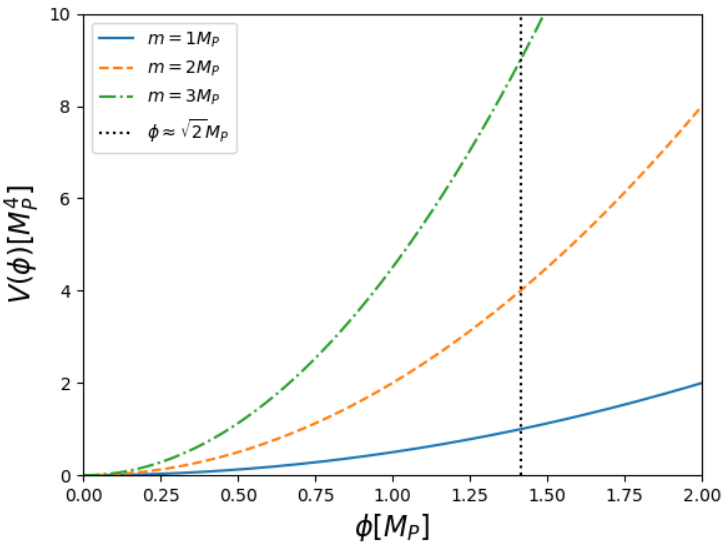


Fig. 2.4. Graphical representation of a quadratic potential $V = \frac{1}{2}m^2\phi^2$, for different values of m. For this particular potential, inflation ends when $\epsilon \approx 1$, which in this case corresponds to $\phi \approx \sqrt{2}M_P$.

Following the first equation in (2.26), in the slow-roll scenario ($\ddot{\phi} \approx 0$), we observe how inflation ends when the slope of the potential V is compensated by the friction term $3H\dot{\phi}$. This can be calculated in detail using the expressions for the slow-roll parameters $\tilde{\epsilon}$ and $\tilde{\eta}$:

$$\tilde{\epsilon} \equiv -\frac{\dot{\phi}^2}{2H^2} \qquad \tilde{\eta} \equiv -\frac{\ddot{\phi}}{H\dot{\phi}} = \epsilon - \frac{1}{2}\frac{\dot{\epsilon}}{\epsilon H} \approx \epsilon - \frac{1}{2}\eta. \tag{2.27}$$

In the limit when these parameters (2.27) tend to 1, we can calculate when inflation would end. In particular, the expression for ϵ can be used to quantify how long inflation needs to last to solve the horizon problem as a function of the number of e-folds

(approximately between 50-60 e-folds). Therefore, the arbitrary expression for V can be constrained as long as it solves the horizon problem (see one example of this potential in Figure 2.4).

2.5 Primordial perturbations

In the previous section, we mentioned that the temperature anisotropies present in the CMB radiation are very small. Still, we cannot ignore them and in fact, they do play a crucial role. They are the earliest imprint in the sky of those inhomogeneities in the young universe, which were the seeds of the current stellar and galactic structure. The accomplishment of inflation resides not only in solving the horizon and flatness problems, but in explaining the origin of these inhomogeneities.

2.5.1 Quantum fluctuations during inflation

The origin of these inhomogeneities lies in the quantum fluctuations of the inflaton with respect to the space-time background (encoded in the FLRW metric) and the field ϕ. In general lines, the overview is as follows. When the universe underwent inflation, quantum fluctuations were stretched to very large scale, accounting for the observed structure of the universe. These fluctuations remained frozen since they left the Hubble radius during this extreme expansion. During later stages of the universe, when the Hubble radius increased, they started to enter the horizon, becoming the initial density fluctuations of the universe.

To study the effect of these quantum fluctuations during inflation with respect to the homogeneous solution, we use a perturbed expression of the FLRW metric in (2.1), $\tilde{g}_{\mu\nu}(t,\mathbf{x}) \equiv g_{\mu\nu} + \delta g_{\mu\nu}(t,\mathbf{x})$, and of the field $\phi(t,\mathbf{x}) \equiv \phi(t) + \delta\phi(t,\mathbf{x})$. To parametrize $\delta g_{\mu\nu}$, we need to fix some freedom in the choice of coordinates. In the case of inflation, a particular choice of coordinates is usually set: the *comoving gauge*. This choice is defined as,

$$\delta\phi = 0, \qquad \delta g_{ij}(t,\mathbf{x}) = a^2\left[(1-2\mathcal{R}(t,\mathbf{x}))\delta_{ij} + h_{ij}(t,\mathbf{x})\right], \qquad \partial_i h_{ij} = 0, \tag{2.28}$$

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 ϕ is unperturbed (the density of the fluid is constant) and the scalar degrees of freedom are parametrized by $\mathcal{R}$. Substituting equation (2.28) into the expression of the action for a single field (2.24), and expanding in powers of $\mathcal{R}$, we find the quadratic action for the scalar metric fluctuation,

$$S_2 = \frac{1}{2}\int d^4x a^3 \frac{\dot{\phi}^2}{H^2}\left[-\dot{\mathcal{R}}^2 + \frac{(\partial_i\mathcal{R})^2}{a^2}\right]. \tag{2.29}$$

Our goal is to derive the equation of motion for $\mathcal{R}$, and show that it has a Simple Harmonic Oscillator (SHO) form. Posteriorly, we promote the classical field $\mathcal{R}$ to a

quantum operator and we quantize it. In order to solve the equations of motion derived from (2.29), the *Mukhanov-Sasaki variable* v is introduced [47]. It is defined as

$$v \equiv z\mathcal{R} \qquad\qquad z^2 \equiv 2a^2\epsilon. \tag{2.30}$$

Transitioning from action (2.29) using the conformal time τ, defined according to equation (2.4), and Mukhanov-Sasaki variable v, we obtain the full quadratic action in perturbations with canonical kinetic terms:

$$S_2 = \frac{1}{2}\int d\tau d^3x \left[(v')^2 + (\partial_i v)^2 + \frac{z''}{z}v^2\right]. \tag{2.31}$$

This action demonstrates that the *Mukhanov-Sasaki variable* v is the one we should use as the canonical quantization variable. In the Fourier space, we obtain the equation of motion for v:

$$v_{\mathbf{k}}'' + \left(k^2 - \frac{z''}{z}\right) v_{\mathbf{k}} = 0. \tag{2.32}$$

where $v_{\mathbf{k}}$ is the amplitude of the Fourier mode, and depends only on the magnitude of k. Equation (2.32) 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 $\hat{v}$,

$$\hat{v}_{\mathbf{k}} = v_{\mathbf{k}}\hat{a}_{\mathbf{k}} + v_{\mathbf{k}}^*\hat{a}_{\mathbf{k}}^\dagger, \tag{2.33}$$

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

$$[\hat{a}_{\mathbf{k}}, \hat{a}_{\mathbf{k}'}^\dagger] = -(2\pi)^3\delta(\mathbf{k} - \mathbf{k}'). \tag{2.34}$$

Solving equation (2.32) is complicated as z depends on the background dynamics. However, some insight can be obtained if we study certain limits and constraints. 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. In this limit, we choose the vacuum state for the fluctuation,

$$\hat{a}_{\mathbf{k}}\,|0\rangle = 0, \tag{2.35}$$

meaning that every mode k is assumed to have started its evolution in the vacuum state, such that there was not 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 [16]. This boundary condition is known as the *Bunch-Davies* vacuum.

In the *quasi-De Sitter* regime, where H and ϵ are approximately constant, the full solution to equation (2.32) is,

$$v_{\mathbf{k}}(\tau) = \frac{e^{-ik\tau}}{\sqrt{2k}}\left(1 - \frac{i}{k\tau}\right). \tag{2.36}$$

As mentioned in the beginning of the section, modes within the Hubble radius exited the horizon during inflation. In the literature, the modes with a wavelength smaller than the Hubble radius are called sub-horizon modes ($aH < k$), whereas modes whose wavelengths are bigger than the Hubble radius are known as super-horizon modes ($aH > k$). We can trace the history of a mode by studying the asymptotic limits of the solution (2.36) in the quasi-De Sitter case ($H, \epsilon \approx$ const):

- $|k\tau| \gg 1 \leftrightarrow k \gg aH$: the solution of $v_{\mathbf{k}}(\tau)$ is dominated by the oscillating exponential part $e^{-ik\tau}$. This is just the result of imposed the above-mentioned Bunch-Davis vacuum.
- $|k\tau| \ll 1 \leftrightarrow k \ll aH$: the dominant contribution to the solution is the divergent factor $1/\tau$. This means that, in this range, equation (2.36) is,

  $$\lim_{-\tau\to 0} v_{\mathbf{k}}(\tau) = \frac{1}{\sqrt{2}k^{3/2}\tau} \propto \frac{aH}{k^{3/2}} \tag{2.37}$$

  where we have used that $\tau \sim (aH)^{-1}$. Recovering the definition of v from equation (2.30), we find that the curvature perturbation $\mathcal{R}$ is constant:

  $$\mathcal{R} = \frac{v}{z} = \frac{v}{a\sqrt{2\epsilon}} \propto \frac{H}{k^{2/3}\sqrt{\epsilon}} \to \lim_{-\tau\to 0} \dot{\mathcal{R}} = 0, \tag{2.38}$$

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

In conclusion, during inflation, quantum fluctuations in the inflaton field led to some parts of the universe being stretched to for a larger period of time than others. These sub-horizon scales exited the horizon due to the accelerated expansion of the universe, becoming super-horizon, where they remained frozen until they re-enter the horizon at later stages of the universe (see Figure 4.1). The before mentioned local time delay for inflation translated into density fluctuations, which follow their evolution during radiation-matter domination. These local differences in density became the CMB temperature anisotropies, which seeded all the current structure of the universe.

2.5.2 Spectrum and bispectrum

As described in the previous subsection, we can relate the CMB temperature anisotropies with perturbations of the field $\mathcal{R}$. In order to characterize these perturbations, it is convenient to define the power spectrum, $P_{\mathcal{R}}(k)$, and the bispectrum, $B_{\mathcal{R}}(k)$. They

are the Fourier transforms of the real-space two-point and three-point correlation functions respectively. To study the probability distribution of these perturbations in our universe, we usually average the ensemble of all possible $\mathcal{R}$. The corresponding expressions are,

$$\langle \mathcal{R}(\mathbf{k}_1)\mathcal{R}(\mathbf{k}_2)\rangle \equiv (2\pi)^3\delta^3(\mathbf{k}_1+\mathbf{k}_2)P_{\mathcal{R}}(k), \tag{2.39}$$

$$\langle \mathcal{R}(\mathbf{k}_1)\mathcal{R}(\mathbf{k}_2)\mathcal{R}(\mathbf{k}_3)\rangle \equiv (2\pi)^3\delta^3(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)B_{\mathcal{R}}(k) \tag{2.40}$$

where both $P_{\mathcal{R}}(k)$ and $B_{\mathcal{R}}(k)$ depend only on the magnitude of the momentum scale $k = |\mathbf{k}|$, and not on the direction.

- The **power spectrum**, $P_{\mathcal{R}}(k)$, is usually redefined as a dimensionless quantity, $\mathcal{P}(k)$, as

$$\mathcal{P}_{\mathcal{R}}(k) = \frac{k^3}{2\pi^2}P_{\mathcal{R}}(k). \tag{2.41}$$

Assuming slow-roll canonical single field inflation; that is, using equation (2.38), $\mathcal{P}(k)$ takes the form

$$\mathcal{P}_{\mathcal{R}}(k) = \frac{1}{8\pi^2}\frac{H_*^2}{\epsilon}, \tag{2.42}$$

where H_* is the Hubble parameter evaluated at horizon crossing $k = aH$. From the observational point of view, a phenomenological parametrization of $\mathcal{P}_{\mathcal{R}}(k)$ is used in terms of a power law. Up to the first order the expression is

$$\mathcal{P}_{\mathcal{R}}(k) = A_s\left(\frac{k}{k_*}\right)^{n_s-1}, \tag{2.43}$$

where k_* is a pivot scale, A_s is the scalar amplitude and n_s is called the *spectral index*. The spectral index quantifies the scale dependence of $\mathcal{P}_{\mathcal{R}}(k)$. The dimensionless power spectrum is expected to be *nearly scale-invariant*. During slow-roll inflation, H and ϵ are not exactly constant. This time-dependence of the Hubble parameter (during inflation, the Hubble parameter decreases as ϵ increases) induces a slightly increment of $\mathcal{P}_{\mathcal{R}}(k)$ for low values of k. In the canonical slow roll inflationary regime, n_s is related to the slow roll kinetic parameters as

$$n_s = 1 - 2\epsilon - \eta. \tag{2.44}$$

We know that ϵ and η are expected to be small, which indicates that n_s should slightly deviate from one. If we assume that the distribution of the density perturbations is Gaussian, the power spectrum contains all information to characterize it. Moreover, the parametrization (2.43) is also used for the perturbations of the metric (tensor modes),

$$\mathcal{P}_t(k) = A_t\left(\frac{k}{k_*}\right)^{n_t}, \tag{2.45}$$

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*:

$$r = \frac{A_t}{A_s}. \tag{2.46}$$

- The **bispectrum**, $\mathcal{B}_\mathcal{R}(k)$, quantifies the deviations from Gaussianity of the distribution of density perturbations. The bispectrum is evaluated on vectors $\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3$. These vectors are constrained under the triangular inequality. The full bispectrum for a canonical single-field, slow-roll inflation was calculated in [45]:

$$\mathcal{B}_\mathcal{R}(k_1,k_2,k_3) = \frac{(2\pi)^4 {\mathcal{P}_\mathcal{R}}^2}{(k_1k_2k_3)^2}\frac{1}{8k_1k_2k_3}\left\{(3\epsilon-2\eta)\sum_{i=1}^{3}k_i^3+\epsilon\sum_{i\neq j=1}^{3}k_ik_j^2+\epsilon\frac{4}{\sum_{i=1}^{3}k_i}\sum_{i=1}^{3}k_i^2k_j^2\right\}. \tag{2.47}$$

where $k_i = |\boldsymbol{k}_i|$. Regarding its shape, $\mathcal{B}_\mathcal{R}(k)$ is classified as,

$$\begin{cases} \text{squeezed} & \text{if } k_3 \ll k_1 \sim k_2 \\ \text{equilateral} & \text{if } k_3 \sim k_1 \sim k_2 \\ \text{flattened} & \text{if } k_3 \sim k_1 + k_2 \end{cases}. \tag{2.48}$$

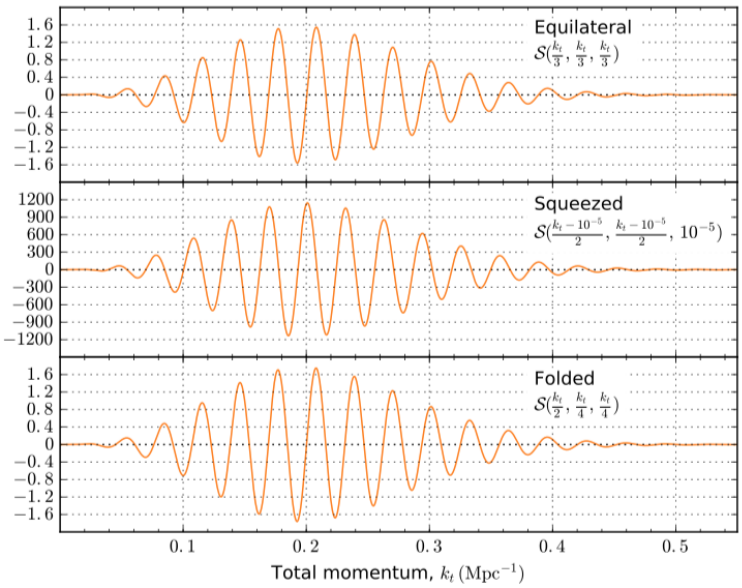


Fig. 2.5. Oscillations in the normalized bispectrum $\mathcal{S}$, related to $\mathcal{B}_\mathcal{R}(k)$ as $\mathcal{B}_\mathcal{R}(k) \equiv (2\pi)^4 A_s^2 M_P^6 (k_1k_2k_3)^{-2}\mathcal{S}$, corresponding to non-gaussianities due to a reduction in the speed of sound (see chapter 3 and appendix C). Notice the oscillatory shape, as well as the $\pi/2$ phase difference between the squeezed and the equilateral and flattened (folded) shapes. [58].

2.6 Cosmological Standard Model

The cosmological observations obtained by the WMAP and Planck missions, which have studied the anisotropies of the CMB, have led to the conclusion that we live in a spatially flat universe. It contains less than 5% of Baryonic Matter (BM), approximately 26% of Cold Dark Matter (CDM) and almost 70% of Dark Energy (DE), making the percentage of radiation negligible today.

The simplest cosmological model that takes into account the predictions made by conventional Big Bang theory and inflation is known as the *Standard Model of Cosmology* or the ΛCDM model. To explain most of the current cosmological observations, this model requires only six free parameters, whose values have been obtained with high accuracy (see in Table 2.1). The procedure and techniques to obtain the values of these parameters using cosmological data coming from the CMB are explained in detail in chapter 4.

Symbol	**Description**	**Value**
$\Omega_b h^2$	density parameter of BM	0.02225 ± 0.00016
$\Omega_c h^2$	density parameter of CDM	0.1198 ± 0.0015
$100\theta_{MC}$	measure of the sound horizon at last scattering	1.04077 ± 0.00032
τ_{reion}	Thomson scattering optical depth due to reionization	0.079 ± 0.017
$\ln(10^{10} A_s)$	scalar amplitude of equation (2.43)	3.094 ± 0.034
n_s	spectral index of equation (2.43)	0.9645 ± 0.0049

Table 2.1. Mean values of the ΛCDM cosmological parameters and their corresponding uncertainties, obtained by statistical inference using CosmoMC and Planck temperature and polarization data (see chapter 4) [51]. The parameter h represents the reduced Hubble constant defined as $h \equiv H_0/(100\, km\, s^{-1} Mpc^{-1})$. 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 $k_* = 0.005\ Mpc^{-1}$.

Chapter 3

Beyond the simplest model of inflation

> "One particular question [...] is whether or not there are *features*, or localized deviations from scale invariance, in the primordial power spectrum"
>
> *J. Chluba, J. Hamann, S. Patil [25]*

3.1 Introduction

In Chapter 2, we have shown how our universe is well described by the ΛCDM model, which is supported by current cosmological observations (see Table 2.1, [51]). We have also pointed out how the inflation paradigm provides a suitable fundamental physical explanation for the origin of the current structure of our universe. In fact, we have defined two of these 6 parameters in the ΛCDM model, the scalar spectral index n_s and the scalar amplitude of the primordial power spectrum A_s, which are in direct connection to inflation.

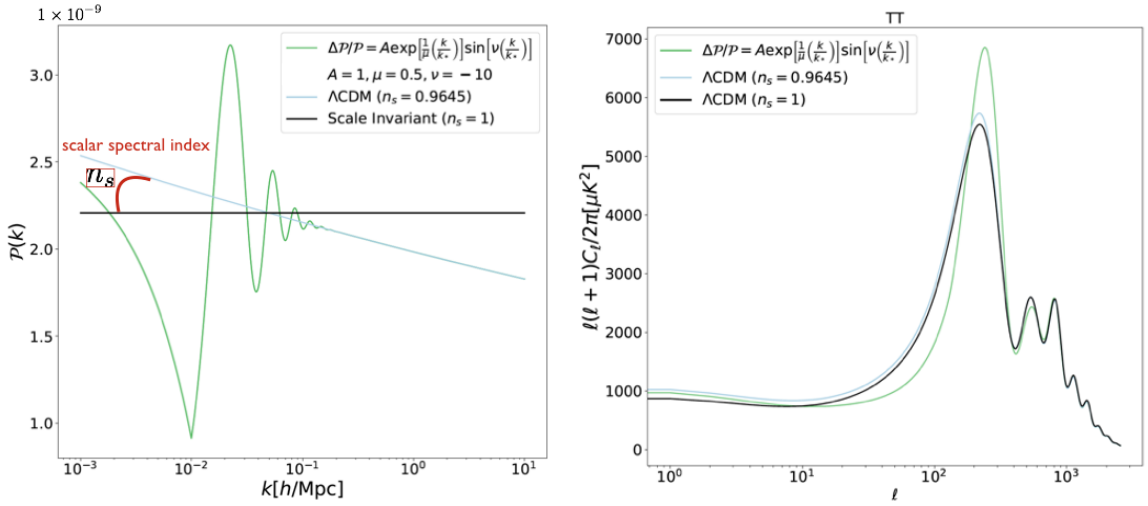
One may argue that statistical tools can be used on the current cosmological data to select the best cosmological model based on an alternative model of inflation, which predicts deviations for any of these inflationary ΛCDM parameters. Thus, the question seems straightforward: *is there any other way of handling the data to increase our skills to distinguish between inflationary models or to finally support the inflationary hypothesis?* [25]. In order to answer this question, we need to study the signatures beyond the canonical single field inflation scenario. Canonical single field inflation is defined as [16]:

- single field slow roll inflation: there is only one degree of freedom controlling the dynamics and the slow-roll parameters defined in chapter 2, which fulfil $\epsilon, \eta \ll 1$.
- the inflaton has canonical kinetic terms with speed of sound $c_s = 1$ in natural units.
- it is minimally coupled to gravity.
- it follows Bunch-Davies initial condition, as described in chapter 2. The sub-horizon modes experienced Mikowski space-time.

One possible starting point is to study deviations in the power-law **Primordial Power Spectrum** (PPS), expressed in equation (2.42). Usually, deviations of the usual power law spectrum are encoded in higher powers of the Taylor expansion of equation (2.42) in terms of $\log(k/k_*)$, such as the parameter α_s:

$$\mathcal{P}_{\mathcal{R}} = A_s \left(\frac{k}{k_*}\right)^{n_s - 1 + \frac{1}{2}\alpha_s \ln(k/k_*)}, \tag{3.1}$$

where α_s is called *spectral running index*, and in the case of slow-roll inflation is expected to be very small. Other inflationary scenarios, as we will see below, predict more complex changes in the PPS (see figure 3.1, where a toy model for a modification of the power-law PPS, $\Delta\mathcal{P}_{\mathcal{R}}/\mathcal{P}_{\mathcal{R}0}$, is shown). These localized deviations from the power-law spectrum, as well as in higher-order correlation functions, are usually known as *features*. The PPS is intimately related with the temperature power spectrum of the CMB[1]). However, most inflationary model predictions at the level of the PPS tend not to be statistically significant enough, which may complicate the analysis of the data to support one inflationary model against other [28].



(a) Primordial Power Spectrum (b) CMB Temperature Power Spectrum

Fig. 3.1. The PPS is the basis for the CMB power spectra: changes with respect the power-tilt law induce noticeable variations in the CMB temperature power spectra. The (localized) deviations from the PPS power-law, $\Delta\mathcal{P}_{\mathcal{R}}/\mathcal{P}_{\mathcal{R}}$, are commonly called *features*. Example created using the development version of CAMB (see Appendix B) and a toy-model from *M. Martinelli* [39].

The canonical single field scenario of inflation states that fluctuations of the inflaton field on large scales correspond to a local shift backwards or forwards in the trajectory of the homogeneous background field. In chapter 2, we argued how these shifts along the trajectory affect the total density in different parts of the universe after inflation; however, it cannot affect the relative number density among the several components: the perturbations are adiabatic, defined by the curvature perturbations $\mathcal{R}$, meaning that the perturbations of all the universe's components came from the same curvature

[1]Further information about how the CMB power spectrum is modelled can be found in chapter 4.

perturbation. If this **adiabaticity** condition is somehow violated, it may point out to different inflationary models.

Finally, in chapter 2, we mentioned that if the curvature perturbation $\mathcal{R}$ is Gaussian, the distribution of the perturbations is fully characterized by the PPS, $\mathcal{P}_{\mathcal{R}}$. Nevertheless, violations of any of the canonical single field conditions detailed before can generate non-gaussian signals. Thus, **non-gaussianity** signatures can be encoded in higher correlation functions such as the bispectrum $\mathcal{B}_{\mathcal{R}}$. Constraints on non-gaussianity are difficult to quantify, but there are mainly four templates for the three-point correlation function which have been historically studied [38].

- One of them describes non-gaussianity by using the following parametrization [19]:
$$\mathcal{R}(x) = \mathcal{R}_g(x) + \frac{3}{5} f_{NL} \mathcal{R}_g^2(x), \tag{3.2}$$
where $\mathcal{R}_g$ is the Gaussian comoving curvature perturbation and f_{NL} controls the skewness of the probability density function. This is a *local* configuration that peaks in the *squeezed* limit in the momentum space (see equation (2.47)). In this limit, the amplitude of the bispectrum is proportional to the tilt of the PPS [45]. This pattern is often produced in multi-field models.
- Another template studies the *equilateral* shape (see again equation (2.47)), for which the non-gaussianity signal is generated by single field inflation models with non-canonical kinetic terms ($c_s \neq 1$), and in models with higher-derivative interactions in the Lagrangian (such as in effective field theories).
- The third template is the so-called *folded* where the signal peaks for *flat* triangles (see equation (2.47)). These non-gaussian signals are produced in single field models initially lying in a non-Bunch-Davies vacuum and in models with higher-derivative interactions.
- There exists a fourth one, known as *orthogonal*, which distinguishes among variants of non-canonical kinetic terms.

3.2 Effective Field Theory of Inflation

The bispectrum, as well as higher correlation functions, can be used to study models beyond the canonical single field scenario. The simplest approach to generate a three-point correlation function, the bispectrum, consists of writing an explicit third order interaction in the Lagrangian. For that, we are going to follow an Effective Field Theory (EFT) methodology to characterize the theory of fluctuations around an inflating cosmological background. This approach is an approximation of a more fundamental theory, based on describing a system through the appropriate degrees of freedom associated to the underlying symmetries. It has been widely used in different fields such as particle physics or condensed matter.

Following EFT of inflationary perturbations [24], we can construct an effective action for the Goldstone boson of time diffeomorphisms $\pi(t, \mathbf{x})$, which is related to

the curvature perturbation $\mathcal{R}$ through the relation $\mathcal{R} = -H\pi$ up to linear order (see appendix C). The effective single field action up to third order, neglecting higher order slow-roll corrections ($\sim \mathcal{O}(\epsilon^2)$) and assuming $\dot{\pi}^3$ to be small and approximately constant is given by,

$$S_2 = \int d^4x a^3 M_P^2 \epsilon H^2 \left[-\frac{\dot{\pi}^2}{c_s^2} + \frac{(\partial_i \pi)^2}{a^2} \right]; \tag{3.3}$$

$$S_3 = \int d^4x a^3 M_P^2 \epsilon H^2 \left[-2Hsc_s^{-2}\pi\dot{\pi}^2 - (1 - c_s^{-2})\dot{\pi}\left(\frac{\dot{\pi}^2}{c_s^2} - \frac{(\partial_i \pi)^2}{a^2} \right) \right], \tag{3.4}$$

where the background information is contained in $\epsilon(t)$ and $c_s(t)$, and $s(t)$ parametrizes the change in the speed of sound $c_s(t)$ such as

$$s \equiv \frac{\dot{c_s}}{(c_s H)}. \tag{3.5}$$

The physical details of the theory are encoded in the speed of sound c_s, and in its corresponding change rate according to s. Note also that equation (3.3) reproduces the action for $\mathcal{R}$ (2.28) in the single field scenario.

This last fact is important. Single field inflation is the simplest theoretical model of inflation, which is able to fit the cosmological data within ΛCDM. Overall, it becomes a suitable and convenient model. Nevertheless, further unification theories have extra degrees of freedom at higher energies, and therefore, may suggest that inflation could be driven not by one single field but by multiple ones. These extra degrees of freedom will cause deviations from the standard canonical single field scenario: features in the primordial power spectrum or in the bispectrum (non-gaussianity) and non-adiabatic perturbations. Detecting any of these signatures will add information about the physics at the very early stages of our universe. Thus, EFT is crucial in this aspect, as it may reconcile the cosmological data with the *unlikely* possibility of having just one degree of freedom.

We now explain how an effective theory for curvature perturbation $\mathcal{R}$, when a strong turn in the inflationary trajectory supported by a heavy field $\mathcal{F}$ with "effective mass"[2] M_{eff}, gives interesting signatures in cosmological observables [10, 23, 17]. In this case, the curvature perturbation is kinetically coupled to the heavy field. We will not discuss the details of the procedure, but instead, we will draw the most important steps when we analyze the evolution of the fields $\mathcal{R}$ and $\mathcal{F}$ when the trajectory is turning at rate[3] Ω. The details of the procedure can be found in appendix C. The goal is to stress that the turning in the trajectory causes a reduction in the speed of sound c_s, which produces a characteristic pattern in the PPS and in the bispectrum, as we will see in section 3.3.

[2]The term "effective mass" is used for historical reasons but it can be misleading, as it is not a proper mass.

[3]Do not mix up the density parameter Ω_i, introduced in chapter 2, with this angular rate Ω

The equations for $\mathcal{R}$ and $\mathcal{F}$ as a function of time form a coupled system of equations whose solutions are expressed in terms of a harmonic expansion, defined in terms of two frequencies ω_+ and ω_-. The values of the frequencies depend on the wave number of the mode k. We can differentiate between two regimes:

- $k/a \gg M_{eff}$: both fields behave as massless and oscillate with frequencies proportional to k/a.
- $M_{eff} \gg k/a \gg H$: the modes are no longer degenerated and the frequencies become $\omega_- \sim k/a$ and $\omega_+ \sim M_{eff}$. In this regime, the effective action for $\mathcal{R}$ can be obtained by integrating out the heavy field $\mathcal{F}$ (see appendix C).

  This effective action is similar to the EFT of inflation equation (3.3), where $\mathcal{R} = \pi H$, but the speed of sound c_s of the adiabatic perturbation $\mathcal{R}$ is given as,

$$c_s^{-2} = 1 + \frac{4\Omega^2}{k^2/a^2 + M_{eff}}, \tag{3.6}$$

  which implies that c_s is related to the angular velocity Ω when there is a turn in the inflationary trajectory, inducing a momentary reduction on the speed of sound c_s.

3.3 Transient Reductions of the Speed of Sound

We have seen that, when the heavy field can be consistently integrated out, we obtain an effectively-single field theory with a variable speed of sound $c_s(t)$. The effect of this variable speed of sound $c_S(t)$ can be seen in the primordial power spectrum $\mathcal{P}_\mathcal{R}$, in the bispectrum $\mathcal{B}_\mathcal{R}$, etc. In particular, transient variations of c_s produce localized oscillatory and correlated features in both $\mathcal{P}_\mathcal{R}$ and $\mathcal{B}_\mathcal{R}$ [13, 15, 25, 58].

In this section, we focus on this transient, moderately sharp and mild reduction in the speed of sound c_s. Concretely, in this case, the effects coming from this reduction of the speed of sound are larger than the slow-roll corrections but small enough to be studied at linear order. This transient reduction in c_s induces corresponding effects on $\mathcal{P}_\mathcal{R}$ and $\mathcal{B}_\mathcal{R}$, which can calculated using the *in-in* formalism. This approach is a generalization of the path integral formalism of quantum field theory [37, 61].

Within this approach, the change in the primordial power spectrum $\Delta\mathcal{P}_\mathcal{R}$ is given by the Fourier transform of the reduction in c_s, and the bispectrum can be calculated up to leading order as a function of the corrected power spectrum $\Delta\mathcal{P}_\mathcal{R}/\mathcal{P}_{\mathcal{R}0}$. This approximation was called, following *Achúcarro et al.*, the *Slow-Roll Fourier Transformation* (SRFT) method [11].

To calculate the modifications in the power spectrum $\Delta\mathcal{P}_\mathcal{R}/\mathcal{P}_{\mathcal{R}0}$, we divide the quadratic action of EFT of inflation, equation (3.3), into a free part (resembling single field inflation) and a small perturbation:

$$S_2 = S_0 + S_{pert} = \\ = \underbrace{\int d^4x a^3 M_P^2 \epsilon H^2 \left[\dot{\pi}^2 - \frac{(\partial_i \pi)^2}{a^2}\right]}_{S_0} - \underbrace{\int d^4x a^3 M_P^2 \epsilon H^2 \left[\dot{\pi}^2 (1 - c_s^{-2})\right]}_{S_{pert}}. \quad (3.7)$$

Transitioning to the conformal time, and using the mentioned *in-in* formalism τ and the following definition of the variable u,

$$u(\tau) \equiv (1 - c_s^{-2}(\tau)), \quad (3.8)$$

we calculate the change in the primordial power spectrum $\Delta\mathcal{P}_\mathcal{R}$ as the Fourier transform of the reduction in the speed of sound c_s:

$$\frac{\Delta\mathcal{P}_\mathcal{R}}{\mathcal{P}_{\mathcal{R}0}} = k \int_{-\infty}^{0} d\tau u(\tau) \sin(2k\tau), \quad (3.9)$$

where $\mathcal{P}_{\mathcal{R}0}$ is the single field inflation featureless power spectrum with $c_s = 1$ defined by equation (2.41).

For the case of the bispectrum, to leading order in u and s, we can also calculate the change $\Delta\mathcal{B}_\mathcal{R}$ using again the in-in formalism:

$$\Delta\mathcal{B}_\mathcal{R} = \frac{(2\pi)^4 \mathcal{P}_{\mathcal{R}0}^2}{(k_1 k_2 k_3)^2} \Bigg\{ -\frac{2}{3}\frac{k_1 k_2}{k_3}\left[\frac{1}{2k}\left(1 + \frac{k_3}{2k}\right)\frac{\Delta\mathcal{P}_\mathcal{R}}{\mathcal{P}_{\mathcal{R}0}} - \frac{k_3}{4k^2}\frac{d}{d\log k}\frac{\Delta\mathcal{P}_\mathcal{R}}{\mathcal{P}_{\mathcal{R}0}}\right] + 2 \text{ perm} + \\ + \frac{1}{4}\frac{k_1^2 + k_2^2 + k_3^2}{k_1 k_2 k_3}\Bigg[\frac{1}{2k}\left(4k^2 - k_1 k_2 - k_2 k_3 - k_3 k_1 - \frac{k_1 k_2 k_3}{2k}\right)\frac{\Delta\mathcal{P}_\mathcal{R}}{\mathcal{P}_{\mathcal{R}0}} \\ - \frac{k_1 k_2 + k_2 k_3 + k_3 k_1}{2k}\frac{d}{d\log k}\left(\frac{\Delta\mathcal{P}_\mathcal{R}}{\mathcal{P}_{\mathcal{R}0}}\right) + \frac{k_1 k_2 k_3}{4k^2}\frac{d^2}{d\log k^2}\left(\frac{\Delta\mathcal{P}_\mathcal{R}}{\mathcal{P}_{\mathcal{R}0}}\right)\Bigg]\Bigg\}\Bigg|_{k=\frac{1}{2}\sum_i^3 k_i} \quad (3.10)$$

where $k_i = |\boldsymbol{k}_i|$ and $k \equiv (k_1 + k_2 + k_3)/2$, and $\Delta\mathcal{P}_\mathcal{R}/\mathcal{P}_{\mathcal{R}0}$ and its corresponding derivatives are evaluated at k. We can observe how, in the squeezed limit set by equation (2.47), equation (3.10) resembles the single field consistency relation (2.45). Moreover, we also see how the bispectrum expression is written in terms of the features of the scalar power spectrum, $\Delta\mathcal{P}_\mathcal{R}/\mathcal{P}_{\mathcal{R}0}$, correlating the primordial power spectrum to the bispectrum.

The action described in equations (3.3) and (3.4) is perturbative in terms of the speed of sound c_s. It implies that the reduction in the speed of sound, c_s, cannot be too big (condition given by $u \ll 1$) nor too fast (condition given by $s \ll 1$). Also, we have already mentioned that the slow-roll regime is kept throughout, and therefore, the contributions of the slow-roll corrections ϵ, η have to be smaller than those of the variable speed of sound c_s. We need to impose these conditions for all values of τ, but it is enough to restrict ourselves at the point τ_m where u and s take their maximum

value $[u(\tau_m), s(\tau_m)] = [u_{max}, s_{max}]$. In short:

$$\epsilon, \eta \ll \max(u, s) \ll 1. \tag{3.11}$$

In chapter 5, we search for this particular set of features due to a transient reduction of the speed of sound in the CMB, after reviewing how to explore the cosmological data thanks to many analysis tools in Chapter 4.

Chapter 4

Data Analysis Tools

> "Creating the tools for
> computer-assisted calculations,
> *programming* or *coding*,
> is (theoretical) physics research in
> the same way as algebraic
> calculations are"
>
> *Jesús Torrado* [60]

In chapters 2 and 3, we have detailed the theoretical background of the standard cosmological model, the framework for the simplest inflationary theory and the extension of the model that we aim to study further.

In order to carry out our research, it is essential to understand and master the data analysis techniques and the necessary computational tools. For this reason, this chapter aims to be self-contained to be used in the future as a reference guide for the basics of cosmological analysis. First, it is explained how to connect inflation and predictions to the Cosmic Microwave Background (CMB) data [19, 27]. Second, we will provide a brief introduction to Bayesian Statistics [55], Machine Learning techniques [54] and tools used in the data analysis of this research project.

In this chapter, we use the notation $\langle f \rangle$ to indicate the average over the ensemble.

4.1 Cosmic Microwave Background

4.1.1 Connection to inflation

In chapter 2, we have mentioned that the power spectrum of density perturbations can be modelled as a power law. The connection between inflation and the current cosmological observations is mainly done by the study of the anisotropies of the CMB and the LSS of our universe. In this thesis, we will focus on the connection between the CMB spectra and the primordial power spectrum. The Primordial Power Spectrum (PPS) (2.43), which encodes the distribution of the comoving curvature perturbation, $\mathcal{R}$, remains constant when it exits the horizon during inflation (see Figure 4.1). When $\mathcal{R}$ re-enters, it starts to evolve according to the physical laws that describe radiation and matter domination. Therefore, we need to know the time evolution of the comoving density curvature perturbation $\mathcal{R}$, so that we can relate it to a cosmological observable $\mathcal{Q}_{\mathbf{k}}(\tau)$ (for instance, the temperature anisotropies of the CMB [19]). Schematically:

$$\mathcal{Q}_{\mathbf{k}}(\tau) \sim \Delta T_{\mathcal{Q}}(k, \tau, \tau_*)\mathcal{R}_{\mathbf{k}}(\tau_*), \tag{4.1}$$

where $\Delta T_{\mathcal{Q}}$ is the so-called *transfer function* between $\mathcal{R}$, at conformal time of horizon crossing τ_*, and a cosmological observable $\mathcal{Q}$, at a later time τ.

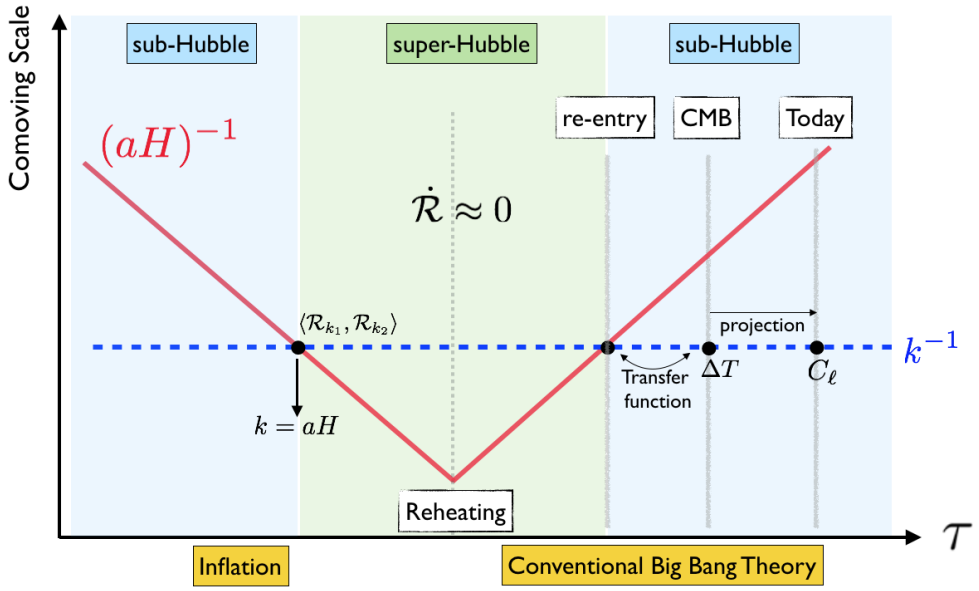


Fig. 4.1. Graphical representation of the history of a particular comoving scale k^{-1} (dashed blue line) leaving the comoving Hubble radius $1/aH$ due to inflation, remaining constant when it is superhorizon, and re-entering during the conventional Big Bang period. The evolution of k^{-1} after re-entering is encoded in the transfer function $\Delta T_{\mathcal{Q}}$, which can be later related to cosmological observations such as the angular power spectrum C_ℓ of the temperature anisotropies of the CMB. Adapted from [19].

How do we relate $\mathcal{Q}_{\boldsymbol{k}}$ to the anisotropies of the CMB map? The procedure consists on decomposing these temperature anisotropies[1], $\Delta T/T_0$, using a harmonic expansion of the map:

$$\frac{\Delta T}{T_0} = \sum_{\ell m} a_{\ell m} Y_{\ell m}, \tag{4.2}$$

where $T_0 \approx 2.7\ K$ is the background temperature of the CMB, $Y_{\ell m}$ are the spherical harmonic basis on a 2-sphere, and $a_{\ell m}$ are the coefficients of the decomposition (usually denominated multipole moments), defined as,

$$a_{\ell m} = \int d\Omega Y^*_{\ell m} \frac{\Delta T}{T_0}, \tag{4.3}$$

with $d\Omega$ being the angular differential in spherical coordinates, $\ell = 0, 1, 2$ corresponding to the monopole, dipole and quadrupole respectively, and $m = -\ell, ..., +\ell$. The multipole ℓ is proportional to the scale k (described in chapter 2) and inverse to physical distance

[1]We have maintained the notation used in [19] for the temperature and the transfer function. We should not confuse the transfer function, $\Delta T_{\mathcal{Q}}$, with the differences in temperature ΔT.

scales (low values of ℓ correspond to two points on the CMB radiation map that are very separated). We can use the information encoded in equation (4.3) to calculate the corresponding angular power spectrum of the CMB temperature map:

$$C_\ell^{TT} = \frac{1}{2\ell+1}\sum_m \langle a_{lm}^*, a_{lm}\rangle \,. \tag{4.4}$$

For low values of ℓ, there is an intrinsic limit related with the accuracy of C_ℓ. This is due to the few possible measurements at large physical distances on the CMB map: the *Cosmic Variance* (see figure 4.2 and observe the large error bars until $\ell \approx 200$).

The angular power spectrum (4.4), obtained experimentally by the Planck mission, can be observed in figure 4.2, together with the theoretical prediction from the standard model of cosmology. However, *how do we establish the relation with ΛCDM in order to obtain the theoretical prediction?* Recovering our original question expressed mathematically in equation (4.1), we find the analogy between $a_{\ell m}$ and $\mathcal{Q}$ with the transfer function ΔT_T as

$$a_{\ell m} \propto \int d^3k (2\pi)^{-3} \Delta T_\ell^T(k) \mathcal{R}(k)_k Y_{\ell m}, \tag{4.5}$$

where the transfer function ΔT_T encodes the physical effects on the evolution of the temperature. Substituting equation (4.5) into (4.4) we obtain,

$$C_\ell^{TT} = \frac{2}{\pi}\int k^2 dk P(k) \Delta T_\ell^T(k) \Delta T_\ell^T(k), \tag{4.6}$$

which is the expression for the theoretical prediction of C_ℓ^{TT}. In general, the CMB background provides more information (more cosmological observables $\mathcal{Q}$) than the temperature anisotropies. In fact, in the CMB map we can also study polarization modes E and B (see next subsection). Thus, the general expression for the angular power spectra of the CMB is

$$C_\ell^{XY} = \frac{2}{\pi}\int k^2 dk P(k) \Delta T_\ell^X(k) \Delta T_\ell^Y(k), \tag{4.7}$$

where X and Y may refer either to the temperature T or the polarization modes E and B. The Planck mission has measured successfully C_ℓ^{TT}, C_ℓ^{TE} and C_ℓ^{EE}. Therefore, the connection with the ΛCDM model is obtained computing theoretically the expression for the transfer function ΔT, which is usually written as a line-of-sight integration that contains the source factor $S^X(k,\tau)$ and the geometric projection $P_\ell^X(k|\tau_0 - \tau|)$ based on Bessel functions:

$$\Delta T_\ell^X(k) = \int_0^{\tau_0} d\tau S^X(k,\tau) P_\ell^X(k|\tau_0 - \tau|). \tag{4.8}$$

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 and CAMB, which will be briefly reviewed in subsection 4.1.3. To understand how these

codes work, we first need to understand the physics of recombination as well as some concepts about polarization.

4.1.2 Understanding the CMB power spectra

This section aims at explaining very briefly the main characteristics of the CMB temperature and polarization spectra.

Temperature anisotropies

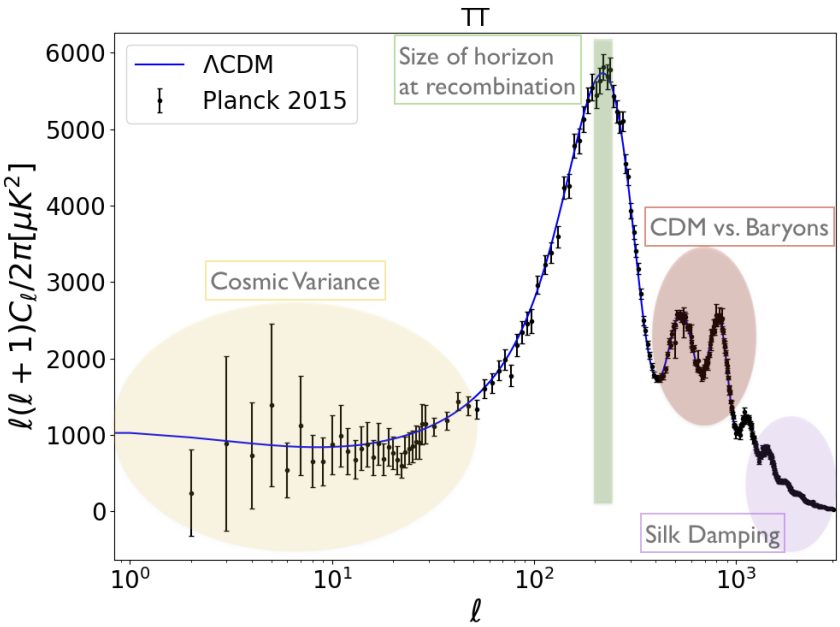


Fig. 4.2. CMB temperature spectrum. The black dots are Planck 2015 data (from [8]) and the solid blue line is the corresponding best fit of ΛCDM obtained using CAMB.

The information that is encoded in the temperature power spectrum of the CMB background corresponds to the epoch of recombination. In that period, the universe was formed by a plasma where photons and baryons[2] were coupled: the photon-baryon plasma. These two species were coupled due to Compton scattering because the mean free path for photons was much smaller than the horizon of the universe and the free electron density was high. This regime is analyzed by using the *Boltzmann distributions*, which describes the interaction of the different species. After recombination, photons are no longer coupled to electrons, and they can move freely through the universe, until today, when we detect them. The Boltzmann distributions for the different species contribute to the source term of the transfer functions ΔT.

[2]In astronomy, the term baryons does not only refer to protons and neutrons, but also to electrons.

Super-horizon modes do not evolve in time[3]. Thus, when we observe the anisotropies on large scales (low ℓ), we are actually observing the perturbations which were larger than the horizon at recombination in their most pristine form. On the other hand, the modes at smaller scales, the sub-horizon modes, evolve over time.

When studying the early stages in the history of our universe, we are interested in photon and matter perturbations. We know how the density scales with respect to the scale factor for radiation and matter. Photon perturbations do not grow as much as matter perturbations with time. Still, the pressure associated to the photon radiation is large, up to the point that it can stop the tendency of matter towards gravitational collapse. Therefore, this system behaves as a *forced harmonic oscillator*, whose simplified equation describing the photon perturbations Θ is,

$$\ddot{\Theta} + k^2 c_s^2 \Theta = F_g, \tag{4.9}$$

where k is the scale, c_s is the speed of sound associated with the photon-baryon plasma and F_g is the gravitational force. The solution of equation (4.9) is oscillatory, which can be clearly observed in the peaks of figure 4.2. Let us focus on a mode that enters the horizon before the time of recombination. The perturbation starts to grow until it reaches its maximum, suppose this happens precisely at the moment of recombination. The scale corresponding to this mode shows large fluctuations with respect to other scales. Thus, we expect to observe a big peak in figure 4.2 corresponding to the time of recombination. Now, let's imagine a mode that enters the horizon slightly earlier in time. This mode peaks earlier and then turns over. Suppose the amplitude of the mode is such that at the moment of recombination the amplitude due to the pressure of the photons is zero. Then, we expect a trough on that angular scale because the fluctuations are very small. A mode that enters even earlier may have time to go through a complete oscillation by the time of recombination, inducing a second peak in the power spectrum.

There are two features of the power spectrum that are worth mentioning. The first one is the fact that odd peaks are higher. This is due to the baryons. Increasing the density of baryons in the universe decreases the speed of sound. From equation (4.9), we observe that the frequency of the oscillations, ω, is proportional to kc_s, and thus, for constant k, the frequency also decreases, making the peaks shift towards larger k. Adding more baryons also enhances the importance of gravity, diminishing the radiation pressure. Therefore, the second peak, which corresponds to an under-density of photons, reflects this fact; it would be harder for photons to escape gravity because of the reduction of the pressure and consequently, the amplitude of the perturbation decreases. Hence, the CMB temperature power spectrum gives a direct measurement of the number of baryons in our universe.

This forced harmonic oscillator system is formed only by baryons and photons. It is also assumed that Cold Dark Matter (CDM) did not interact with photons. The amount of DM has a direct impact on the peaks of the temperature power spectrum.

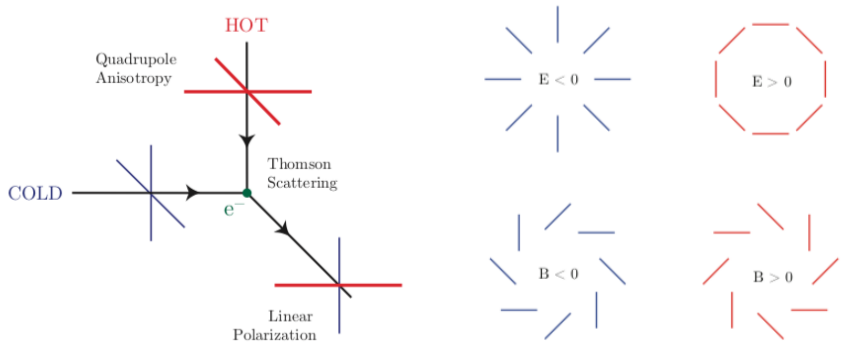
[3]This statement is explained in chapter 2, in subsection 2.5.1.

An increasing density of DM in our universe reduces the overall amplitude of the peaks. Moreover, the difference between the third and the second peaks can give an estimation of the amount of DM.

The second characteristic is that for multipoles $\ell \gtrsim 2000$, oscillations are *damped*. This is the result of the scattering of photons through a medium full of electrons (the baryon-photon fluid is just an approximation that works good in large scales, but in reality, the photons travel a finite mean distance between scatterings). This scattering induces a random walk for a photon, that washes out perturbations smaller than the mean distance between scatterings. This is called *Silk Damping*.

Finally, we need to mention that, in this scenario, we are assuming that the last scattered photons have been travelling without any further incidents until today. We have to take into consideration that recombination takes place after the transition from radiation domination to matter domination. Therefore, the gravitational potentials associated to matter had time to evolve, inducing a change in the redshift of the last scattered photons. Furthermore, since few millions of years, we are Dark Energy (DE) dominated, which implies that the gravitational potentials have decreased inducing new effects over these photons (such as modifications in their redshift). These phenomena are considered in the *Sachs-Wolfe effects* [27]. They are also taken into account in the geometric projections of the transfer functions ΔT.

Polarization anisotropies



(a) Thomson scattering schema (b) Polarization patterns

Fig. 4.3. The CMB radiation gets polarized through Thomson Scattering (*left*), showing different polarization patterns with E and B modes (*right*) [19].

Apart from the CMB temperature anisotropies, the CMB carries out polarization information. The CMB photons are expected to become polarized due to Thomson scattering. Let's assume that the incident radiation (photons) is isotropic. When these photons interact with free electrons, the obtained radiation pattern is unpolarized, since orthogonal polarization directions cancel with each other. Nevertheless, it is

possible to obtain an outgoing polarized radiation if the incoming radiation has a quadrupole component, as shown in figure 4.3.

In order to characterize the polarization anisotropies, we can follow the same procedure that we pursued for the temperature anisotropies, described by equation (4.2). However, we need to use the expansion on the sphere as a function of tensor (spin-2) spherical harmonics, as the polarization is not a scalar field. Instead of using the analogous moments to $a_{\ell m}$ for temperature anisotropies based on the spin-2 quantities, we define two new scalar fields, denominated E (curl-free) and B (divergence-free). These two mode patterns specify completely the linear polarization field.

In principle, we should consider four types of correlations between the temperature and polarization anisotropies when calculating C_ℓ: TT, BB, EE, TE (these last two one are measured by Planck and can be seen in figure 4.10). In the standard model, it is assumed that the correlations TB and EB vanish because of symmetry properties. In the case of BB, Planck has measured the lensing B modes and have set some upper bound for Gravitational Waves at large scales. The BB spectrum is also measured by other experiments, such as the Background Imaging of Cosmic Extragalactic Polarization (BICEP).

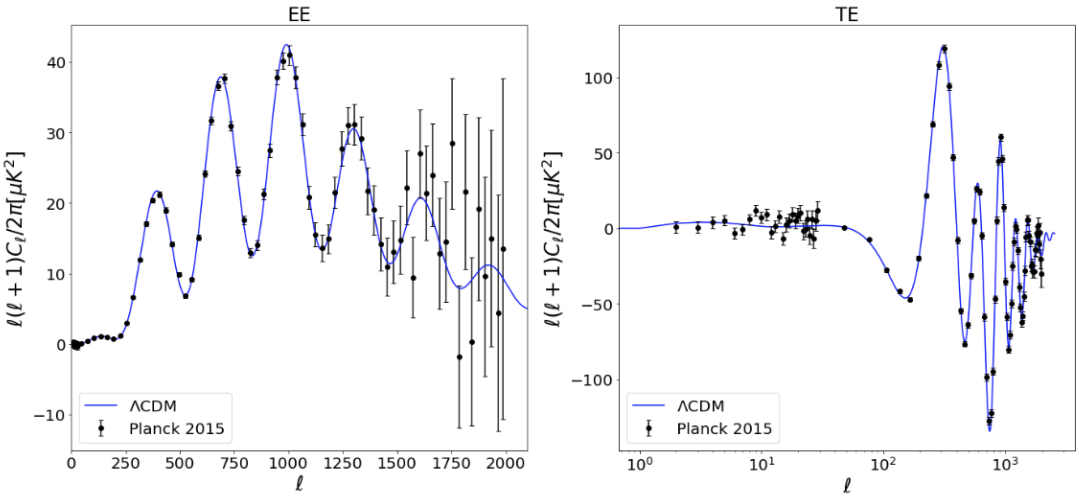


Fig. 4.4. CMB E-mode spectrum (*left*) and cross-spectrum of temperature and E-modes (*right*). The black dots are Planck 2015 data (from [8]) and the solid blue line is the corresponding best fit of ΛCDM obtained using CAMB.

4.1.3 Cosmological Theory Codes

As previously mentioned, the transfer functions account for the evolution in time of each Fourier mode by means of solving Boltzmann equations for each species. The solution is captured by using a *Legendre* expansion (encoded in the geometrical projection P_ℓ of equation (4.8)). The main problem resides in solving thousands of factors of the *Legendre* expansion accurately in time until today. In order to reduce the number of

equations, symmetries and a suitable system of coordinates are used. Moreover, the geometrical part of ΔT encodes the information of the *line-of-sight integral*, where effects such as the *Sachs-Wolfe effects* and Doppler effects are included. If polarization effects are included, this makes the line-of-sight more complicated as one needs to generalize equations to vector/tensors. [39]

Feature	**camb**	**class**
Language	FORTRAN	C
Wrapper	PYTHON	PYTHON, C++, more complete
New PPS	Arbitrary, not available publicly yet	Arbitrary, publicly available

Table 4.1. Main differences between CAMB and CLASS regarding the feature of interest for our current research. The acronym PPS refers to Primordial Power Spectrum.

Thus, to calculate C_ℓ^{XY} (where X and Y can be the temperature T and the modes E or B), we need to compute the transfer functions ΔT to convolve them with the PPS. There are several cosmological theory codes that are precisely designed for this goal. In 1995, *Bertschinger* released the first code named COSMICS, which only solved Boltzmann equations without implementing all corrections in the line-of-sight, which was later included by *Seljak & Zaldarriaga* in a faster code denominated CMBFAST. In 1999, *Lewis & Challinor* reorganized CMBFAST, including new improvements in the implementation of sources and recombination physics from codes such as RECFAST by *Seager*. This new code was called the Code for Anisotropies in the Microwave Background (CAMB). Finally, in 2011 (to get prepared for Planck data analysis), *Lesgourgues & Tram* released a new code, the Cosmic Linear Anisotropy Solving System (CLASS), written from scratch and adding new parametrizations, polarization equations, and algorithms. There are probably a few other codes (mostly private), but CAMB and CLASS are the only publicly available ones that are still maintained and improved to reach higher precision levels.

4.2 Statistics and Model Selection

Cosmological theory codes rely on parameters in order to solve the transfer functions and obtain the CMB power spectra, among others. In fact, in order to measure them efficiently with data, you need to use some refined techniques. Most of these techniques are based on statistics. Dealing with statistics to analyze any experimental data is essential. Statistics is crucial to evaluate the fit of cosmology observations and any theoretical model. It is important to define what you want to test. For example, you may be interested in finding the parameters that fit a chosen theoretical model better, or you may want to discover if a model, with given parameters, is better than another one. These questions in cosmology are usually answered in the Bayesian statistical

framework. This section aims to be a practical guide on how to use Bayesian statistics in the context of cosmology.

4.2.1 Bayesian Statistics

Bayesian statistics is based on producing probability distributions of the parameters of interest, given that the model parameters are treated as random variables [35]. Bayesian statistical analysis relies on *Bayes' Theorem*.

This theorem provides the probability distribution of the parameters θ given the observed data d and a given model M. This probability distribution, $P(\theta|d, M)$, is called *posterior distribution*, and is defined as

$$P(\theta|d, M) = \frac{\mathcal{L}(d|\theta, M)\Pi(\theta|M)}{\mathcal{Z}(d|M)}. \tag{4.10}$$

In equation (4.10), 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 set of parameters θ and the model M. 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 θ given some external information. For instance, this information may come from previously collected data, limits imposed by theory, beliefs of the researchers, 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 matches the experimental data given a set of parameters θ. The likelihood considers the assumptions of the model as well.

In cosmology, we are especially interested in using the data from the Planck mission to test a cosmological model M, such as ΛCDM or any other (such as the inflationary model explained in chapter 3). Therefore, a relation between the experimental data, d, and the theory model defined according to equation (4.7) is needed. The experimental data used in this research, d, can be expressed in terms of the observed signal s and the experimental noise n so that

$$d_i = s_i + n_i, \tag{4.11}$$

where the sub-index i indicates the values of the variables for the $i - th$ experimental observation. In this case, n_i is the electronic noise associated to the measurement and

it is assumed to be normally distributed with zero-mean and variance N_{ij}. Moreover, we also assume the signal s to be normally distributed with zero-mean and variance related to the CMB source measurements according to the equation

$$\langle s_i s_j \rangle = \sum_{\ell} \frac{2\ell + 1}{4\pi} C_\ell P_\ell, \tag{4.12}$$

where P_ℓ is the geometric projection of equation (4.8). Thus, the distribution of the data d is also normal as the sum of two normal distributions is also normal. As a consequence, the corresponding likelihood $\mathcal{L}_{Planck}$ will be a multivariate normal function defined as

$$\mathcal{L}_{Planck} = \frac{1}{(2\pi)^k \sqrt{|\Sigma|}} e^{-\frac{1}{2} d \Sigma^{-1} d}, \tag{4.13}$$

where Σ is the covariance matrix of the data d (which combines the variances of the noise n and signal s) and k is the number of variables measured by observations.

This approach of using a multivariate normal distribution to model the likelihood corresponding to Planck data is very basic. In reality, more complex algorithms and other assumptions are used, especially for $\ell < 29$ (called the `lowTEB likelihood`) and for the TT, TE and EE likelihood at high ℓ (denominated `plikHM_TTTEEE` likelihood). For the sake of simplicity, these further methods will not be mentioned in this thesis but can be found at [1, 52].

Prior

The already known information about the parameters of interest (previous to the research) is summarized in a probability distribution: the prior. The information coming from previous research experiments can be used as prior distributions, such as Gaussian priors resulting from previous results. In general, to define a prior for the set of parameters θ, we need to set the distribution and the hyper-parameters[4] of the distribution.

In our research, in some cases, we restrict the hyper-parameters of the prior distribution according to the theory. In these cases, we assign equal probability to each θ, so that we do not have any preferred values (the values follow a uniform distribution and are sometimes called uninformative priors). 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 5, in subsection 5.1.2 for a sampling method described in the next section).

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 normal distribution and the likelihood is defined as in equation (4.13), the

[4]We call hyper-parameters the parameters of the prior distributions. Examples of hyper-parameters are, for instance, the bounds of a uniform distribution.

posterior distribution is also normal). Otherwise, for non-conjugate priors, the posterior distribution is not of the same family as the prior, which increases the difficulty of the sampling procedure (see subsection 4.2.2).

A different approach: Frequentist Statistics

It is worth mentioning that one could have taken a different approach to obtain quantitative statements about features in the data. For instance, the probability could have been a direct measure of the proportion of outcomes in an experiment. This approach is called *Frequentist approach*, and it is very common in particle physics and other fields in science. This approach is, in general, based on defining a *null hypothesis*, such as "the Primordial Power Spectrum (PPS) is featureless", and an *alternative hypothesis*, for instance, "the PPS has some features described by the model M" [25].

In this case, the 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 should be rejected or not. Associated to the statistic we have a p-value. This value is defined as the probability of obtaining a statistic value equal to or more extreme than the observed one. When the associated p-value is smaller than the probability of rejecting the null hypothesis when it is actually true[5], α, we reject the null hypothesis in favour of the alternative one. In general, α is set to 0.05.

There is a notorious debate regarding the choice of a Bayesian or Frequentist approach, which will not be discussed here, as both of them offer pros/cons (many researchers criticise the use of a prior in Bayesian statistics, as well as the focus on obtaining a p-value smaller or equal than 0.05 have induced a tendency to change models and hypotheses on the fly during research, which is considered a malpractice called *p-hacking*). What it is true is that the use of Bayesian statistics in the research of features in the PPS is more popular [25], due to the demanding numerical simulations and other procedures required to analyse the data in the Frequentist approach. One of this requirements is the large number of needed experiments, such as in particle physics. Unfortunately, in cosmology, we cannot perform several experiments for the same research question.

4.2.2 Parameter Inference

For the estimation of the values of the set of parameters θ, we use Bayes' Theorem (4.10), but the normalization $\mathcal{Z}(d|M)$ is usually not considered, as it only provides a re-scaling of the normalization of the distribution. Thus, in parameter extraction:

$$P(\theta|d, M) \propto \mathcal{L}(d|\theta, M)\Pi(\theta|M). \tag{4.14}$$

To infer the best value of θ 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.

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

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 (4.14). Therefore, we need to use different tools to compute the posterior than evaluating the whole set of parameters in 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. The name Monte Carlo simply means obtaining a representation of a distribution by sampling it in a random way, and the Markov Chain improves the sampling efficiency.

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 only depends on the outcome of the random variable in the previous step. 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 the sample number. 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. This state is reached after discarding a percentage of the initial samples of the chain: the *burn-in* phase.

The simplest scenario for a MCMC is the following. We compute the set of parameters iteratively. At each iteration, we check whether the new set of parameters fits better than the previous one. This step is done by means of checking a selection criterion. Regarding the selection criteria, we define several sampling algorithms. If the selection criterion is fulfilled, we select the new set of parameters. Otherwise, we keep the previous one. At the end of each iteration, we save the set of parameters. The ordered set of saved iterations form what we call a *chain*. The density of points in the chain gives the posterior distributions of the parameters.

There are several sampling algorithms. In this thesis, we focus on Metropolis-Hastings and Nested Sampling.

Metropolis-Hastings algorithm

This algorithm 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 θ', or to come back to the previous step θ. This criterion is based on calculating a ratio, the so-called *acceptance ratio*, a, defined as

$$a = \frac{p(\theta')q(\theta|\theta')}{p(\theta)q(\theta'|\theta)}, \tag{4.15}$$

where $p(\theta')$ is the posterior distribution and $q(\theta'|\theta)$ is the distribution that suggests a new candidate θ' for the next sample value. In general, q is proposed to be a normal distribution. At each step, we must draw a realization of θ' from $q(\theta'|\theta)$, and a random number v from a uniform distribution between 0 and 1. If $a \geq v$, we accept θ', and it becomes a new state of the chain. Otherwise, we reject θ' and the new state of the chain is θ again. This is shown in figure 4.5. In this figure, it is easy to observe that

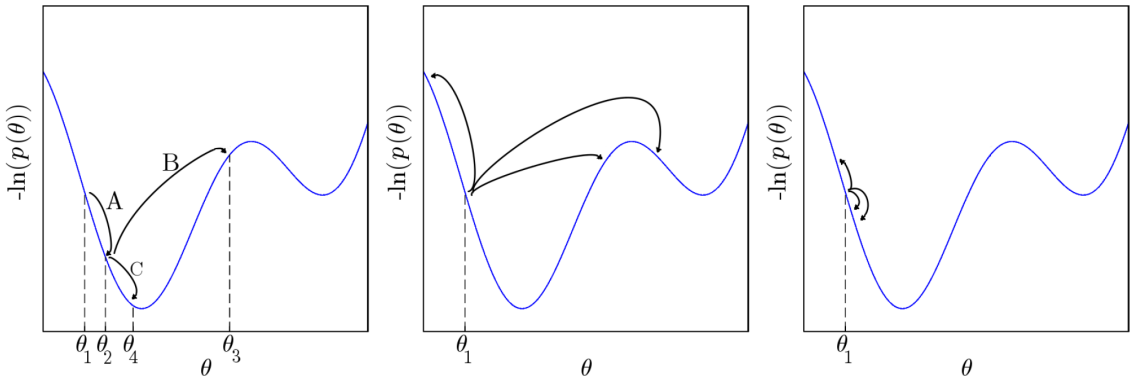


Fig. 4.5. *Left*: An example of a Markov chain constructed by the Metropolis-Hastings algorithm. It starts at 1, then 2 is proposed and accepted (step A), 3 is proposed and refused (step B), and finally, 4 is proposed and accepted (step C). The resulting chain is $(\theta_1, \theta_2, \theta_2, \theta_4)$. *Central*: An example of what happens with a jump size too broad: the chain lacks mobility because all the proposals are unlikely. *Right*: An example of what happens with a jump size too narrow: the chain samples the parameter space very slowly. Figure from [38].

each step only depends on the previous step and that it is independent of the number of steps, making the chain a Markov Chain.

Still, the major disadvantage of this algorithm is the fact that the sampled sets of parameters are correlated, and that we only use the previous sample to obtain the current one. This means that, if we start 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, it is convenient to discard a small percentage of the initial values of the chain, the so-called *burn-in* phase, in order to achieve the stationary property.

The Metropolis-Hastings algorithm allows us to construct a Markov Chain containing as many values as parameters we are trying to infer per step. In general, the more samples we draw, the better the chain will reflect the underlying marginal posterior distribution for each parameter. Therefore, a question arises: *when should we stop sampling?* A simple answer is not available, as there is no test that can affirm whether a chain has converged or not. Still, there are some convergence diagnostics that 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 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 slowing converging; however, if the acceptance is too small, it may indicate that the chain is locally stuck.

- To run several chains at the same time with different starting points and observe whether they give similar results. This similarity is evaluated using the *Gelman-Rubin* statistical criterion R. 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.

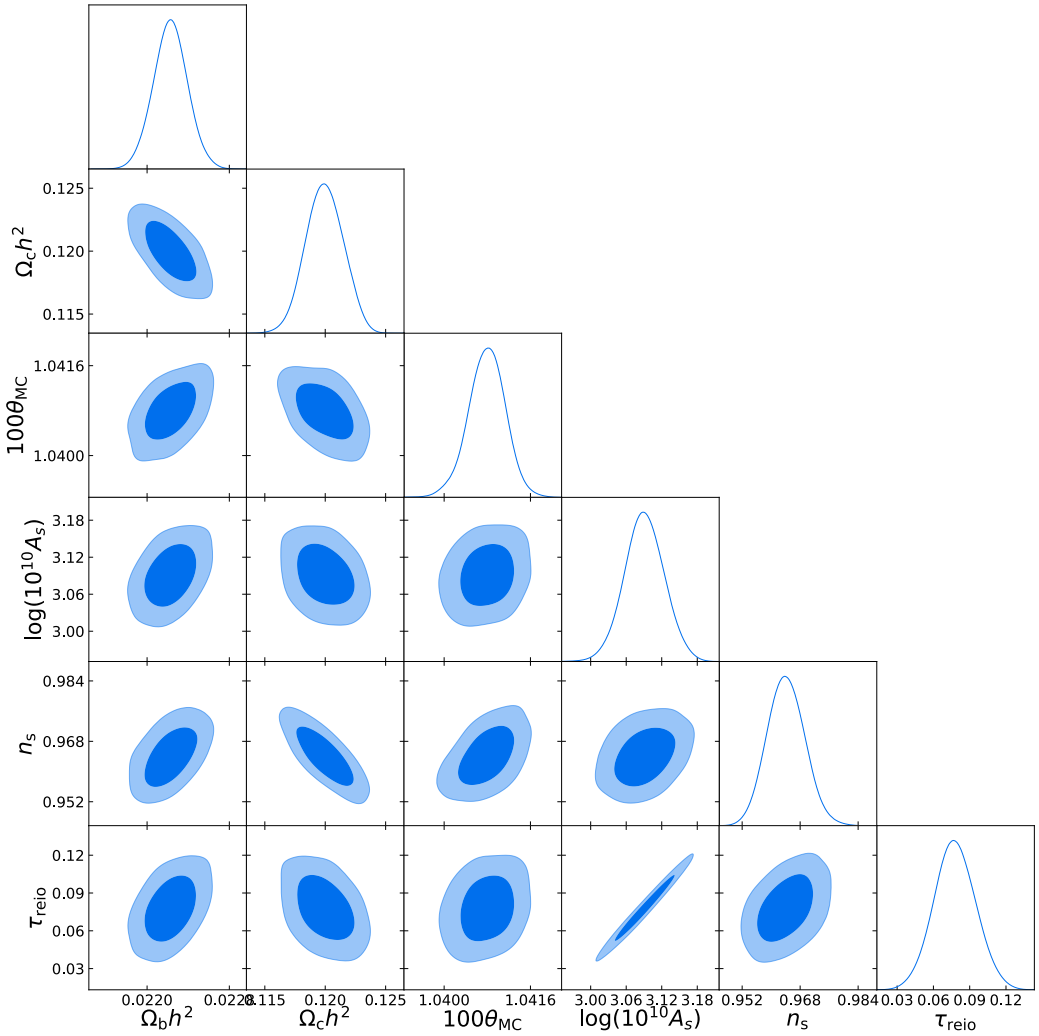


Fig. 4.6. 1D and 2D posterior distributions of the six ΛCDM parameters, running COBAYA (using a MCMC sampler and CAMB as theoretical cosmology code) with 16 chains (Gelman-Rubin convergence criterion $R - 1 \approx 0.03$) and using Planck 2015 temperature and polarization data (TT, TE, EE). This figure is a self-made computation analogous to figure 6 of [51], where CosmoMC was used instead.

The most important cosmological MCMC codes, such as MontePython [18] and CosmoMC [42] (and its new version COBAYA[6]) use this algorithm to explore the

[6]For more information about the new cosmological code COBAYA, see Appendix A.

full range of ΛCDM parameters (see figure 4.6). They use per default a normally distributed q. In higher dimensions, as is the case for ΛCDM, a covariance matrix Σ can be given from a previous MCMC simulation where the stationary property was almost achieved. The codes will use it to decompose the parameters into uncorrelated orthogonal base parameters to find the high probability regions more easily.

Finally, once the posterior distribution is sampled, we usually need to marginalise it to show the results of a multi-parameter fit, for instance, the fit of the ΛCDM model, in order to generate contour plots (see figure 4.6). Marginalization involves projecting the distributions down all other dimensions, and it is usually achieved by means of summing or integrating over the unwanted distributions of the parameters. Mathematically, this can be written as

$$P_{\Theta_1}(\theta_1) = \int_{\theta_2} P_{\Theta_1,\Theta_2}(\theta_1, \theta_2) d\theta_2 \tag{4.16}$$

where Θ_1 are the set of random variables with values θ_1, which are the variables that we want to keep, and Θ_2 are the set of random variables with values θ_2 that we aim to get rid of.

Nested Sampling

The Metropolis-Hastings algorithm tends to struggle with several problems. The first one was already mentioned, and it is the need of eliminating from the chain the burn-in phase. Moreover, because we need to define a step, the Metropolis-Hastings algorithm is inefficient when the posterior distribution is multimodal; that is, if the posterior has several minima or peaks (in figure 4.5, using the Metropolis-Hastings, it is likely that the computation will get stuck in the absolute minimum). 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).

Therefore, in order to circumvent these issues, an alternative way of sampling the posterior was developed by J. Skilling [57]: *Nested Sampling*. This algorithm is based on calculating the evidence Z (marginalized likelihood). From the Bayes' theorem (4.10):

$$Z(d|M) \times P(\theta|d, M) = \mathcal{L}(d|\theta, M) \times \Pi(\theta|M) \tag{4.17}$$

$$\int Z(\theta) \times p(\theta) d\theta = \int \mathcal{L}(\theta) \times \Pi(\theta) d\theta \quad \rightarrow \quad Z = \int \mathcal{L} dX. \tag{4.18}$$

where we have used the fact that $dX = \Pi(\theta)d\theta$ can be modelled into the posterior $dP = p(\theta)d\theta$, and the prior and posterior are normalised to the unit total. From this equation (4.18), we observe that Z, which is the normalization in equation (4.10), is defined as the area below the curve $\mathcal{L}dX$ (see right panel of figure 4.7). 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.

The algorithm works as follows in 1-dimension. In the first step, we generate a starting set S_0 of n samples uniformly distributed over the space and allowed by the

prior Π. Next, we delete the lowest likelihood sample $\mathcal{L}_0$ in S_0, and replace it with a new uniform sample with higher likelihood $\mathcal{L}_1 > \mathcal{L}_0$ (this is called a hard constraint on likelihood value), moving to step S_1.

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 4.7, 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, in order to determine if convergence is achieved or not. The value of this fraction is what we call "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 4.7, red points). This procedure can be generalized to multiple dimensions [57].

In cosmology, there are mainly two scientific Nested Sampling algorithm codes: MultiNest [29], which uses a rejection sampling technique to approach to the hard constraint on the likelihood value (it suffers in high dimensions) and Polychord [30], which is slice-based and the tool used for the main research content of this project.

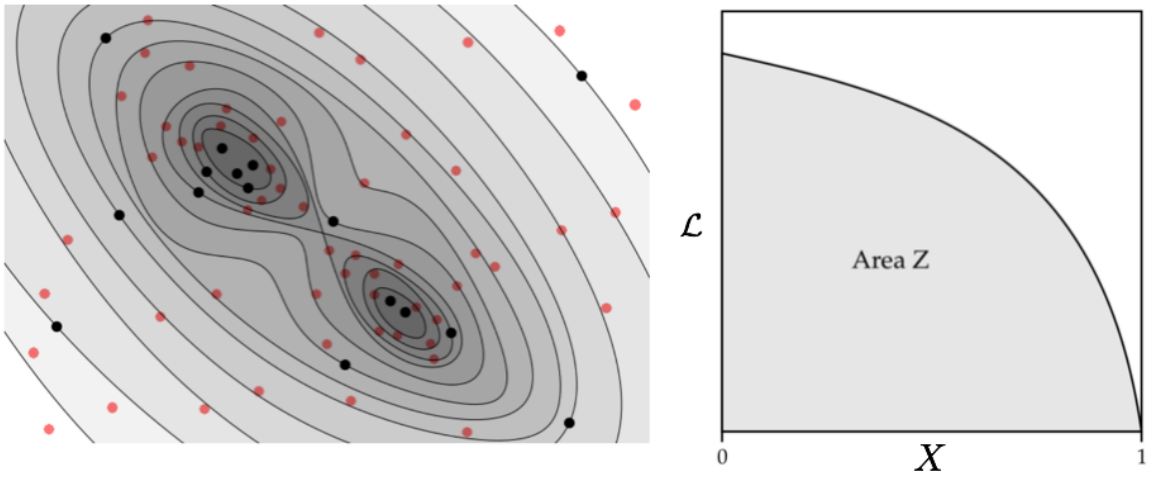


Fig. 4.7. *Left*: Graphical interpretation of the Nested Sampling algorithm by means of looking at the parameter space: the contours specify the areas of the search after each step, where the point with the lowest likelihood was discarded. In every step, the contours are reduced, moving closer to the 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 [39]. *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 [57].

4.2.3 Model Comparison

The selection of a model M_1 over a different model M_2 in Bayesian statistics is usually computed by calculating the *Bayes Factor*, which is inferred from applying model comparison to Bayes' theorem [57]:

$$B = \frac{P(M_1|d)}{P(M_2|d)} = \frac{Z(d|M_1)\Pi(M_1)}{Z(d|M_2)\Pi(M_2)}, \tag{4.19}$$

where $P(M|d)$ are the posterior probability distributions integrated over all the parameter space Θ. 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 assuming that the ratio of priors $\Pi(M)$ is 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.

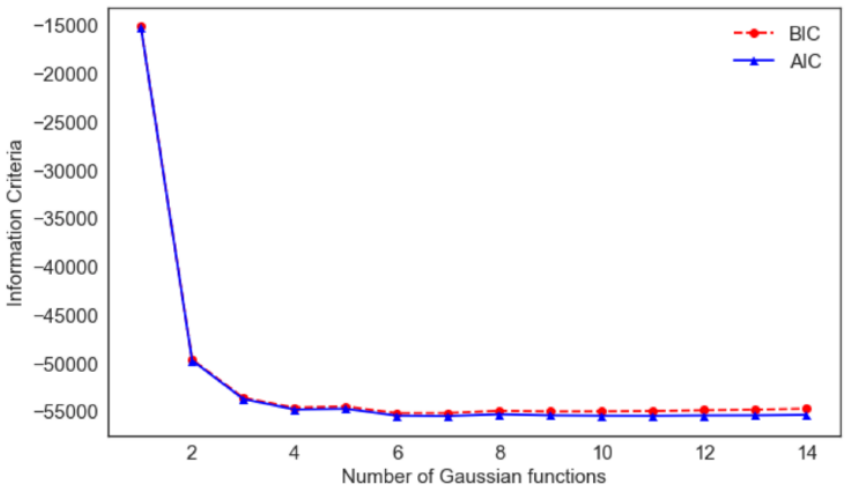


Fig. 4.8. Illustrative example of the asymptotic behaviour of the AIC, in red with circles, and BIC, in blue with triangles, with respect to the number of parameters of a model (in this case, the number of Gaussian functions needed to fit a set of data). In this example, we observe that the most statistical convenient number of Gaussian functions is three. From that point onwards, the information criteria do not continue decreasing. Example from the course [7].

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 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. The information criteria penalize the model taking into account the number of parameters k the model has.

The most popular information criteria are the *Akaike Information Criterion* (AIC) and the *Bayesian Information Criterion* (BIC). The AIC is defined as

$$AIC = -2\ln\mathcal{L} + 2k, \tag{4.20}$$

where $-2\ln\mathcal{L}$ is also denoted by χ^2 and sometimes it is called *deviance*[7].

The BIC is defined as,

$$BIC = -2\ln\mathcal{L} + k\ln N, \tag{4.21}$$

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 parameter (see figure (4.8)).

4.3 Gaussian Processes

Gaussian Processes (GPs) are a Machine Learning technique. Machine Learning (ML) is a field in the interface between computer science and modern statistics that aims to use statistical techniques to give computational algorithms the ability to improve their performance on a task progressively. In the last two decades, there has been an explosion in the implementation of ML techniques, which have become essential in the analysis and interpretation of data (for example, finding patterns in large data sets or comparing data to models) [35].

The world of ML continues to increase every day and this makes it complicated to find an up-to-date definition or try to set the frontier between ML and other fields. In fact, being a *data fluent* person has become an indispensable skill not only for academic environments but also for companies from almost all backgrounds [8]. However, if there are two fields where ML has been used for years, these are particle physics and astronomy, due to the dramatic increase in data volume since the 90's. In modern cosmology, it is getting more and more popular, as demonstrated by the last program of the Cosmo21 conference "Statistical Challenges in 21st Century Cosmology" [6].

In particular, the basic mechanism used by every ML technique is the following: a data set is separated into the *training data* and *generalization data* (this is called *cross-validation*). The training data is used to learn a ML technique, which is denominated commonly as training process (for instance, a regression method to fit the data to a model, or to classify a set of data). Later, the method's performance is tested using the generalization data, which is called the validation process. The separation of the original data set into these two data sets is also studied by several ML techniques. The goal is to reduce as much as possible the error associated with the real data values and

[7]It receives this name because $-2\ln\mathcal{L}$ follows a χ^2 distribution

[8]The positions *Data Scientist* or and *ML expert* have become very popular in the industry, such as accounting firms, airlines, insurance companies or even food delivery companies. If the reader is interested in learning ML or data mining techniques, you can follow seminars at [4].

the values estimated using the tested method during the validation process. This error is called generalization error.

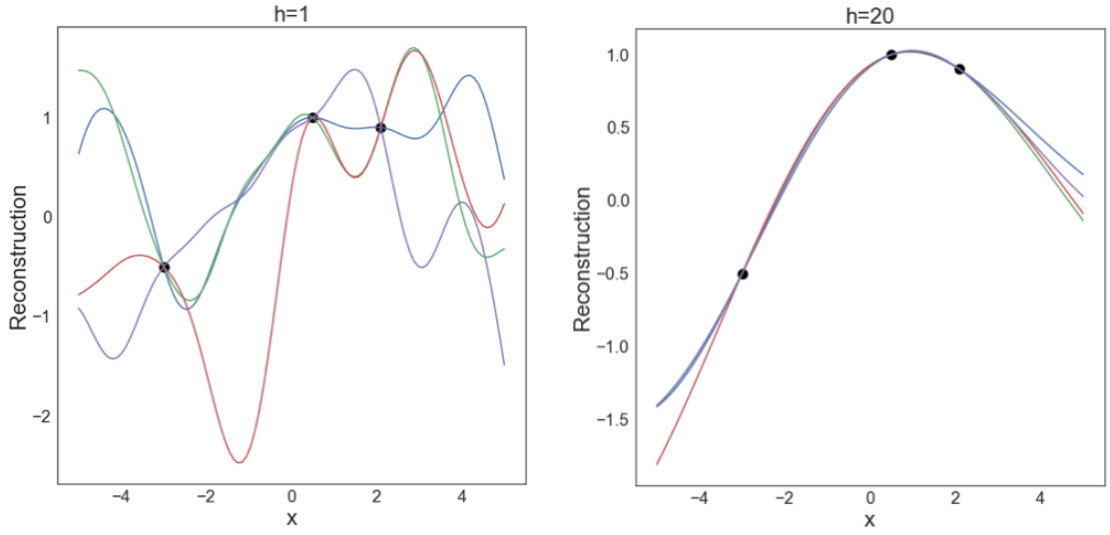


Fig. 4.9. Representation of a collection of random variables generated using Gaussian Processes with different correlation lengths h [7].

The regression methods can be classified, for instance, according to the function they approximate to (linear or non-linear), or by the use of a parameter in the fitting process (parametric or non-parametric). Gaussian Processes are a non-linear and non-parametric generic supervised learning method designed to solve regression and probabilistic classification problems [9]. Formally, they are defined to be a collection of processes based on random variables indexed by time or space, so that every finite collection of those random variables has a multivariate normal distribution defined by the kernel or covariance matrix [54]. In 1-dimension, the kernel is defined as,

$$K(x, x') = \exp\left(-\frac{|x - x'|}{2h}\right), \tag{4.22}$$

where x is the random variable and h is the correlation length, which encodes the length-scale "influence" of each point on the next one (see figure 4.9). In this figure, the left panel shows a more "wavy" reconstruction, because the correlation length is smaller. Every collection of Gaussian Processes is different because, although h fixes the scale, the election of each x is random (in figure 4.9, every colour is a different collection of random variables with the same h). Therefore, we can use Gaussian Processes to define a regression method in which the parameter h of the kernel is fitted by using maximum likelihood estimation as in any common regression method (we generate a Gaussian Process that goes through the training data for a given h, calculate the likelihood, and repeat this process until the optimal h is found). For the kernel of

equation (4.22), if the training data contains only one point, the function resulting from the regression method will be a simple 1-dimensional Gaussian function.

4.4 Data evidence for inflation

In chapter 2, we argued that inflation is a hypothetical period in the history of the universe. For this reason, we strongly look for observational evidence that support the fact that inflation was a real phase in the early universe. At this point of the thesis, we have already described all data analysis tools needed to understand how the current cosmological observations are studied, and how we can argue that some of that evidence does add confidence in supporting inflation.

The first observational evidence in favour of inflation is the measurement of the total parameter density Ω. The Planck collaboration [51] (see again figure 4.6) calculated the ΛCDM parameters using Bayesian parameter inference as described in section 4.2 with the MCMC tool CosmoMC and the TT, TE and EE likelihoods. They concluded that our universe is flat ($\Omega = 1 \pm 10^{-5}$). Although remarkable, inflation predicts flatness by construction, and therefore, we need to continue looking for new evidence.

In the parameter estimation study carried out by the Planck collaboration, two main parameters related to predictions from inflation are included in the ΛCDM model (see table 2.1): n_s and A_s. Simplest models of inflation predict an almost scale invariant and slightly red (higher power at low k) spectrum of primordial perturbations ($n_s \approx 1$). The experimental value is $n_s = 0.9645 \pm 0.0049$, which is almost 1, as predicted. We also know that the collaboration was able to set an upper limit for $r < 0.11$ (when only Planck 2015 data is used) and $r < 0.009$ (when Planck 2015 data is combined with BICEP), which is consistent with single field inflation [51].

Moreover, we have to keep in mind that inflation gives a mechanism to explain the underlying primordial density field at the time of recombination. The CMB power spectra are a simple snapshot of this primordial density field, whose main characteristics were already predicted in subsection 4.1.2. According to the simplest inflationary model, fluctuations froze when exiting the horizon, making the initial phases of the Fourier modes were fixed before the modes entered the horizon again, providing a way to explain the phase coherence of all the Fourier modes. Without this coherence, the CMB power spectra would not show the peaks we were able to explain above, but just white noise. Therefore, the peak-structure of the CMB spectra is one of the main predictions of inflation.

Still, someone may argue that the coherent phases shown in the CMB temperature power spectrum for $\ell > 200$ (see figure 4.10) can be obtained by recalling different causal theories of structure formation with modes well within the horizon. However, when we study the polarization spectra, these theories are also ruled out. In the right panel of figure 4.3, we observe that the cross-correlated spectrum between the CMB temperature and the E-mode shows a peak around $\ell \approx 100$, already detected by the WMAP mission [20]. This minimum between the temperature and polarization spectra comes from the

fact that the velocity of the photons is out of phase with the temperature, and since the polarization spectra are the projection of the quadrupole temperature anisotropy, the polarization peaks are also correlated with the temperature perturbations. This scale $\ell \approx 100$ was not within the horizon at recombination. Consequently, there is no causal mechanism that can describe this phenomenon and it has to be the result of phase coherence.

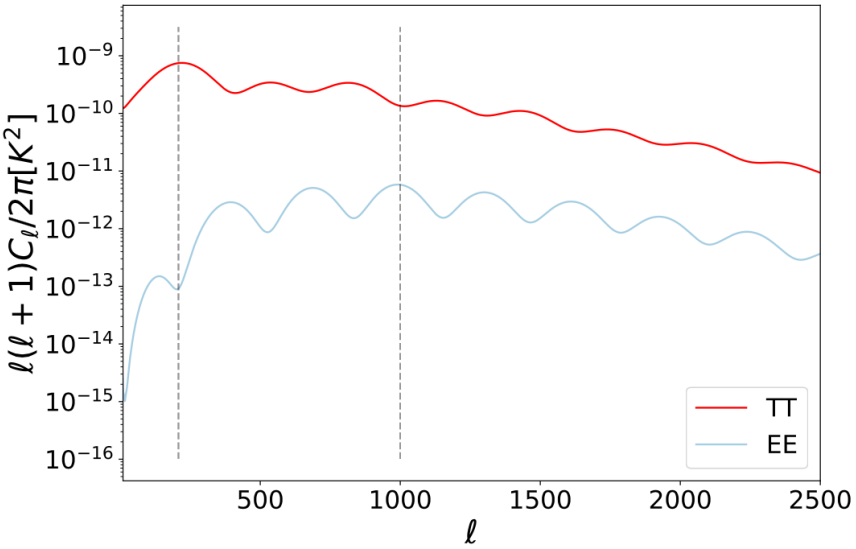


Fig. 4.10. CMB temperature (TT, above line) and polarization (E-modes, below line) spectra obtained from the corresponding best fit of ΛCDM obtained using CLASS. The vertical black dashed lines indicate the counter-phase between the EE and TT spectra.

Chapter 5

Search for features in the CMB

> "Torture the data, and it will confess to anything"
>
> *Ronald Coase* [55]

In this chapter, we update the search for localized oscillatory features in the CMB power spectrum due to transient reductions in the speed of sound using Planck 2015 data, carried out by *J. Torrado et al.* [33, 58]. We have proposed new parametrizations for the reduction of the speed of sound and we have performed parameter estimation for the modified ΛCDM model following the same methodology explained in [58]. For convenience, we have dropped the subindex $\mathcal{R}$ when the features in the primordial power spectrum $\Delta\mathcal{P}/\mathcal{P}$ are shown.

5.1 Research set-up

5.1.1 Ansatz for the reduction

In chapter 3, we characterized the reduction of the speed of sound by the parameters u_{max}, s_{max} and the moment of maximum reduction, τ_0. To compute the modification of the PPS with respect to the power-law parametrization $\Delta\mathcal{P}/\mathcal{P}$ following equation (3.7), we need a parametrization for the reduction of the speed of sound $u(\tau)$. In previous works [12, 33, 58], a Gaussian as a function of e-folds N (or equivalently, in conformal time τ) was proposed:

$$u = A\exp\left\{-\beta(N-N_0)^2\right\} = A\exp\left\{-\beta\left(\ln\frac{\tau}{\tau_0}\right)^2\right\}. \tag{5.1}$$

This reduction is parametrized naturally by 3 parameters: the amplitude $A < 0$ (maximum reduction in intensity), the width $\beta > 0$ (related to the sharpness), and the instant of maximum reduction $N_0 > 0$ (or $\tau_0 < 0$). However, we aim to use the independent parameters of our Slow Reduction Fourier Transform (SRFT) method (u_{max}, s_{max}, τ_0), described in chapter 3. The parameter u is related to the reduction of the speed of sound c_s as,

$$u(\tau) \equiv 1 - \frac{1}{c_s^2(\tau)}. \tag{5.2}$$

The value u_{max} is defined as the maximum reduction of u taking place at τ_0:

$$u_{max} = |u(\tau_0)|, \tag{5.3}$$

where we are using the absolute value of u so that u_{max} is always positive. The parameter s is defined as,

$$s(N) \equiv \frac{\dot{c_s}}{Hc_s} = \frac{1}{c_s}\frac{dc_s}{dN} = \sqrt{1-u(N)}\frac{d}{dN}\left(\sqrt{\frac{1}{1-u(N)}}\right). \tag{5.4}$$

The maximum value of s, named s_{max}, is the maximum of the function $s(N)$ calculated by solving the following equation:

$$\frac{d}{dN}s(N) = \frac{d}{dN}\left[\sqrt{1-u(N)}\frac{d}{dN}\left(\sqrt{\frac{1}{1-u(N)}}\right)\right] = 0. \tag{5.5}$$

Solving for N the equation (5.5), we find the value N_m for which s has a maximum. The function s evaluated at N_m is what we call s_{max}. In short:

$$\frac{d^2}{dN^2}s(N_m) > 0 \rightarrow s_{max} = |s(N_m)|, \tag{5.6}$$

where we have used again the absolute value so that s_{max} is always positive. To use a concrete ansatz, we need to establish a relation between u_{max} and s_{max}, and the amplitude, A, and the sharpness, β. From equation (5.1), we calculate the amplitude:

$$A = -u_{max}. \tag{5.7}$$

To obtain β as a function of A and s_{max}, we use Taylor's expansion for $1/\sqrt{1-u(N)}$, where $u(N)$ is the parametrization of our choice described by A and β, in terms of u up to first order to solve the equation (5.5). This procedure is analogous of the one used in [58].

In this work, we have proposed two other ansätze for $u(\tau)$, which are also naturally described by 3 parameters and softly depart from zero and return, and has an absolute minimum at τ_0 (see figure 5.1). These two functions[1] can be seen in table 5.1. Observe how both parametrizations also fulfil that u_{max} is described by equation (5.7). For later convenience, we defined a new variable to refer to the constants appearing in the obtained values of s_{max} as a function of β and u_{max}, for the new parametrization of $u(\tau)$. This variable is γ. The value of γ for both parametrizations can be seen in 5.1.

The cosine case, which is restricted to only one period, looks similar to the Gaussian case for almost the whole range of values for (A, β, τ_0), and has been mainly used as a test to verify the already-obtained results with the Gaussian ansatz. The hyperbolic secant case only shows some differences with respect to the studied Gaussian parametrization for certain combinations of high values of τ_0 and β (for instance, $\tau_0 < -1000$).

[1]The choice of such a high power for the cosine is motivated by numerical reasons.

	6th power Cosine	Hyperbolic Secant
$u(\tau)$	$u(\tau) = A\cos^6\left[\beta_c(\ln(\tau/\tau_0))\right]$	$u(\tau) = A\text{sech}[\beta_s(\ln(\tau/\tau_0))]$
β	$\beta_c = \lvert s_{max}\rvert/(-3A\gamma_c)$	$\beta_s = \lvert s_{max}\rvert/(-A\gamma_s)$
γ	$\gamma_c = \cos^5\left(\frac{1}{2}\arccos(2/3)\right)\sin\left(\frac{1}{2}\arcsin(2/3)\right)$	$\gamma_s = \tanh[\ln(\sqrt{2}+1)][1/\cosh[\ln(\sqrt{2}+1)]$

Table 5.1. Expressions for β given the different proposed parametrizations for $u(\tau)$. The constant γ is obtained by calculating the maximum of the function (5.4) with respect to N solving equation (5.5).

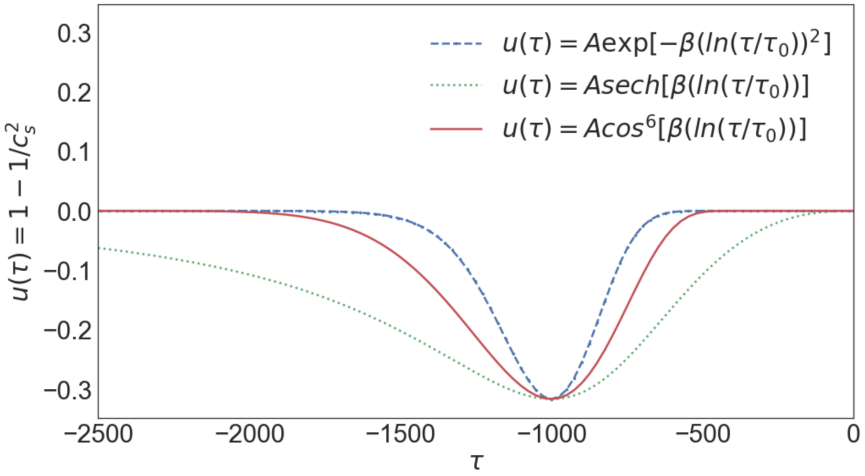


Fig. 5.1. Parametrizations of the reduction of the speed of sound in conformal time $u(\tau)$: a Gaussian function (dashed-blue), hyperbolic secant (solid-red) and a power of the cosine (pointed-green). The values used for the three parametrizations are: $u_{max} = 0.3$, $s_{max} = 0.4$ and $\tau_0 = -1000$.

5.1.2 Prior distributions for the parameters of the features

The prior information is given in equation (3.11). In principle, any given value for u_{max} and s_{max} within that range is permissible, so we could think of using an uninformative prior such as uniform distribution. However, since s_{max} depends on both β and A, setting a uniform prior for both β and A is not enough. This is due to the fact that this uniform rectangular region does not map correctly into a uniform region for u_{max} and s_{max}. Therefore, if we follow this approach, we would search for allowed possible features inefficiently. For this reason, following [58], we use as a prior a *framing square* shape, where the diagonal is $\log[u_{max}] = \log[s_{max}]$ (see figure 5.2). Above the diagonal, the limits given by equation (3.11) are imposed on s_{max}, whereas below, the limits are on u_{max}. A function that fulfils this characteristics is the symmetric log-Beta distribution defined as,

$$\max(\log[u_{max}], \log[s_{max}]) \sim \text{Beta}(a, a); \tag{5.8}$$

$$Beta(a,a) = \int_0^1 x^{a-1}(1-x)^{a-1}dx, \tag{5.9}$$

where $a > 1$ is called the shape parameter and we have chosen $a = 5$, so that we have approximately the 95% confidence level at half the boundaries. We sample in logarithmic scale because we expect the difference in possible values for $\max(\log[u_{max}], \log[s_{max}])$ to be more than two orders of magnitude. In fact, we fix the limits for possible values of $(\log[u_{max}], \log[s_{max}]) \in [-4, 0]$. This choice is coherent with the fact that both u_{max} and s_{max} have to be larger than the slow-roll parameters, according to the theoretical bounds.

In fact, we could have restricted ourselves to use as the lowest limit for $(\log[u_{max}]$, $\log[s_{max}])$ the value -2, given by the central value of the largest slow-roll parameter $\eta \approx 0.02$ according to Planck 2015 [51]. This limit is favoured by previous searches of features [12, 33, 58], where all interesting modes where found in the range $(\log[u_{max}], \log[s_{max}]) \in [-2, 0]$. Still, we decide to make the range broader in order to (1) verify previously found results, and (2) open the possibility of finding new modes.

Regarding the prior information for the instant of the maximum reduction in the speed of sound τ_0, the theoretical framework does not impose any range as far as all possible values of the conformal time used for the parametrizations, τ_i, are negative $(\tau_i < 0)$, because inflation ends at $\tau = 0$. Thus, a uniform prior on τ_0 is chosen[2].

In principle, the decisions taken up to this point allow a sampling from the parameter space consistent with the theory. Still, from a practical point of view, the prior can be simplified: not all the combinations allowed by the prior distributions described above for $\log[u_{max}], \log[s_{max}]$ and τ_0 produce features in the PPS which are visible on the CMB power spectra. First, reductions in the speed of sound which took place at an earlier conformal time, $\tau_0 < -10000$, do not leave imprints on the CMB power spectra, as long as $(\log[u_{max}], \log[s_{max}])$ take values allowed by the prior distribution. This effect can also happen if $\beta_c > 14$ and $\beta_s > 10$.

Second, we need to discard features that are degenerated with changes in the slow-roll parameters. To identify the features we are interested in, we chose sets of values for the parameters that produce, minimum, four full oscillations within the CMB power spectra. For the case of the cosine-like function, we need to impose a limit $\beta_c > 0$, and for the hyperbolic secant $\beta_s > 1$. On top of this requirement, apart from the constraints on β, we also need to constrain the moment of maximum reduction, τ_0, by an upper limit $\tau_0 \leq -70$ (we need the function $u(\tau)$ to return smoothly to 0 for $\tau = 0$, so that we need to restrict the moment of maximum reduction τ_0 not to take place very close to $\tau = 0$). In conclusion, the prior given for the parameters $(\log[u_{max}], \log[s_{max}], \tau_0)$ is,

$$\max(\log[u_{max}], \log[s_{max}]) \sim \text{Beta}(5, 5) \tag{5.10}$$

[2]A uniform prior on $\log[|\tau_0|]$ is also suitable for a search. However, as $\log[|\tau_0|]$ is a *physical time* instead of a *conformal time*, we maintain the sampling on τ_0 as the conformal time is the natural time scale for inflation. Moreover, both kinds of priors were used for the previous search in [58], concluding that both choices are consistent.

$$\log[u_{max}], \log[s_{max}] \in [-4, 0]; \;\; \tau_0 \in [-10000, -70]; \tag{5.11}$$

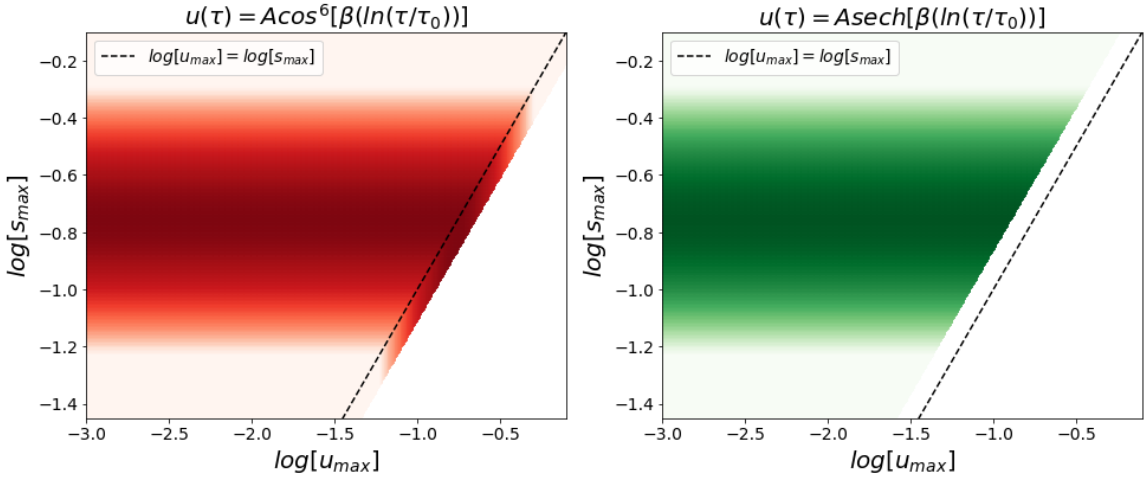


Fig. 5.2. Prior density distributions for the parameters $\log[u_{max}]$ and $\log[s_{max}]$. The limits in the distribution are determined by the line of $log(\beta_c) = 0$ (cosine case, *left*) and $log(\beta_s) = 1$ (hyperbolic secant case, *right*). We dashed black line shows the diagonal $\log[u_{max}] = \log[s_{max}]$.

5.1.3 Sampling Methodology

The features due to the reduction of the speed of sound are calculated by means of taking a fast Fourier Transform of the parametrizations $u(\tau)$, according to equation (3.9). Later, the features are added to the power-law PPS using a modified version of CAMB[3] and a pivot scale $k_* = 0.005\ Mpc^{-1}$.

The features are fitted using the Bayesian parameter estimation method explained in chapter 4. To do so, we use the unbinned likelihood[4] of the CMB TT, TE and EE power spectra data from the Planck 2015 release: `lowTEB + plikHM_TTTEEE_unbinned`. The use of the unbinned likelihoods is very important. For example, we have features for values $\tau_0 < -1500$ that oscillate faster than the predefined bins of $\Delta k \approx 10^{-3}\ Mpc^{-1}$. Therefore, if we use the binned likelihoods, our method will be blind to those features. The used sampling tool is CosmoChord, as previous explorations already showed that the posterior distribution is multimodal.CosmoChord is a modified version of CosmoMC that contains CAMB as cosmological theory code. It also includes Polychord as sampler. The free fitting parameters are the 6 parameters of ΛCDM model, meaning $(\Omega_b h^2, \Omega_c h^2, \theta_{MC}, \tau_{reio}, \ln A_s, n_s)$ (explained in chapter 2), for which we have used a uniform prior, plus the 3 feature parameters $(\log[u_{max}], \log[s_{max}], \tau_0)$.

[3]For more information about the modification of CAMB, the reader can find information in Appendix B.

[4]A basic introduction to the Planck 2015 likelihoods can be found at chapter 4, subsection 4.2.1.

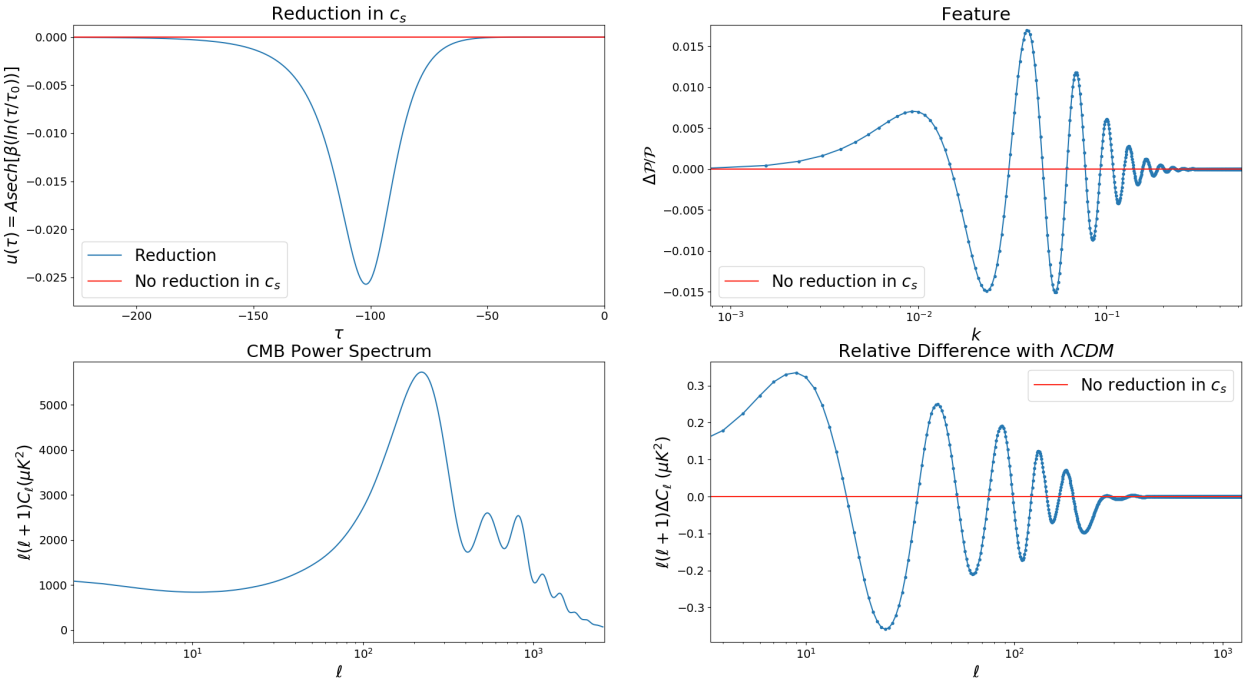


Fig. 5.3. Graphical summary of the sampling methodology. *Up left panel*: the reduction of the speed of sound in terms of $u(\tau)$. *Up right panel*: feature calculated using a fast Fourier Transform of $u(\tau)$. *Bottom left panel*: CMB TT power spectrum corresponding to the modified PPS. *Bottom right panel*: relative difference between the modified CMB TT power spectrum and the ΛCDM case. Example calculated using the hyperbolic secant case (same applies for other parametrizations), using the best-fit values for the mode 100, shown in table 5.2.

We do expect some degeneracy with respect to the PPS parameters $(\log A_s, n_s)$, which motivates our decision of setting all ΛCDM model parameters free. To obtain an estimation of $\Delta\chi^2$ with respect to the unmodified ΛCDM, we have run CosmoChord with a featureless PPS using the same likelihoods of Planck 2015: `lowTEB + plikHM_TTTEEE_unbinned`. The reference value, this is, the value corresponding to the unmodified ΛCDM model, is $\chi^2 = 34654.9$. We have fixed the value of the nuisance parameters[5] of the Planck 2015 likelihoods to their best fit obtained by the Planck MCMC with the binned likelihoods `lowTEB + plikHM_TTTEEE`. After some tests, we have not found any evidence of degeneracies between the feature parameters and the nuisance parameters, which support our decision of fixing them. In the future, more complete studies to evaluate the effect of each of these nuisance parameters on the features parameters are suggested, although they are currently beyond the scope of this research. We have run CosmoChord with 16 MPI processes, fixing the number

[5]The nuisance parameters are associated with the experimental calibration of the instrument and other effects. For more information read [51].

of live points to 1000 and with stopping criterion[6] 0.02. With this sampling set-up, the runs converged after four or five days.

5.2 Results of the fit

Observing the marginalised posterior for $(\log[u_{max}], \log[s_{max}], \tau_0)$ using the hyperbolic secant ansatz for $u(\tau)$ (see figure 5.4), we identify two main modes at low τ_0 (see table 5.2). These modes are approximately at $\tau_0 \approx -100, -200$, and are associated to amplitude values of around $u_{max} \approx 0.017$. These modes are in the range where $u_{max} > s_{max}$.

Comparing these results with previous research [12, 33, 58], we conclude that we have recovered the two main modes already-found at low τ_0. However, in [58], *J. Torrado et al.* observed two modes at $\tau_0 \approx -800, -1000$, for which $u_{max} \approx s_{max}$. We were not able to recover with the hyperbolic secant ansatz these two modes at high τ_0. The disappearance of the modes at high τ_0 can be due to the fact that, for this particular range, $\tau_0 > -500$ and $u_{max} \approx s_{max}$, the hyperbolic secant shape is very different to the Gaussian ansatz (see figure 5.1), producing a complete new feature pattern which is not favoured by the data.

Mode name	Ansatz	$\log[u_{max}]$	$\log[s_{max}]$	τ_0
100	Hyperbolic Secant	[-1.79 (-1.57) -1.45]	[-1.07 (-0.78) -0.58]	[-101 (-103) -105]
	6th power Cosine	[-1.85 (-1.62) -1.41]	[-1.18 (-0.94) -0.48]	[-99 (-105) -110]
200	Hyperbolic Secant	[-1.77 (-1.61) -1.42]	[-0.75 (-0.43) -0.24]	[-199 (-205) -209]
	6th power Cosine	[-1.73 (-1.59) -1.39]	[-1.22 (-0.61) -0.39]	[194 (203) 209]
800	6th power Cosine	[-1.79 (-1.57) -1.45]	[-0.93 (-0.41) -0.09]	[-770 (-864) -912]
1000	6th power Cosine	[-1.93 (-1.67) -1.41]	[-0.83 (-0.35) -0.12]	[-946 (-1083) -1108]

Table 5.2. Intervals for $(\log[u_{max}], \log[s_{max}], \tau_0)$ corresponding to the modes at 68% confidence level, with the maxima a posteriori in parenthesis, in the posterior distributions of figures 5.4 and 5.5. The mode names are chosen in agreement with [58].

The results obtained using the cosine ansatz (figure 5.5) does not provide new information with respect to previous searches [12, 58], as we recovered exactly the same four modes that were already observed with the Gaussian parametrization (see figure 3 of [58]). This result is reasonable, as the 6th power cosine and the Gaussian parametrizations have a similar shape for the most range of values for the feature parameters. We also see traces of a faint mode around $\tau_0 \approx -370$, which was already observed in [12] (under the name $\mathcal{A}$ mode) using Planck 2013 data, but not observed in the update search [58]. The cause of the partial re-appearance of this mode may be the subtle differences in shape between two ansätze.

[6]The stopping criterion is defined in chapter 4, in subsection 4.2.2 when the Nested Sampling algorithm is explained.

Our modified ΛCDM model with the three features' parameters is not statistically favoured with respect to the simplest ΛCDM model when the Bayes factor is calculated, for any of the two different ansätze. This result was already found when the Gaussian parametrization was used [58]. However, we cannot miss what the final goal of this project is: these modes should be the starting point for searches of features in the bispectrum, which can make our modified ΛCDM model statistically more significant and robust in the future.

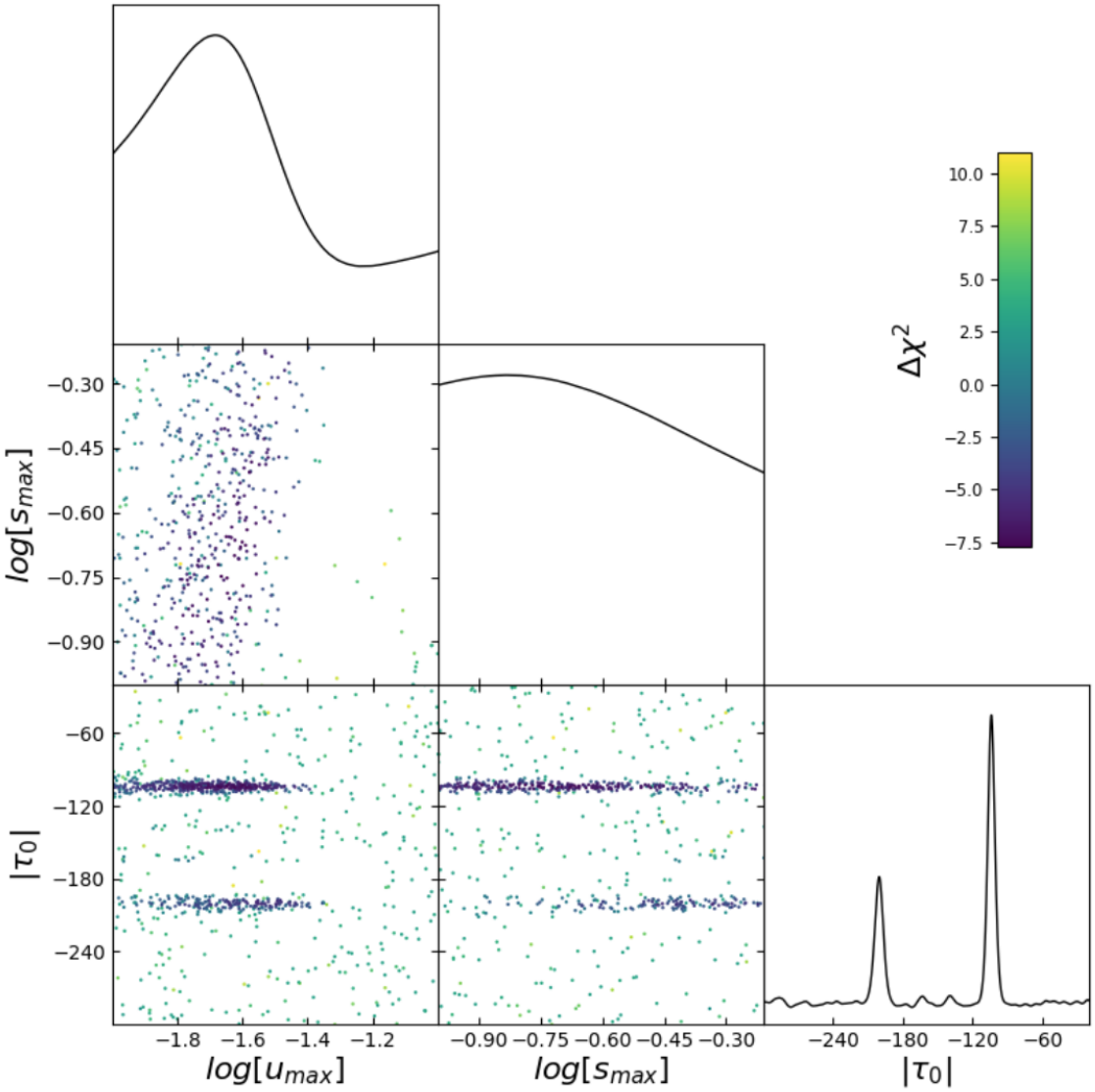


Fig. 5.4. 1D and 2D marginalized posterior distributions of the feature parameters $(\log[u_{max}], \log[s_{max}], \tau_0)$, using the hyperbolic secant ansatz. The colour-bar indicates the difference in χ^2 with respect to the best fit of the featureless ΛCDM model.

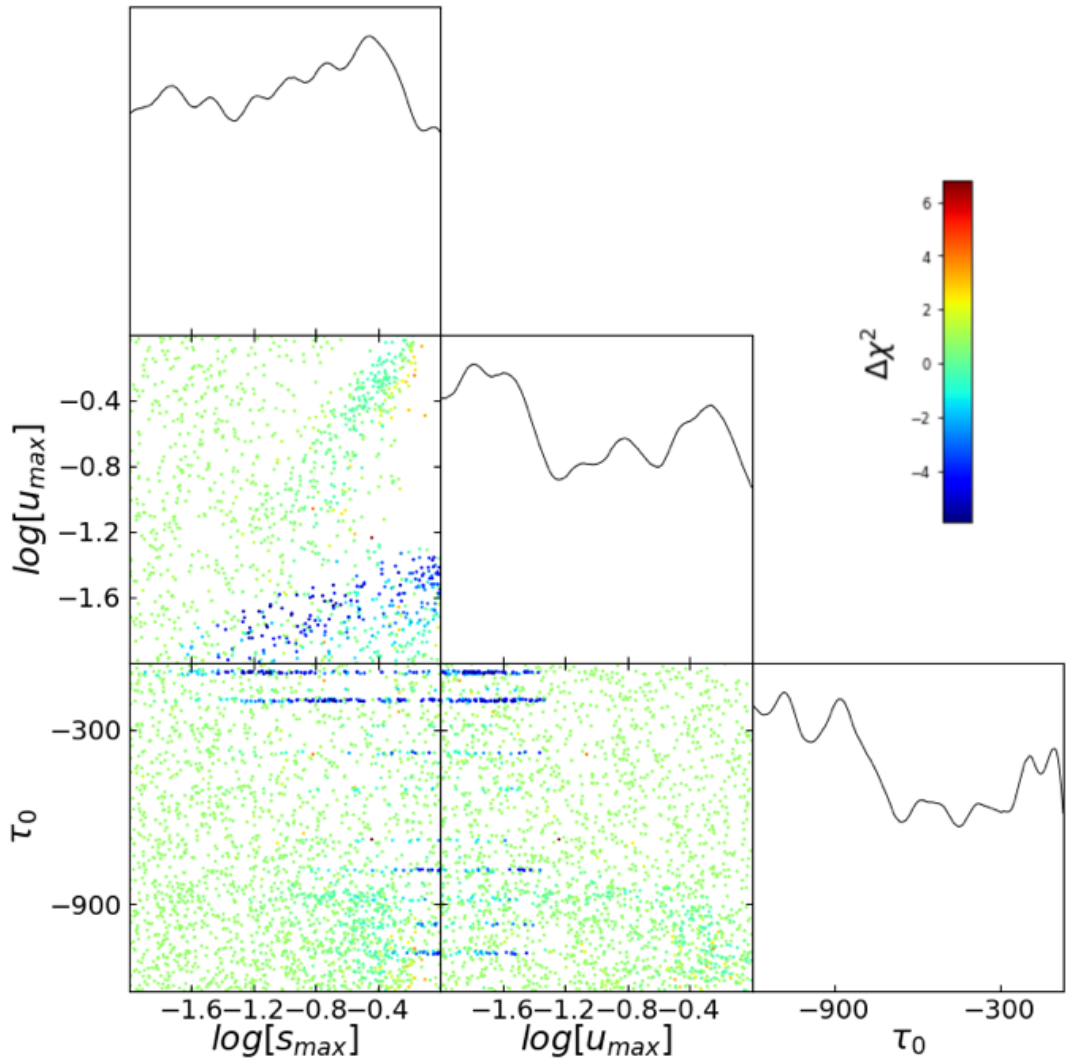


Fig. 5.5. 1D and 2D marginalized posterior distributions of the feature parameters $(\log[u_{max}], \log[s_{max}], \tau_0)$, using the 6th power cosine parametrization. The colour-bar indicates the difference in χ^2 with respect to the best fit of the featureless ΛCDM model.

5.3 Reconstruction of $u(\tau)$

5.3.1 Motivation

In section 5.2, we have summarized the several modes found when two different parametrizations for $u(\tau)$ are used. We have also concluded that we did not find any trace of modes at $\tau_0 \approx -800$ and $\tau_0 = -1000$ when we used the hyperbolic secant parametrization. One can argue that the reason lies on the integration method used

by CAMB for the PPS to obtain C_ℓ. Nevertheless, after some testing, it seems that the most likely reason behind the loss of these modes at high τ_0 is the difference in shape between the hyperbolic secant with respect to the Gaussian and the 6th power cosine at high value of τ_0 when u_{max} and s_{max} are reasonable small. Moreover, the re-appearance of traces of the so-called $\mathcal{A}$ mode when the 6th power cosine is used with respect to the Gaussian parametrization is also a signal that the dependence of our model on the ansatz for $u(\tau)$ is noticeable.

This fact also remarks the intrinsic need of our theory to define a priori what a *good* ansatz is, which does not only model the dependence of the reduction of the speed of sound in time, but also fulfils the numerical needs of our computation (for instance, numerical stability during integration processes).

Therefore, a major improvement in our research, which can make our model be more robust and consistent, is to avoid using a predetermined function for $u(\tau)$. Instead, we aim to reconstruct the shape of $u(\tau)$ favoured by the data and allowed by the theoretical bounds at the same time. For this, we use a hyper-parametric Machine Learning (ML) technique, called Gaussian Processes (GPs) and explained in chapter 4 in detail. Apart from overcoming the weaknesses of using a predetermined parametrization function for $u(\tau)$, we can also explore more complex features with this new method. In fact, it was already pointed out in [58] that the possibility of having reductions in the speed of sound at two different values of τ_0 cannot be excluded. For instance, a combinations of the modes $\tau_0 \approx -800$ or $\tau_0 \approx -1000$, and $\tau_0 \approx -200$ or $\tau_0 \approx -100$ are allowed by the theory, and these situations arise naturally when GPs are implemented.

Reconstructions at the level of the PPS have already been attempted [50]. However, the degeneracy with respect to the data is very big. A reconstruction of $u(\tau)$ instead of reconstructing directly PPS allows us to include extra theoretical limits and conditions, which should minimize the impact of only using the data.

5.3.2 Methodology

In light of the limitations in the use of a predetermined parametrization for $u(\tau)$, we reconstruct the shape of $u(\tau)$ using GPs. To reconstruct $u(\tau)$, we set a number k of training nodes $(\tau_{0,k}, u_k)$ (see figure 5.6), where $\tau_k < \tau_{k+1}$, and we use the Gaussian kernel defined in equation (4.25) with a fixed correlation length h to generate the random collection of points. If we had 3 free parameters when we were using a concrete parametrization for $u(\tau)$, in this case we have $2k+1$ parameters (k pairs $[\tau_{0,k}, u_k]$ and h).

To verify that the parameters of the reconstruction $(\tau_{0,k}, u_k)$ fulfil the EFT conditions of equation (3.11), we establish a relation with respect to u_{max} and s_{max}, which are calculated *a posteriori*. Once the reconstruction is done, we look for the maximum value of $u(\tau)$ and we verify that $u_{max} < 1$. The value of u_{max} is related to the reconstruction

through equation (5.4). If we re-write that equation we find

$$s_{max} = \max\left[\frac{|\tau|}{2(1-u(\tau))}\frac{du(\tau)}{d\tau}\right] < 1, \tag{5.12}$$

where $du(\tau)/d\tau$ is calculated numerically and later, the maximum of s is found computationally. If either u_{max} or s_{max} (or both) do not fulfil the requirements, the reconstruction is discarded.

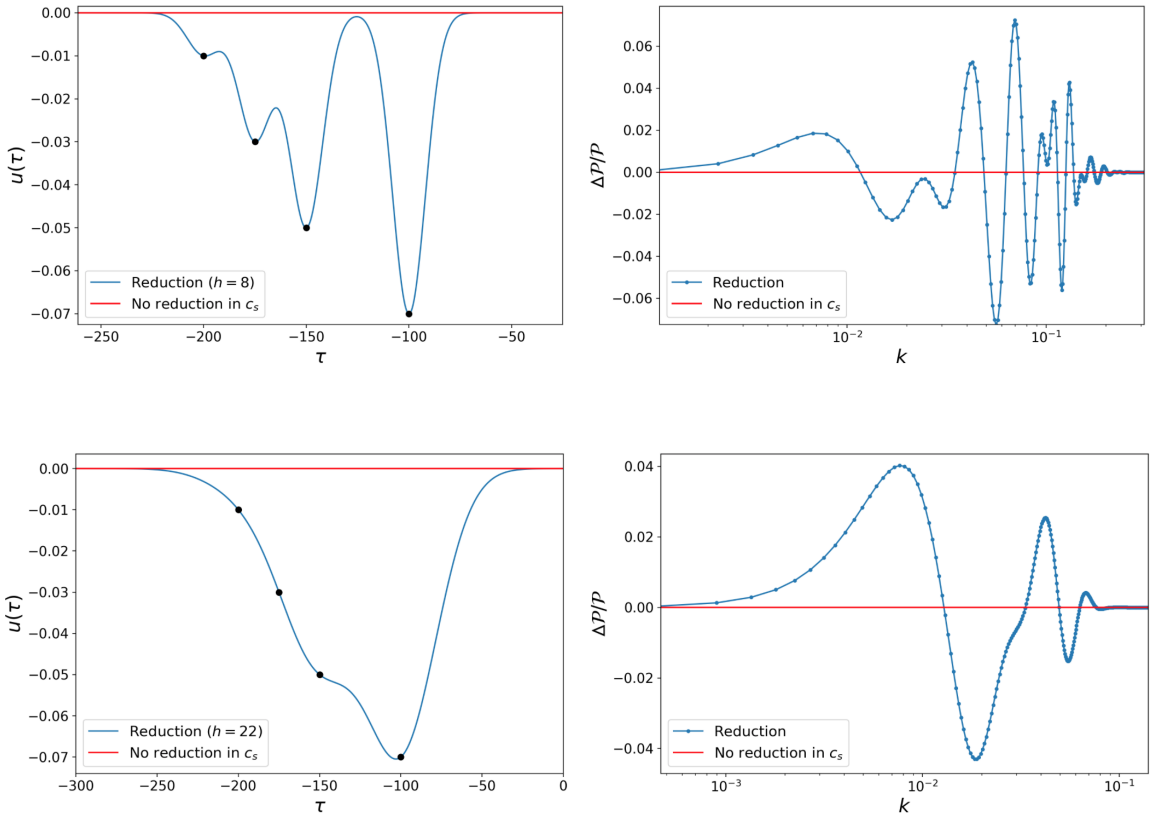


Fig. 5.6. *Left*: reconstruction of $u(\tau)$ using Gaussian Processes, with four training nodes (black dots), a Gaussian kernel and two different values of the correlation length h. *Right*: corresponding features in the primordial power spectrum $\Delta\mathcal{P}/\mathcal{P}$ from the reconstruction.

We also impose a preliminary prior distribution on the derived values of $\log[u_{max}]$ and $\log[s_{max}]$, which follows a Beta distribution (see figure 5.7) as in the case of the hyperbolic secant and the 6th power cosine. We discard the reconstruction if at any moment $u(\tau)$ becomes positive. In this work in progress, we are using the new Bayesian tool COBAYA and the sampling algorithm POLYCHORD. We use the unbinned likelihoods `lowTEB + plikHM_TTTEEE_unbinned` corresponding the Planck 2015 data, and we run with 16 MPI processes, fixing the number of live points to 1000 and with stopping criterion 0.02. At this moment of the preliminary research process, we are

fixing the ΛCDM model parameters as well as the nuisance parameters to the best value of the Planck 2015 MCMC fit, for computational reasons (we are interested in this moment in the convergence of the features' parameters in a reasonable period of time).

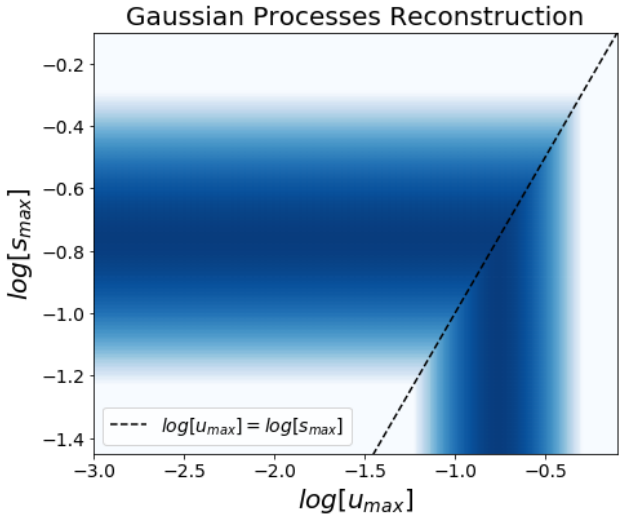


Fig. 5.7. Prior density distribution for the parameters $\log[u_{max}]$ and $\log[s_{max}]$. The diagonal sets $\log[u_{max}] = \log[s_{max}]$.

5.3.3 Preliminary results

In figure 5.8, we observe the marginalised posterior for $(\log[u_{max}], \log[s_{max}], \tau_0)$ using the reconstruction of $u(\tau)$ obtained with Gaussian Processes. At the time of writing this report, the sampling method had not yet converged. Still, we are able to identify the two main modes already observed at low $\tau \approx -100$ and $\tau \approx -200$, corresponding to values of $u_{max} > s_{max}$. These results shows how these modes at low τ are clearly favoured by the data, independently of the parametrization used. It also points out that Taylor's expansion carried out to calculate s_max in the three parametrizations of $u(\tau)$ works, as long as the amplitude of the reduction, which is directly related to u_{max}, is sampled from the allowed prior distribution, which assures that u_{max} is small (we are using the same prior for both the fixed parametrizations and the reconstruction).

As mentioned above, at the time of writing this report, the sampling method has not yet converged, as PolyChord gets stuck for high values of τ_0. Further research has to be done regarding this matter. We need to check if the prior distribution that we are using for our feature parameters is suitable in the case of the Gaussian Processes' reconstruction, as well as evaluate if our stopping criterion of 0.02 for PolyChord is too demanding to find new clusters, which will be candidates for possible modes.

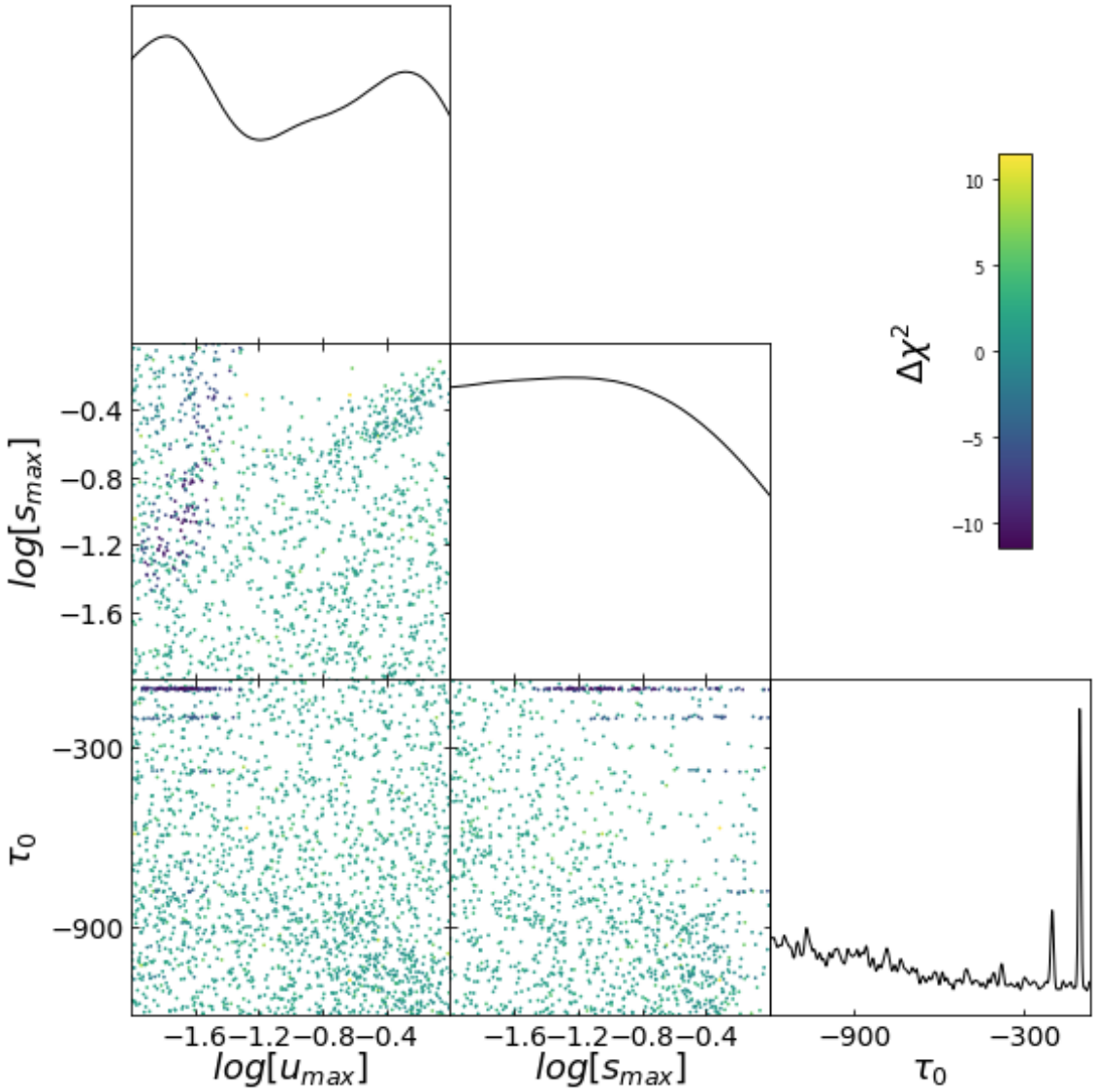


Fig. 5.8. Preliminary 1D and 2D marginalized posterior distributions of the feature parameters $(\log[u_{max}], \log[s_{max}], \tau_0)$, using the GPs reconstruction for two nodes. The colour-bar indicates the difference in χ^2 with respect to the best fit of the featureless ΛCDM model.

Chapter 6

Conclusion and outlook

> "The beginning is the most important part of the work"
>
> *Plato* [53]

In this thesis, we have detailed the theoretical framework of the current ΛCDM model of cosmology, which includes the inflationary paradigm. Furthermore, we have showed how the analysis of Cosmic Microwave Background (CMB) cosmological data is usually done from the statistical point of view, pointing out the most important available analysis tools.

We have also studied possible extensions of the single field inflation theory from the Effective Field Theory (EFT) point of view, analysing a concrete model within this theory that predicts features in the correlation functions, such as the primordial power spectrum (PPS) and in the bispectrum. These features are a consequence of a reduction in the speed of sound of the curvature perturbation due to, for instance, a turn in the field trajectory.

Our main goal has been to detect those possible features in the PPS analysing current CMB data. For that, we have proposed a modified ΛCDM model with nine parameters: the six ΛCDM parameters and three other parameters that define the shape of the reduction of the speed of sound. Within this approach and using parameter estimation, we have identified certain possible fits of our modified model with respect to the data, which we shall call *modes*.

We were able to identify some of these modes with already-detected ones showed in previous researches [33, 58]. We have observed that some of these modes disappear when a concrete parametrization for the reduction of the speed of sound $u(\tau)$ (the hyperbolic secant) is used, concluding that our results are very dependent on the chosen parametrization for $u(\tau)$. In addition, we have observed that, in fact, our theory does not restrict the possibility of having two reductions in the speed of sound at different times. Thus, the data could favour another kind of feature patterns corresponding to several relative reductions of the speed of sound, as it was already mentioned in [58].

For this reason, as an extension to this project, we are working on reconstructing $u(\tau)$ using Gaussian Processes (GPs). We are carrying out a preliminary search for features using a prior similar to the one used for previous parametrizations, where the theory parameters $(\log[u_{max}], \log[s_{max}])$ are checked numerically a posteriori. After analysing some preliminary results of the posterior distribution, we have observed a

problem in the convergence of the sampling algorithm. A possible explanation to this problem is a large number of small clusters detected by POLYCHORD for high values of τ_0. This may suggest that we need to assess the quality of our current prior in future steps. Also, we still need to evaluate information criteria to estimate the improvement when new training nodes are added in the reconstruction.

Besides, as explained in chapter 3, our extension of the simplest inflationary model gives observable predictions in the bispectrum. Hence, we expect to perform a joint search for features not only in the Primordial Power Spectrum (PPS) but in the bispectrum as well. To this aim, we need the likelihoods for the bispectrum, which in principle could be released once that Planck mission has concluded. This aspect is crucial: our modified model of ΛCDM with parameters $(\log[u_{max}], \log[s_{max}], \tau)$, which determine the feature, is disfavoured with respect to the simplest ΛCDM model when the Bayes' factor is calculated. However, the detected modes in the PPS are crucial to give a prediction of the whole bispectrum. Therefore, in order to increase the significance of our research, we need to verify these current promising modes in the bispectrum as well.

With the perspective of the imminent final release of Planck data, we have continued to increase our computational expertise to repeat the analysis in the future using up-to-date techniques (a new implementation on a highly-used code such as CAMB and beta-test of COBAYA, the new cosmological Bayesian tool). For this reason, brief appendices including the major advantages and implementations of these codes have been presented in this thesis.

Finally, another important point that we can investigate in the future is to repeat the statistical analysis from a frequentist point of view, using re-sampling techniques such as bootstrap during the analysis of the results. Furthermore, with the ending of the Planck mission and the incoming launch of the Euclid satellite, it seems more than convenient to start thinking about how to incorporate Large Scale Structure cosmological datasets into our current research, as they look promising to study non-Gaussianity, following the methodology used by [33].

Appendix A

Cobaya: the new cosmological Bayesian analysis tool

A.1 Software Information

cobaya[1] [59], the *COde for BAYesian Analysis*, is a Bayesian statistical analysis tool designed for sampling and modelling in the framework of cosmology. The code is able to explore priors and posteriors using several samplers and theoretical cosmology codes. The results (sampling chains) can be analysed with GetDist [41]. Its authors are *J. Torrado* and *A. Lewis*, and at the moment, it is under beta test. It is planned to be released during the summer of 2018.

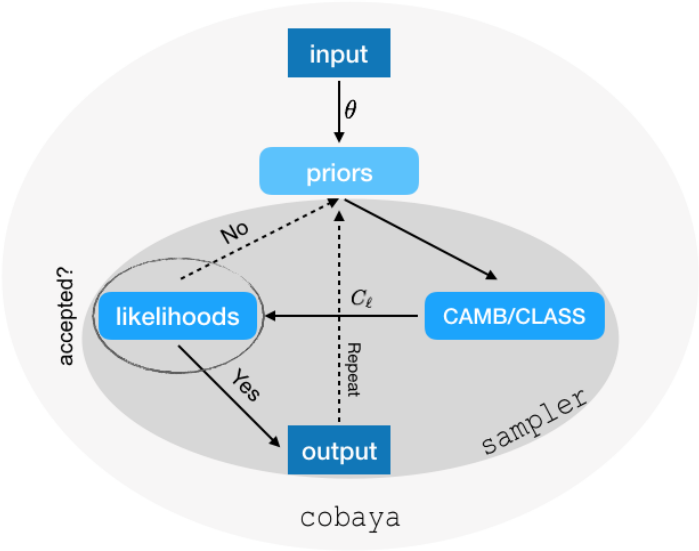


Fig. A.1. cobaya working scheme. The input yaml file is set with the priors, likelihoods and the sampler that we decide to use. A set of parameters, θ, allowed by the prior, is passed to the theoretical cosmology code, which calculates the CMB power spectra or any other cosmological observable. The result is compared with the data likelihood, and by the sampling algorithm, it is decided whether θ is accepted and written in the chain. The sampler repeats this procedure until convergence is reached.

cobaya, which was designed for general purpose statistical tool, is fully written in python (compatible with both v.2.7 and v.3.0), callabled from the shell or from a jupyter notebook and uses yaml language for the input files.

[1]The word "cobaya" means in Spanish "guinea pig".

A.2 Improvements

With respect to other tools, COBAYA gives:

- Automatic installation script of cosmological codes (theoretical and samplers) and data (and their dependencies): only needs a YAML file indicating desired modules.
- flexibility, as it is easily extensible: possibility of defining own likelihoods and priors distributions effortlessly.
- the choice between the two major cosmological theory codes CLASS [21] and CAMB [44, 32] indistinctly.
- the possibility of choosing from several samplers, for instance the MCMC CosmoMC [43, 40, 48] and the Nested Sampling PolyChord [30, 31].
- an easy use of modified versions of theoretical cosmology codes (COBAYA only needs the path to their corresponding PYTHON wrappers).
- the possibility of create citations easily: automatic script to create citation references of the used modules.
- the opportunity of analysing derived parameters directly calculated by COBAYA.

A.3 List of bugs and other implemented suggestions

Within the detailed beta test carried out for the test of COBAYA, we suggested new improvements and we also found and fixed some bugs:

- COBAYA failed to finish its algorithm when convergence was reached but only one chain was being computed.
- COBAYA did not die in cluster computations as some errors were not classified as *error log outputs*.
- COBAYA stopped every time you re-run computations in the same folder and a chain file is also present. Currently, it only stops if the chain file is not empty.
- Included derived parameters on the fly, weighted by the prior, after the computation of the theoretical cosmology code.
- Tested automatic installation script of COBAYA modules. Included possibility of changing the path to common libraries (i.e: Lapack, OpenBlas) from the YAML file.

- Easy check included in COBAYA to verify which libraries NUMPY is using, and change them if desired.

A.4 Acknowledgements

I acknowledge Santander Supercomputacion support group at the University of Cantabria who provided access to the supercomputer Altamira Supercomputer at the Institute of Physics of Cantabria (IFCA-CSIC), member of the Spanish Supercomputing Network, for performing simulations/analyses.

COBAYA takes advantage of multiple PYTHON modules such as MATPLOTLIB [34], SCIPY [36] and PANDAS [46].

Appendix B

Arbitrary Change of the PPS Modification for CAMB

B.1 Software information

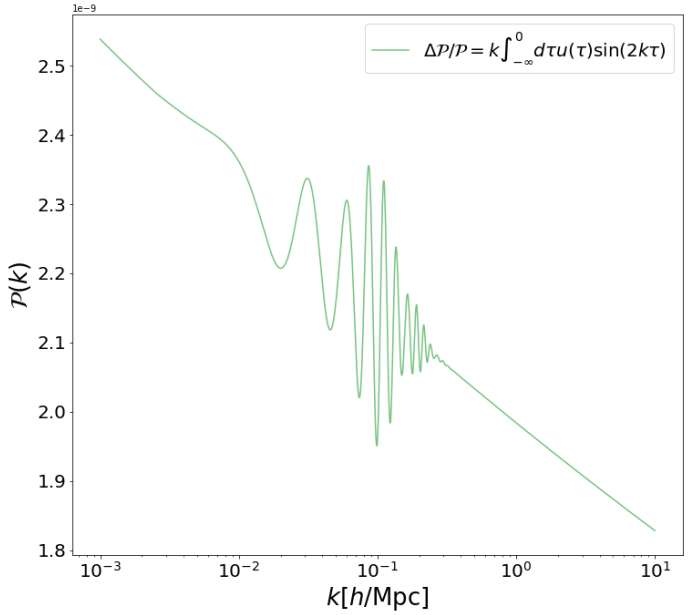


Fig. B.1. Example of a feature (predicted for a transient and mild reduction of the speed of sound $u(\tau)$ parametrized using Gaussian Processes, with training nodes $\tau_t = [-150, 120]$ and $u_t = [-0.03, -0.1]$) calculated with the new modification of CAMB.

In order to calculate the predicted features in the Primordial Power Spectrum (PPS), it has been necessary to modify CAMB accordingly. This modification was first introduced by *J. Torrado* [33] to allow for features in the PPS from reductions of the speed of sound (see chapter 3), parametrized using a Gaussian ansatz for $u(\tau)$. During the course of this project, we have extended the modification not only to include any prospective new parametrizations for $u(\tau)$, but an arbitrary modification of the PPS, $\Delta\mathcal{P}/\mathcal{P}$ (see figure B.1).

This modification of CAMB can be found in https://gitlab.strw.leidenuniv.nl/SRFT/CAMB_external_PS.git, and at the moment, it is not publicly available[1]. This modification is planned to be merged with the official incoming version of CAMB in the near future.

[1] If the reader is interested in the project, an e-mail can be sent to *canasherrera@lorentz.leidenuniv.nl*

B.2 Structure

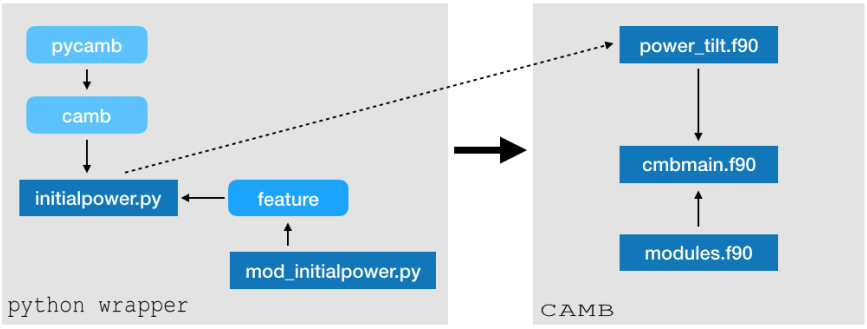


Fig. B.2. CAMB modification to include an arbitrary modification of the power-law type PPS. The feature is generated via the PYTHON wrapper of CAMB, PYCAMB, which includes the class `InitPower`. If the external mode for an arbitrary $\Delta\mathcal{P}/\mathcal{P}$ is activated (`Mode` $= 1$), the `initialpower.py` module calls the submodule `feature`, which includes the self-written functions of your model predictions for features. The feature is passed to the FORTRAN main core of CAMB, to the module `power_tilt.f90`, which sums the feature to the power-law PPS.

The arbitrary feature in the PPS is computed using the CAMB PYTHON wrapper, called PYCAMB, which includes the `InitPower` class in the module `initialpower.py`. The mode is activated by the attribute of the `InitPower` class: `InitPower.Mode = 1`. If this mode is activated, it will call the submodule `feature` within the PYTHON wrapper. This module includes the PYTHON file `mod_initialpower.py`, which can be arbitrary modified to include any predictions for modifications of the PPS. At the moment, a toy-model designed by *M. Martinelli* is included in the DEMO case. The parameters that define the feature will be saved as attributes of the class `InitPower`. This step is essential, as the new sampler COBAYA automatically detects all attributes of any class from PYCAMB, and it is able to perform sampling over them.

For convenience, some automatic sampling functions to define the range of k have been included. Nevertheless, the user is able to define its own range and density for the k. The function `mod_initialpower.py` also calculates the minimum distance within the k, Δk_{min}, which is directly determined by the density of the sampling the user performs. This value is necessary for the computation of the C_ℓ at later stages.

Once that the feature $\Delta\mathcal{P}/\mathcal{P}$ is created, it is passed to the FORTRAN core of CAMB, concretely to `init_power.f90`. This routine adds $\Delta\mathcal{P}/\mathcal{P}$ to the power-law PPS. To calculate the C_ℓ, it is necessary to integrate over the modified PPS (see chapter 4). However, if the interval dk for integration is too big, the feature will be missed. Therefore, we set dk to be the the minimum distance within the k, Δk_{min}. For a $\Delta k_{min} = 10^{-5}$, CAMB needs approximately 15 seconds to calculate the CMB power spectra, in a computer *i*7 with 16 GB of RAM. We also set the corresponding $d\ell$ to be the corresponding minimum distance $\Delta\ell_{min} \approx 13911.6\Delta k_{min}$.

B.3 Improvements for publication

In order to publish our modification in the new version release of CAMB, we need to perform further changes in the code such as:

- Include an easy reading routine in `power_tilt.f90` to obtain from a text file a random PPS for those users that still would like to use the full-written FORTRAN code and avoid the PYTHON.
- Include interpolation routines from `subroutines.f90` from the development branch, instead of the ones available at `equations.f90`.
- Re-define dynamically the parameters for the integration density `k_boost` and `l_boost` in `cmbmain.f90` and `modules.f90`.
- Include modifications to take into account the effects in the lensed power spectrum.
- Link COBAYA with a general function not necessary under the module `mod_initialpower.py` (under development, as we need COBAYA to finish its integration with the new development branch of CAMB).
- Change variable names for more recognizable ones such as `primordial_k` instead of `dPoP0`.

B.4 Acknowledgements

We would like to thank *Antony Lewis* for interesting discussions about how to incorporate the modification in CAMB for a future publication.

Appendix C

From multifield inflation to EFT

The aim of this appendix is to obtain the Effective Field Theory (EFT) expansion of adiabatic perturbations for a single field inflationary model, represented in the action of equations (3.3) and (3.4) in chapter 3, and calculated in [24]. We start from a multiple field model. This appendix is based on the previous work of EFT of inflation [14, 23, 24] and on basic concepts of multifield inflation [17, 16]. For convenience, we set the reduced Planck mass $M_P = \sqrt{1/8\pi G} = 1$.

C.1 Basics of multifield inflation

Let us begin by setting the action for a set of multiple scalar fields ϕ^a minimally coupled to gravity,

$$S = \int d^4x\sqrt{-g}\left[\frac{1}{2}R - \frac{1}{2}g^{\mu\nu}\gamma_{ab}(\phi)\partial_\mu\phi^a\partial_\nu\phi^b - V(\phi)\right], \tag{C.1}$$

where $a = (1, ..., N)$ being N the total number of fields, R is the Ricci scalar of the space-time metric $g_{\mu\nu}$, γ_{ab} is the space metric spanned by the fields ϕ^a, and V is the scalar density potential. The equation of motion is,

$$\Box\phi^a + \Gamma^a_{bc}\partial_\mu\phi^b\partial^\mu\phi^c - V^a = 0 \tag{C.2}$$

where $\Box = \nabla_\mu\nabla^\mu$, ∇_μ being the covariant derivative and $V^a \equiv \gamma^{ab}V_b$. The Christoffel symbols Γ^a_{bc} are defined according to the field space metric γ_{ab} as

$$\Gamma^a_{bc} = \frac{\gamma^{ad}}{2}(\partial_b\gamma_{dc} + \partial_c\gamma_{bd} - \partial_d\gamma_{bc}). \tag{C.3}$$

Given this set-up, we can study the solutions of the system (C.2).

C.1.1 Background

We focus first on the equations corresponding to a homogeneous and isotropic cosmological background, where $\phi^a = \phi^a_0(t)$ is a scalar field solution, which defines the trajectory in field space as a function of time t. Let $g_{\mu\nu}$ be the flat Friedmann-Lemaitre-Robertson-Walker (FLRW) metric,

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a(t)\delta_{ij}dx^i dx^j. \tag{C.4}$$

where $a(t)$ is the scale factor defined in chapter 1. Using equation (C.4), the equations of motion (C.2), which describe the evolution of the background, are

$$D_t\dot{\phi}_0^a + 3H\dot{\phi}_0^a + V^a = 0, \tag{C.5}$$

$$3H^2 - \dot{\phi}_0^2/2 - V = 0, \tag{C.6}$$

where D_t is the covariant time derivative defined as,

$$D_t Y^a \equiv \dot{Y}^a + \Gamma^a_{bc}\dot{\phi}^b Y^c, \tag{C.7}$$

and $\dot{\phi}_0^2/2$ is the total kinetic energy expressed as,

$$\dot{\phi}_0^2 \equiv \gamma_{ab}\dot{\phi}_0^a\dot{\phi}_0^b \leftrightarrow \dot{\phi}_0 = \sqrt{\gamma_{ab}\dot{\phi}^a\dot{\phi}^b}. \tag{C.8}$$

where, from it, we find:

$$\dot{H} = -\frac{\dot{\phi}_0^2}{2}. \tag{C.9}$$

These equations of motion (C.5) and (C.6) are written in terms of the basis formed by the fields ϕ^a, and they do not give easily any insight about how the system behaves. For this reason, a different basis is used, called *kinematic* basis, which is formed by the tangent $T^a = T^a(t)$ and normal $N^a = N^a(t)$ direction vectors to the trajectory. Let us assume that we are dealing with a system formed by only two fields. In this case, these vectors are defined as,

$$T^a \equiv \frac{\dot{\phi}_0^a}{\dot{\phi}_0} \qquad\qquad N^a = \gamma^{ab}N_b = \gamma^{ab}\sqrt{\gamma}\epsilon_{bc}T^c, \tag{C.10}$$

where $\gamma = \det(\gamma_{ab})$ and ϵ_{bc} is the Levi-Civita symbol, so that

$$N^a T_a = 0 \qquad\qquad N_a N^a = T_a T^a = 1, \tag{C.11}$$

Projecting the equations of motion along T^a we obtain

$$\ddot{\phi}_0 + 3H\dot{\phi}_0 + V_T = 0, \tag{C.12}$$

where $V_T \equiv T^a V_a$. Projecting along N^a,

$$D_t T^a = -\frac{V_N}{\dot{\phi}_0}N^a, \tag{C.13}$$

where $V_N \equiv N^a\partial_a V$. Following the same procedure used for single field inflation in chapter 2, we define the slow-roll parameters, ϵ, η^a, which take into account the

background time evolution:

$$\epsilon \equiv -\frac{\dot{H}}{H^2} \qquad\qquad \eta^a \equiv -\frac{1}{H\dot{\phi}_0} D_t \dot{\phi}_0^a, \tag{C.14}$$

where η^a is a two-dimensional field that gives the rate of change of $\dot{\phi}_0^a$ in time. Decomposing η^a in terms of the new vectors of the kinetic basis $\eta^a = \eta_\parallel T^a + \eta_\perp N^a$, we find

$$\eta_\parallel = -\frac{\ddot{\phi}_0}{H\dot{\phi}_0} \qquad\qquad \eta_\perp = -\frac{V_N}{H\dot{\phi}_0}, \tag{C.15}$$

where $\eta_\parallel$ agrees with the definition of $\tilde{\eta}$ given in chapter 2 in equation (2.27), for single field inflation, and $\eta_\perp$ accounts for the rate of turning in the trajectory, telling how fast the tangent vector T^a changes in time. This rotation in time is encoded in the angular velocity Ω of the trajectory defined as,

$$\Omega \equiv -N_a \dot{T}^a = \eta_\perp H = \frac{V_N}{\dot{\phi}_0}. \tag{C.16}$$

To summarize, and as in the single field inflationary case, slow-roll inflation will take place as long as,

$$\epsilon \ll 1 \qquad\qquad \eta_\parallel \ll 1, \tag{C.17}$$

where a large value for $\eta_\perp$ is consistent with the slow-roll condition. In fact, we will see later how $\eta_\perp$, and so $\dot{\phi}$, is related to a change in the speed of sound c_s.

C.1.2 Perturbations

We now study perturbations around the background solution $\phi_0^a(t)$. We parametrize the perturbations as

$$\delta\phi^a(t, \mathbf{x}) = \phi^a(t, \mathbf{x}) - \phi_0^a(t), \tag{C.18}$$

and we decide to work with the gauge invariant quantities $\mathcal{R}$ and $\mathcal{F}$ defined as

$$\mathcal{R} \equiv \frac{H}{\dot{\phi}} T_a \delta\phi^a + \psi \qquad\qquad \mathcal{F} \equiv N_a \delta\phi^a, \tag{C.19}$$

where ψ is the scalar perturbation of the spatial part of the metric, and in the *comoving* gauge described in chapter 2, agrees with the comoving curvature perturbation $\mathcal{R}$. The quadratic action for $\mathcal{R}$ and $\mathcal{F}$ is,

$$S_2 = \frac{1}{2}\int d^4x a^3 \left[\frac{\dot{\phi}_0^2}{H^2}\dot{\mathcal{R}}^2 - \frac{\dot{\phi}_0^2}{a^2 H^2}(\nabla\mathcal{R})^2 + \dot{\mathcal{F}}^2 - \frac{(\nabla\mathcal{F})^2}{a^2} - M_{eff}^2\mathcal{F}^2 - 4\Omega\frac{\dot{\phi}_0}{H}\mathcal{F}\dot{\mathcal{R}}\right], \tag{C.20}$$

where M_{eff} is called the "effective" mass[1] of $\mathcal{F}$ defined as,

$$M_{eff}^2 \equiv m^2 - \Omega^2, \tag{C.21}$$

and $m^2 \equiv V_{NN} + \epsilon H^2 \mathbb{R}$, where $\mathbb{R}$ is the Ricci scalar associated to γ_{ab}.

C.2 EFT description when $\mathcal{F}$ is heavy

The action (C.20) shows how Ω couples both perturbations $\mathcal{R}$ and $\mathcal{F}$. The term Ω appearing in the effective mass is due to a correction in the potential because of the centripetal force suffered by the turn. If $\mathcal{F}$ is heavy, the coupling reduces the effective mass M_{eff}, suggesting that there may be a single field effective description of the theory when $\Omega^2 \sim m^2$, as the mass hierarchy would have broken.

We calculate the equations of motion associated to the system described by (C.20) in Fourier space:

$$\ddot{\mathcal{R}} + (3 + 2\epsilon - 2\eta_{\parallel})H\dot{\mathcal{R}} + \frac{k^2}{a^2}\mathcal{R} = 2\Omega\frac{H}{\dot{\phi}_0}\left[\dot{\mathcal{F}} + \left(3 - \eta_{\parallel} - \epsilon + \frac{\dot{\Omega}}{H\Omega}\right)H\mathcal{F}\right], \tag{C.22}$$

$$\ddot{\mathcal{F}} + 3H\dot{\mathcal{F}} + \frac{k^2}{a^2}\mathcal{F} + M_{eff}^2\mathcal{F} = -2\Omega\frac{\dot{\phi}_0}{H}\mathcal{R}. \tag{C.23}$$

The values $\mathcal{R} = \text{const}$ and $\mathcal{F} = 0$ (for an arbitrary value of Ω) are non-trivial solutions of equation (C.22) and (C.23). In fact, if $\mathcal{F}$ is heavy, $\mathcal{F} \to 0$ and $\mathcal{R} \to \text{const}$ after horizon exit during inflation. Then, in the regime when $M_{eff} \gg (k/a)$, we can ignore the time derivatives of $\mathcal{F}$ in equation (C.23). Taking into consideration that $M_{eff} \gg H$, we can disregard the friction term $3H\dot{\mathcal{F}}$ as well. Therefore, equation (C.23) becomes,

$$\frac{k^2}{a^2}\mathcal{F} + M_{eff}^2\mathcal{F} = 2\dot{\phi}_0\eta_{\perp}\mathcal{R} \leftrightarrow \mathcal{F} = \frac{2\dot{\phi}_0\eta_{\perp}}{\frac{k^2}{a^2} + M_{eff}^2}\mathcal{R}, \tag{C.24}$$

where we have used the definition of Ω from equation (C.16). Substituting relation (C.24) into the action (C.20) and neglecting the terms $\dot{\mathcal{F}}$, we obtain an effective action for the comoving curvature perturbation only,

$$S_{eff} = \frac{1}{2}\int d^4x a^3 \frac{\dot{\phi}_0^2}{H^2}\left[\frac{\dot{\mathcal{R}}}{c_s^2} - \frac{k^2\mathcal{R}^2}{a}\right], \tag{C.25}$$

where the speed of sound c_s of the adiabatic mode is given by,

$$c_s = 1 + \frac{4\Omega^2}{\frac{k^2}{a^2} + M_{eff}^2}. \tag{C.26}$$

[1]The name effective mass is used only for historical reasons, but it is not a proper mass.

This effective field theory is valid as far as we can continue ignoring the kinetic terms of $\mathcal{F}$ in equation (C.23):

$$\ddot{\mathcal{F}} \ll M_{eff}^2 \mathcal{F}. \tag{C.27}$$

Equation (C.27) represents the adiabatic condition that every mode k needs to fulfil for the effective field theory to be valid.

C.3 Generalization

We have obtained an effective action for the comoving curvature $\mathcal{R}$. Now, we study a generalization of this framework. For that, we are going to identify the Goldstone boson corresponding to the broken time translational symmetry $\pi = \pi(t, \mathbf{x})$, with the adiabatic mode of our system $\mathcal{R}$. To obtain the effective action, we are eventually interested in integrating out the heavy field $\mathcal{F}$.

C.3.1 Gauge relation between $\mathcal{R}$ and π

In this section, we will remark in which manner π and $\mathcal{R}$ are related by gauge transformations. We begin by using the Arnowitt-Deser-Misner (ADM) formulation[2] to write the metric as,

$$ds^2 = -N^2 dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt) \tag{C.28}$$

where N is the lapse function, N^i is the shift vector and h_{ij} is the metric on the spatial slices. To study the deviations from the homogeneous and isotropic background, we follow the same procedure carried out in subsection C.1.2, parametrizing the perturbations in this system using:

$$\phi^a(t, \mathbf{x}) = \phi_0^a(t + \pi) + N^a(t + \pi)\mathcal{F}, \tag{C.29}$$

$$h_{ij} = a^2(t + \pi)e^{2\mathcal{R}}\delta_{ij}, \tag{C.30}$$

where π parametrizes the deviations along the trajectory and $\mathcal{F}$ the displacements from the trajectory. There are non-trivial solutions to the system of equations like in the previous case, where $\pi = \text{const}$, $\mathcal{R} = \text{const}$ and $\mathcal{F} = 0$.

Now, we analyse how $\mathcal{R}$ and π are related by a gauge transformation. Let us consider the general coordinate transformation:

$$x^\mu \to \hat{x^\mu} = x^\mu + \xi^\mu. \tag{C.31}$$

[2]The ADM formalism is a Hamiltonian formulation of general relativity.

We expand equations (C.29) and (C.30) in terms of $\mathcal{R}$ and π. For simplicity, we restrict to the single field case, where we use $\delta\phi(t,\boldsymbol{x}) \equiv \phi(t,\boldsymbol{x}) - \phi_0(t)$ and $h_{ij} \equiv a^2(t)(1+2\varphi)\delta_{ij}$:

$$\delta\phi(t,\mathbf{x}) = \dot{\phi}_0\pi + \frac{\ddot{\phi}_0}{2}\pi^2 + ..., \tag{C.32}$$

$$\varphi(t,\mathbf{x}) = H\pi + \mathcal{R} + \left(H^2 + \frac{\dot{H}}{2}\right)\pi^2 + 2H\pi\mathcal{R} + \mathcal{R}^2 + ..., \tag{C.33}$$

Transforming t according to equation (C.31), so that $t \to \hat{t} = t + \xi^0$, we observe that equations (C.32) and (C.33) transform up to first order as

$$\hat{\delta\phi} = \dot{\phi}_0(\pi - \xi^0) \qquad\qquad \hat{\varphi} = H(\pi - \xi^0) + \mathcal{R}, \tag{C.34}$$

where we can see that π is the Goldstone boson associated to the spontaneous time translational symmetry breaking $\hat{\pi} = \pi - \xi^0$. If we choose to work on a flat slicing, the spatial part of the metric will remain unperturbed so that $\varphi = 0$. Thus, from equation (C.34) we realize that,

$$\xi^0 = \pi + \frac{\mathcal{R}}{H} \tag{C.35}$$

and therefore, we arrive at the relation between π and $\mathcal{R}$,

$$\hat{\pi} = -\frac{\mathcal{R}}{H}. \tag{C.36}$$

C.3.2 EFT action for π

To obtain the effective action for the generalized framework, we use the flat gauge described above. In this gauge, π is the dynamical field of interest but is not longer constant. In this framework, equation (C.30) becomes,

$$h_{ij} = a^2(t)\delta_{ij}. \tag{C.37}$$

Introducing equations (C.29) and (C.37) into the action (C.1), we obtain

$$\begin{aligned} S_{eff} = -\int d^4x a^3 \frac{N}{2}\Big\{ &-\frac{6H^2}{N^2} + \frac{4H}{N^2}N^i_{,i} + \frac{1}{2N^2}\left(N^i_{,j}N^j_{,i} + \delta_{ij}N^{i,k}N^j_{,k} - 2N^i_{,i}N^j_{,j}\right) \\ &+ \frac{1}{N^2}(\dot{\phi}_0 + \Omega\mathcal{F})^2\left[(1+\dot{\pi} - N^i\pi_{,i})^2 - \frac{N^2}{a^2}(\nabla\pi)^2\right] \\ &\frac{1}{N^2}(\dot{\mathcal{F}} - N^i\mathcal{F}_{,i})^2 - \frac{(\nabla\mathcal{F})^2}{a^2} - 2V(\phi_0^a + N^a\mathcal{F})\Big\}. \end{aligned} \tag{C.38}$$

Calculating the equation of motion and ignoring from them the kinetic terms of $\mathcal{F}$ as we did before, we find a relation between $\mathcal{F}$ and π:

$$\mathcal{F} = \frac{\gamma\dot{\phi}_0\Omega}{M_{eff}^2 - \gamma\Omega}, \tag{C.39}$$

$$\gamma \equiv \frac{1}{N^2} - 1 - \frac{1}{N^2}\left[2(1+\dot{\pi}-N^i\pi_{,i}) + (1+\dot{\pi}-N^i\pi_{,i})^2 - N^2\frac{(\nabla\pi)^2}{a^2}\right]. \tag{C.40}$$

Introducing the obtained expression for $\mathcal{F}$, equation (C.39), back into the action (C.38) and ignoring the kinetic terms of $\mathcal{F}$ in the action, we obtain

$$S_{eff} = \int d^x \frac{a^3}{2}\dot{\phi}_0^2\left[\dot{\pi} - \frac{(\nabla\pi)^2}{a^2} + \frac{(c_s^{-2}-1)[2\dot{\pi}+\dot{\pi}^2-(\nabla\pi)^2/a^2]^2}{4-2\dot{\pi}+\dot{\pi}^2-(\nabla\pi)^2/a^2}\right]. \tag{C.41}$$

To compare the action (C.41) with the action (3.5) of chapter 3, we expand (C.41) up to third order in π and use the relation $\dot{H} = -\dot{\phi}_0^2/2$,

$$S_{eff} = -\int d^4x a^3\dot{H}\Big\{c_s^{-2}\dot{\pi}^2 - \frac{(\nabla\pi)^2}{a^2} + (c_s^{-2}-1)\dot{\pi}\left[\dot{\pi}^2 - \frac{(\nabla\pi)^2}{a^2}\right] + (c_s^{-2}-1)^2\frac{\dot{\pi}^3}{2} - 2\frac{\dot{c}_s}{c_s^3}\pi\dot{\pi}^2 - 2H\eta_{\|}\pi\left[c_s^{-2}\dot{\pi}^2 - \frac{(\nabla\pi)^2}{a^2}\right]\Big\}. \tag{C.42}$$

If we neglect the higher order slow-roll corrections from action (C.42), we finally obtain the EFT action of chapter 3.

$$S = \int d^4x a^3\epsilon H^2\left[-\frac{\dot{\pi}^2}{c_s^2} + \frac{(\partial_i\pi)^2}{a^2} - 2Hsc_s^{-2}\pi\dot{\pi}^2 - (1-c_s^{-2})\dot{\pi}\left(\frac{\dot{\pi}^2}{c_s^2} - \frac{(\partial_i\pi)^2}{a^2}\right)\right]; \tag{C.43}$$

Acknowledgements

So grateful I feel with Ana, who has taught me not only so many lessons related with researching, but has taught me to have a mind prepared to start to carry out research myself. Firstly, because of the opportunity of pursuing a one year research project, with all the benefits this has brought for my "learning curve". Secondly, for giving me the invaluable chance to attend schools and visit Jesús at the University of Sussex. Thirdly, for all the priceless advice regarding research and presentations. And finally, but not least, for making me go through *"the minimum amount of possible stress finding a PhD"*, and allowing me to stay here, with you, with the group and with the research line.

A Matteo, grazie mille per l'aiuto che mi hai dato ed il tempo che mi hai dedicato: ho davvero apprezzato che tu ti sia unito a me in questa avventura. Durante il lavoro ti sei dimostrato il miglior esempio di ottimismo e buon spirito che potessi avere e mi hai insegnato molto.

Gracias Jesús, por enseñarme y trabajar con ilusión en lo que ha sido, y es, la extensión a tu doctorado. También te agradezco haberme dado la oportunidad de ser la cobaya de COBAYA.

Special thanks to the cosmology group in Leiden, which helped me to prepare for crucial moments. To my friends from the *Matt Damon office*, and to the ones of my own office (Fré, Lizzy, Elena), thank you for sharing those enjoyable moments while working and studying together. Gracias especiales van a mi grupo de amigos hispano-hablantes (Marco, Felipe, Jorge y Juan Manuel) por molestarme cuando trabajaba y obligarme a hacer descansos.

For my partner, soul-mate and personal statistician, Juan, I have a huge *thanks* for teaching me basics of statistics and checking with precision and mathematical coherence this thesis. I still do not believe that, after choosing different professional paths, we are coming back to the origins when we used to work together and learning from each other.

A vosotros, papá y mamá, agradeceros cada segundo de esfuerzo dedicado a hacerme feliz. Ya está. Ya me habéis hecho la mejor persona adulta que podríais esperar. Ahora dedicaré mis esfuerzos a haceros felices yo a vosotros: ¡disfrutad!

To Alessandra: if you have got here, thank you so much for being my second reader.

References

[1] (2015). 2015 CMB spectra and likelihood code. https://wiki.cosmos.esa.int/planckpla2015/index.php/CMB_spectrum_%26_Likelihood_Code. Accessed: 29-06-2018.

[2] (2015). Calculadora Cosmologica: code that solves Friedman Eqs. numerically given values for components of universe. https://github.com/gcanasherrera/CalculadoraCosmologica. Accessed: 27-06-2018.

[3] (2015). CMB map reconstruction using Planck + WMAP. http://www.cosmostat.org/derpicated/research-topics-old/cmb. Accessed: 18-04-2018.

[4] (2015). DataCamp: Learn Data Science Online. https://www.datacamp.com. Accessed: 24-04-2018.

[5] (2018). Avenues OPEN: Big Bang. https://open.avenues.org/elements/big-bang. Accessed: 19-01-2018.

[6] (2018). Cosmo21 2018: "Statistical Challenges in 21st Century Cosmology". http://cosmo21.cosmostat.org/programme/. Accessed: 29-05-2018.

[7] (2018). Github Repository of "Data Bases and Data Mining in Astronomy". https://github.com/gcanasherrera/DDM2017.git. Accessed: 10-05-2018.

[8] (2018). LAMBDA - Data Products. https://lambda.gsfc.nasa.gov/graphics/. Accessed: 2018-05-01.

[9] (2018). Scikit-learn 0.19.1 documentation: Gaussian Processes. http://scikit-learn.org/stable/modules/gaussian_process.html. Accessed: 10-02-2018.

[10] Achúcarro, A., Atal, V., Céspedes, S., Gong, J.-O., Palma, G. A., and Patil, S. P. (2012a). Heavy fields, reduced speeds of sound, and decoupling during inflation. Phys. Rev. D, 86(12):121301.

[11] Achúcarro, A., Atal, V., Hu, B., Ortiz, P., and Torrado, J. (2014a). Inflation with moderately sharp features in the speed of sound: Generalized slow roll and in-in formalism for power spectrum and bispectrum. Phys. Rev. D, 90(2):023511.

[12] Achúcarro, A., Atal, V., Ortiz, P., and Torrado, J. (2014b). Localized correlated features in the CMB power spectrum and primordial bispectrum from a transient reduction in the speed of sound. Phys. Rev. D, 89(10):103006.

[13] Achúcarro, A., Gong, J.-O., Hardeman, S., Palma, G. A., and Patil, S. P. (2011). Features of heavy physics in the CMB power spectrum. J. Cosmology Astropart. Phys., 1:030.

[14] Achúcarro, A., Gong, J.-O., Hardeman, S., Palma, G. A., and Patil, S. P. (2012b). Effective theories of single field inflation when heavy fields matter. *Journal of High Energy Physics*, 5:66.

[15] Achúcarro, A., Gong, J.-O., Palma, G. A., and Patil, S. P. (2013). Correlating features in the primordial spectra. Phys. Rev. D, 87(12):121301.

[16] Achúcarro, A. and Welling, Y. (2015). Multiple Field Inflation and Signatures of Heavy Physics in the CMB. *ArXiv e-prints*.

[17] Atal, V. (2016). On multifield inflation, adiabaticity and the speed of sound of the curvature perturbations. *Casimir PhD Series*.

[18] Audren, B., Lesgourgues, J., Benabed, K., and Prunet, S. (2013). Conservative Constraints on Early Cosmology: an illustration of the Monte Python cosmological parameter inference code. *JCAP*, 1302:001.

[19] Baumann, D. (2009). TASI Lectures on Inflation. *ArXiv e-prints*.

[20] Bennett, C. L., Larson, D., Weiland, J. L., Jarosik, N., Hinshaw, G., Odegard, N., Smith, K. M., Hill, R. S., Gold, B., Halpern, M., Komatsu, E., Nolta, M. R., Page, L., Spergel, D. N., Wollack, E., Dunkley, J., Kogut, A., Limon, M., Meyer, S. S., Tucker, G. S., and Wright, E. L. (2013). Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. ApJS, 208:20.

[21] Blas, D., Lesgourgues, J., and Tram, T. (2011). The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes. J. Cosmology Astropart. Phys., 7:034.

[22] Carroll, B. W. and Ostlie, D. A. (2014). *An Introduction to Modern Astrophysics*. Pearson, second edition.

[23] Céspedes, S., Atal, V., and Palma, G. A. (2012). On the importance of heavy fields during inflation. J. Cosmology Astropart. Phys., 5:008.

[24] Cheung, C., Fitzpatrick, A. L., Kaplan, J., Senatore, L., and Creminelli, P. (2008). The effective field theory of inflation. *Journal of High Energy Physics*, 3:014–014.

[25] Chluba, J., Hamann, J., and Patil, S. P. (2015). Features and new physical scales in primordial observables: Theory and observation. *International Journal of Modern Physics D*, 24:1530023.

[26] Cyburt, R. H., Fields, B. D., and Olive, K. A. (2003). Primordial nucleosynthesis in light of WMAP. *Physics Letters B*, 567:227–234.

[27] Dodelson, S. (2003). *Modern cosmology*. Academic Press, San Diego, CA.

[28] Easson, D. A. and Powell, B. A. (2011). Identifying the Inflaton with Primordial Gravitational Waves. *Physical Review Letters*, 106(19):191302.

[29] Feroz, F., Hobson, M. P., and Bridges, M. (2009). MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics. MNRAS, 398:1601–1614.

[30] Handley, W. J., Hobson, M. P., and Lasenby, A. N. (2015a). POLYCHORD: nested sampling for cosmology. MNRAS, 450:L61–L65.

[31] Handley, W. J., Hobson, M. P., and Lasenby, A. N. (2015b). POLYCHORD: next-generation nested sampling. MNRAS, 453:4384–4398.

[32] Howlett, C., Lewis, A., Hall, A., and Challinor, A. (2012). CMB power spectrum parameter degeneracies in the era of precision cosmology. *JCAP*, 1204:027.

[33] Hu, B. and Torrado, J. (2015). Searching for primordial localized features with CMB and LSS spectra. Phys. Rev. D, 91(6):064039.

[34] Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. *Computing In Science & Engineering*, 9(3):90–95.

[35] Ivezic, Z., Connolly, A. J., VanderPlas, J. T., and Gray, A. (2014). *Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data*. Princeton University Press, Princeton, NJ, USA.

[36] Jones, E., Oliphant, T., Peterson, P., et al. (2001–). SciPy: Open source scientific tools for Python. [Online; accessed <today>].

[37] Keldysh, L. V. (1964). Diagram technique for nonequilibrium processes. *Zh. Eksp. Teor. Fiz.*, 47:1515–1527. [Sov. Phys. JETP20,1018(1965)].

[38] Leclercq, F., Pisani, A., and Wandelt, B. D. (2014). Cosmology: from theory to data, from data to theory. *ArXiv e-prints*.

[39] Lesgourgues, J. (2018). Cosmo Tools 2018: A school designed to help young researchers understand the black box of cosmology codes. https://indico.cern.ch/event/688110/timetable/?print=1&view=standard. Accessed: 2018-05-18.

[40] Lewis, A. (2013). Efficient sampling of fast and slow cosmological parameters. *Phys. Rev.*, D87(10):103529.

[41] Lewis, A. (2018). GetDist: MCMC sample analysis, kernel densities, plotting, and GUI. https://github.com/cmbant/getdist. Accessed: 2018-06-18.

[42] Lewis, A. and Bridle, S. (2002a). Cosmological parameters from CMB and other data: A Monte Carlo approach. *Phys. Rev.*, D66:103511.

[43] Lewis, A. and Bridle, S. (2002b). Cosmological parameters from CMB and other data: A Monte Carlo approach. *Phys. Rev.*, D66:103511.

[44] Lewis, A., Challinor, A., and Lasenby, A. (2000). Efficient computation of CMB anisotropies in closed FRW models. *Astrophys. J.*, 538:473–476.

[45] Maldacena, J. (2003). Non-gaussian features of primordial fluctuations in single field inflationary models. *Journal of High Energy Physics*, 5:013.

[46] McKinney, W. (2010). Data structures for statistical computing in python. In van der Walt, S. and Millman, J., editors, *Proceedings of the 9th Python in Science Conference*, pages 51 – 56.

[47] Mukhanov, V. F. (1988). Quantum Theory of Gauge Invariant Cosmological Perturbations. *Sov. Phys. JETP*, 67:1297–1302. [Zh. Eksp. Teor. Fiz.94N7,1(1988)].

[48] Neal, R. M. (2005). Taking Bigger Metropolis Steps by Dragging Fast Variables. *ArXiv Mathematics e-prints*.

[49] Penzias, A. A. and Wilson, R. W. (1965). A Measurement of Excess Antenna Temperature at 4080 Mc/s. ApJ, 142:419–421.

[50] Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Arroja, F., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., and et al. (2016a). Planck 2015 results. XX. Constraints on inflation. A&A, 594:A20.

[51] Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Bartlett, J. G., and et al. (2016b). Planck 2015 results. XIII. Cosmological parameters. A&A, 594:A13.

[52] Planck Collaboration, Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Bartlett, J. G., Bartolo, N., and et al. (2016c). Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters. A&A, 594:A11.

[53] Plato and Lee, H. D. P. (1974). *The Republic*. Penguin - Harmondsworth.

[54] Rasmussen, C. E. and Williams, C. K. I. (2005). *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press.

[55] Riemer-Sorensen, S. (2018). LCDM and Beyond: Cosmology Tools in Theory and in Practice: "Statistics and model selection in cosmology". http://icg.port.ac.uk/~jschewts/cantata/L5/Statistics_Notes.pdf. Accessed: 2018-05-18.

[56] Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Schmidt, B. P., Schommer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B., and Tonry, J. (1998). Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. AJ, 116:1009–1038.

[57] Skilling, J. (2006). Nested sampling for general bayesian computation. *Bayesian Anal.*, 1(4):833–859.

[58] Torrado, J., Hu, B., and Achúcarro, A. (2017). Robust predictions for an oscillatory bispectrum in Planck 2015 data from transient reductions in the speed of sound of the inflaton. Phys. Rev. D, 96(8):083515.

[59] Torrado, J. and Lewis, A. (2018). cobaya, a code for Bayesian analysis in Cosmology. https://github.com/JesusTorrado/cobaya. Accessed: 2018-06-18.

[60] Torrado Cacho, J. (2015). A search for transient reductions in the speed of sound of the inflaton in cosmological data, and other topics. *Casimir PhD Series*.

[61] Weinberg, S. (2005). Quantum contributions to cosmological correlations. Phys. Rev. D, 72(4):043514.