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

BAYESIAN INFERENCE FOR PRECISION COSMOLOGY: A MACHINE LEARNING APPROACH

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

DANIELA IGNACIA GRANDÓN SILVA

PROFESOR GUÍA:
DOMENICO SAPONE

MIEMBROS DE LA COMISIÓN:
NELSON PADILLA
GONZALO PALMA QUILODRÁN
CRISTOBAL SIFÓN ANDALAFT

Este trabajo ha sido parcialmente financiado por ANID Doctorado Nacional/2019-21191886 - Fondecyt Regular N. 1200171 - Universidad de Chile - L'Oréal-UNESCO For Women in Science

SANTIAGO DE CHILE
2023

RESUMEN DE LA TESIS PARA OPTAR
AL GRADO DE DOCTORA EN CIENCIAS, MENCIÓN FÍSICA
POR: DANIELA IGNACIA GRANDÓN SILVA
FECHA: 2023
PROF. GUÍA: DOMENICO SAPONE

INFERENCIA BAYESIANA PARA LA COSMOLOGÍA DE PRECISIÓN: UN ENFOQUE DESDE APRENDIZAJE DE MÁQUINAS

La cosmología se encuentra actualmente en el estado afortunado de recopilar cada vez más datos y, en particular, datos cada vez más precisos. Esto, a su vez, implica que la estimación de numerosos parámetros cosmológicos será cada vez más precisa. Lo anterior ofrece el potencial para una comprensión más profunda de la naturaleza de la energía oscura y sobre la importancia estadística de las tensiones observadas en el modelo cosmológico estándar, conocido como Λ*Cold Dark Matter* (ΛCDM). Con el fin de prepararse para la llegada de estos datos tan precisos, una variedad métodos estadísticos y de *machine learning* (redes neuronales, procesos gaussianos, entre otros) se han empleado como ayuda para procesar los datos y acelerar la inferencia estadística de parámetros cosmológicos. Implementar estas herramientas permite explorar diferentes estrategias para el análisis de datos, como el estudio de efectos instrumentales y físicos que pueden introducir errores sistemáticos y la exploración de modelos más allá de ΛCDM.

En esta tesis, se emplea por primera vez un método estadístico robusto que permite hacer inferencia estadística en el caso que las evaluaciones de la función de verosimilitud o *likelihood* de los datos sean asistidas por redes neuronales artificiales. Dado que las redes neuronales son capaces de *aproximar* las cantidades aprendidas, un problema no resuelto es cómo evitar que el error de aproximación introduzca errores en la inferencia de parámetros. Nuestro método (Bayesiano) provee la solución a este problema, que consiste en una nueva función de probabilidad *a posteriori* que elimina cualquier riesgo de que el resultado de los parámetros cosmológicos esté sesgado por errores introducidos por las herramientas de machine learning. Por otro lado, y siguiendo el estudio de los sesgos en la inferencia de parámetros cosmológicos, esta tesis presenta un estudio de los sesgos que podrían introducir los efectos del feedback de bariones, los cuales corresponden a procesos astrofísicos que pueden contaminar la señal del efecto lente gravitacional débil. Para esto, se emplean datos simulados y reales del *survey* de galaxias Hyper Suprime-Cam (HSC) y se implementan herramientas de machine learning que son entrenadas en base a simulaciones de N-cuerpos. Logramos demostrar que el análisis en base a datos simulados introduce sesgos severos en la inferencia del parámetro S_8. Sin embargo, el análisis en base a datos reales es consistente con un escenario donde los bariones poseen un mecanismo de feedback débil, y por ende, no introducen sesgos importantes en la inferencia de parámetros cosmológicos para los surveys de galaxias actuales, como HSC.

Finalmente, en esta tesis se estudian métodos paramétricos y no paramétricos con el fin de develar las propiedades de la energía oscura, y entender si esta cantidad existe en forma de constante cosmológica o un fluido dinámico. Para este análisis, usamos un set de datos observacionales del Universo tardío, y distintos *priors* en el parámetro de Hubble H_0. Encontramos que distintos métodos convergen a una energía oscura dinámica, aunque su significancia estadística no es suficiente para poder descartar completamente un modelo con constante cosmológica.

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE DOCTORA EN CIENCIAS, MENCIÓN FÍSICA
POR: DANIELA IGNACIA GRANDÓN SILVA
FECHA: 2023
PROF. GUÍA: DOMENICO SAPONE

BAYESIAN INFERENCE FOR PRECISION COSMOLOGY: A MACHINE LEARNING APPROACH

Cosmology is currently experiencing a fortunate growth in the collection of increasingly precise data, yielding sub-percent precision on numerous cosmological parameters. Such constraining power offers the potential for a deeper understanding of the nature of dark energy and sheds light on the statistical significance of tensions observed in the standard cosmological model, known as Λ Cold Dark Matter (ΛCDM). In order to get ready for the arrival of such precise data, a variety of statistical methods and machine learning techniques have been applied to assist the data processing and accelerate the statistical inference of cosmological parameters. These tools enable the exploration of multiple analysis strategies, including investigating systematic effects that may impact parameter inference and exploring cosmological models beyond ΛCDM.

In this thesis, we implement for the first time a robust statistical method that enables the inference of cosmological parameters when the likelihood evaluations are assisted by artificial neural networks. As neural networks are, by design, an approximation of some target function, a far unsolved problem is how to avoid that the neural net's approximation error will bias the parameter inference. We propose an adapted posterior function, which is the Bayesian solution to the wish of wanting to eliminate a neural net's bias and propagate the uncertainty of its prediction. We demonstrate our method in a cosmological analysis and show that our solution works successfully for both well-trained and poorly-trained neural networks. Following the study of biases in cosmological inference, we also examine the bias introduced by baryonic processes in the weak gravitational lensing signal. These are astrophysical processes that can contaminate weak lensing data vectors and lead to bias in our cosmological parameter estimates. To investigate this, we employ weak lensing mocks that follow the Hyper Suprime-Cam (HSC) Year 1 properties. The mocks are based on cosmological N-body simulations and hydrodynamical simulations, using the N-body simulations to build a likelihood pipeline and train a Gaussian Process regressor that learns the statistics for arbitrary cosmologies. We demonstrate that baryonic feedback mechanisms can introduce severe biases in the amplitude of matter clustering parameter S_8 for a simulation-based analysis. We also perform statistical inference on HSC-Y1 real data and demonstrate that a weak feedback scenario is preferred by the HSC data.

Finally, in this thesis we study dark energy through two approaches: a parametric method and a non-parametric method based on Gaussian Processes, with the aim of determining whether this exotic component is consistent with a cosmological constant or a dynamical fluid. We use late-time cosmological data and consider different priors on the Hubble parameter H_0. We found that both approaches employed in this work support a deviation from standard cosmology through an evolving dark energy component, although its statistical significance is not enough to rule out the cosmological constant.

Cuando emprendas tu viaje a Itaca
pide que el camino sea largo
Que muchas sean las mañanas de verano
en que llegues -¡con qué placer y alegría!-
a puertos nunca antes vistos.
Konstantino Kavafis, Ithaca (1911)

Agradecimientos

Quiero agradecer primero a mi supervisor de Doctorado, Domenico Sapone. Su visión científica ha sido fundamental para guiar mis pasos en esta etapa, como también el espacio entregado para poder desarrollar mis intereses en cosmología. No menos importante, agradezco el constante aliento en cada desafío presentado en estos 4.5 años de Doctorado. Al comité de postgrado del Departamento de Física por su ayuda en momentos turbulentos donde la burocracia se tornaba más complicada de lo normal. Especialmente, agradezco a Luis Foa, Gonzalo Palma y a Claudia Urrutia por toda la ayuda brindada. A los miembros de la comisión: Gonzalo Palma, Nelson Padilla y Cristóbal Sifón, por su disposición en aceptar y revisar mi trabajo.

A Nikolaus Vogt, Amelia Bayo, Víctor Cárdenas y Alfredo Vega. Por las inumerables lecciones que me han dejado durante estos años, las cuales me han inspirado y marcado positivamente. De eso y más, estoy muy agradecida.

Mis agradecimientos también hacia Elena Sellentin, quien compartió conmigo su conocimiento en astroestadística durante mi estadía en Leiden y me retroalimento de manera constante durante este proceso. A Jia Liu, por su guía y colaboración en todos los aspectos de mi carrera científica. Me siento en deuda con ustedes por todas las oportunidades brindadas.

Agradezco las fructíferas conversaciones que pude tener con distintos expertos/as en el área: Julien Lesgourgues (y todo el grupo en Aachen), Michelle Lochner, Elisa Chisari, Joachim Harnois-Deraps, David Camarena, Samuel Lepe, Ken Osato, Xiangchong Li y Raúl Angulo. A Javier Silva por el apoyo durante mi aventura con las escaleras holandesas, a Gabriela Marques por el trabajo en equipo y su excelente disposición a contestar todas mis preguntas, a Leander Thiele por su compañia en las frustraciones con los datos durante nuestra estadía en Japón, y a Reggie por el trabajo codo a codo.

Esta tesis no pude haber sido posible sin el financimiento de la Agencia Nacional de Investigación y Desarrollo de Chile ANID-Doctorado Nacional -2019-2119188. Agradezco también al staff de la Universidad de Leiden y Kavli IPMU en Tokyo por la hospitalidad durante mis estadías, y por el apoyo financiero para poder presentar mi investigación en distintas instancias.

Finalmente, agradezco a mi familia, mi Totín y amigos por el amor incondicional a lo largo de los años.

Table of Content

List of Tables

List of Figures

Introduction

Our understanding of the Universe has undergone dramatic advances during the last century. This rapid progress is attributed to the development of physical theories that aim to describe the thermal history of the Universe, as well as the rapid advancement of technology that has enabled the collection of numerous observational probes. Since the publication of Albert Einstein's Theory of General Relativity in 1916 [11], many observational efforts have been dedicated to testing and confirming the predictions of this theory. In 1919, an expedition led by Sir Arthur Eddington confirmed the predictions of General Relativity for the first time, specifically in the deflection of light by the gravitational field of the Sun. In the subsequent years, measurements of the time-delay effect in the Solar System further corroborated the theory. Even in the past few decades, the direct detection of gravitational waves by the LIGO detectors [12] has also provided confirmation of the predictions of this theory. Given the precision of all these observations, General Relativity stands as one of the fundamental pieces in our current understanding of the dynamics of the Universe and its structures. During the 20th century, other major breakthrough discoveries took place. The detection of the cosmic microwave background (CMB) radiation in the early 1960s and the evidence of accelerated expansion, thanks to Type-Ia supernovae observations [13, 14], initiated the era of *Modern Cosmology*. The current concordance model of cosmology, the Λ Cold Dark Matter (Λ CDM) model, has been established and precisely determined based on observations of the Cosmic Microwave Background (CMB) at microwave and infrared frequencies [2], the large-scale structure probes [15], among others. From these observations, we know that 5% of the energy-content in the Universe is composed by ordinary matter and radiation, described by the standard model of particle physics. However, the energy-budget is dominated in a 95% by two exotic and unknown components: *dark matter* and *dark energy*. The former is crucial to describe the formation of structures in the Universe, while the latter is considered as the source of the accelerated expansion. To elucidate the properties and the nature of these components, and also characterize the expansion history of the Universe, cosmology is preparing for the advent of a new era of experiments. Ground-based galaxy surveys such as Rubin Observatory's Legacy Survey of Space and Time (LSST) [15], and space missions such as the Euclid satellite [16], will provide measurements of the matter distribution with unprecedented precision. Furthermore, galaxy surveys will provide large volumes of data, conducting research into the era of precision cosmology and big-data. As for the study of General Relativity during the past century, we hope a new *data-driven* era will shed on light on the current cosmological paradigm. In order to prepare for the analysis of these data sets, the implementation of novel machine learning techniques have been considered as a promising tool to maximize the scientific-return of cosmological data.

In this thesis, we investigate how statistical tools and machine learning techniques aim to provide further insights into the current challenges of modern cosmology, such as biases introduced in parameter inference due to important aspects of data analysis in cosmology: the role of machine learning in cosmology, and the role of systematicsc in the study of weak gravitational lensing.

In this context, neural networks have become popular to accelerate parameter inferences in cosmology, especially for the upcoming generation of galaxy surveys. Many cosmology-emulators based on neural networks have been designed for the analysis of the upcoming Euclid and LSST data [17, 18, 19]. However, it is known that neural networks emulators are only capable of approximate physical functions, and hence a recurrent question has been how to propagate the approximation error of neural networks, in order to avoid biases in the parameter inference. In this thesis, we present a Bayesian solution to propagating a neural net's approximation error and thereby debiasing parameter inference. We capture the reported approximation errors, allowing us to eliminate the neural net's bias during inference, and propagating its uncertainties. We demonstrate that our method is quickly implemented and successfully infers cosmological parameters even for strongly biased neural nets.

The precision and volume of data provided by the upcoming galaxy surveys not only necessitates the incorporation of novel methods to improve the process of parameter inference, but also demands a detailed study of the systematic errors that can easily contaminate this precise cosmological measurements. The study of the weak gravitational lensing of galaxies, i.e. the bending of light due to intervining mass between the source and the observer, involves the statistical detection of a coherent weak signal imprinted in the galaxy shapes. It has been demonstrated that at small scales, astrophysical effects can impact the weak lensing signal and lead to biased cosmological constraints if not accounted for [20, 21]. Moreover, these astrophysical effects can be mimicked by varying cosmological parameters, such as the neutrino mass [21]. Therefore, an analysis of the impact of astrophysical effects such as baryonic feedback, is an important consideration currently in galaxy survey collaborations. In this thesis, we investigate the effects of baryonic physics on weak lensing measurements by employing statistical tools that go beyond the traditional power spectrum. This approach offers several benefits. Different statistical tools are expected to be influenced by systematics in distinct ways. Moreover, the constraining power of summary statistics varies, making it essential to determine which summary statistics can yield the tightest constraints on cosmological parameters while remaining robust against systematics. Rubin LSST [15] is currently developing different strategies to account for the effects of baryons. Since the Hyper-Suprime Cam (HSC) depth is 1 magnitude less depth than that of Rubin-LSST Year 10 [22] and possess an excellent image quality, studying the systematic effects on HSC data is important to understand the potential challenges that LSST may face in the future. It is noteworthy that machine learning methods are gaining popularity in the analysis of weak lensing data. These methods range from deep learning techniques that directly extract cosmological information from weak lensing maps [23] to the training of Gaussian Process or Neural Network emulators for modeling non-Gaussian statistics. The latter is implemented in this thesis as well.

The content of this thesis is organized as follows: In Chapter 1 we provide an overview of the standard model of cosmology at the background and matter perturbation level. We discuss some of the cosmological probes that have placed constraints on the model parameters

and address the challenges that the current cosmological paradigm is still encountering. In Chapter 2 we present the elements of statistics and data analysis that are used in the following chapters, encompassing descriptive statistics, statistical inference and machine learning techniques. Chapter 3 reviews the weak gravitational lensing of galaxies due to the large-scale structure in the Universe. Chapter 4, Chapter 5, and Chapter 6 present the contributions of this PhD thesis. Chapter 4 and Chapter 5 aim to understand the biases in parameter inference. The results of our paper Grandón et al. (2022) [24] are presented in Chapter 4, where a novel Bayesian formalism provides the framework for the statistical inference in cosmology based on artificial neural networks. Chapter 5 is based on Grandón et al. (2023) [25] and Marques et al. (incl. Grandón) (2023) [26]. It shows the result of the study of non-Gaussian statistics and baryonic physics in a weak lensing survey, the Subaru Hyper-Suprime Cam survey [27]. In Chapter 6, based on Grandón et al. (2021) [28] and Bernardo, Grandón, Levi, Cárdenas (2021) [29] we investigate the dynamics of dark energy from a phenomenological approach, using parametric and non-parametric analyses of late-time cosmological data. As these methods are inherently different in how they analyze the data, the fact that similar evolution for the dark energy densities are found suggests that the dynamics is driven by the data, and is not an artifact of the reconstruction methods. Finally, our concluding remarks and future research directions are presented in Chapter 6.7.

Chapter 1

The standard model of cosmology

Our current understanding of the Universe is summarized in the standard model of cosmology, the ΛCDM model, whose foundations are based on the Theory of General Relativity published by Albert Einstein in 1916 [11]. The ΛCDM model has emerged as the concordance model of cosmology as it successfully fits most of the cosmological observations. Nevertheless, in order to fit those observations, ΛCDM incorporates three phenomenological, intriguing and yet unknown mechanisms: *cold dark matter* (CDM), *dark energy* in the form of the cosmological constant Λ and *inflation*. Revealing the true nature of these components constitutes the forefront of ongoing research in this field, and we aim to understand better their properties by designing new ground-based and space-based cosmological experiments.

In this chapter we provide an overview of the ΛCDM model, alongside the presentation of essential tools needed for the rest of this thesis. At the end of this chapter, we also highlight the open questions that challenge the current cosmological paradigm. The concepts studied in this chapter are based on the textbooks [30, 31, 32] and *lecture notes* [33], where a detailed derivation of all the equations can be found.

1.1 Einstein's field equations

General Relativity (GR) describes gravity as a property of a dynamical and curved space-time instead of an external gravitational force as assumed in classical Newtonian gravity. Hence, particles do not move in a gravitational field, instead they are free-falling in a curved space-time.
The fundamental element that encodes the effects of gravity is the metric tensor $g_{\mu\nu}(t, \vec{x})$, which generalizes the Minkowski metric from special relativity to curved space-times. It also allows us to compute intervals between two events in a curved geometry through the line element $\mathrm{d}s^2 \equiv g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$. The dynamics of the Universe come up by solving the Einstein field equations[1]

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} \,, \tag{1.1}$$

[1] speed of light is set to 1, $c = 1$.

where G is the Newton gravitational constant. In the left-hand side (l.h.s.), the Einstein tensor $G_{\mu\nu}$ describes the geometry of space-time through partial derivatives of the metric tensor encoded in the Ricci tensor $R_{\mu\nu}$ and Ricci scalar R. The right-hand side (r.h.s.) contains the energy-momentum tensor $T_{\mu\nu}$ and describes the properties of the matter components in the Universe as

$$T_{\mu\nu} = (p + \rho)u_\mu u_\nu + p g_{\mu\nu} + \Sigma_{\mu\nu}, \tag{1.2}$$

with u^μ is the four-velocity of the fluid in comoving coordinates. The pressure and energy density of the fluid are represented by p and ρ, respectively, and $\Sigma_{\mu\nu}$ is the viscous stresses. This tensor incorporates the constituents in the Universe that source the curvature of space-time.
The Einstein field equations correspond to ten coupled and non-linear partial differential equations, in four independent variables $(t, \vec{x})$. The metric tensor $g_{\mu\nu}$ is a solution to Eq. (1.1), which can be obtained by assuming highly symmetric space-times with simple forms of $T_{\mu\nu}$, such as the spherically symmetric vacuum solution known as Schwarzschild metric for black holes. In what follows, we are interested in the *Friedmann-Lemaitre-Robertson-Walker* (FLRW) metric to study the evolution and dynamics of the Universe. This is presented in the next section.

1.2 The homogeneous and isotropic Universe

The central premise of the standard model of cosmology is that the Universe is statistically isotropic and homogeneous at large scales (> 100 Mpc). This is called the *cosmological principle* and is strongly supported by astronomical observations. Isotropy means that the Universe looks the same in all directions from Earth position. As we do not occupy any privileged position in the Universe, consequently we obtain spatial homogeneity[2].
The standard model pose another pinnacle since 1929. Edwin Hubble [34], based on the measurements of the distance and radial velocities of galaxies, discovered that the Universe is expanding and the expansion rate can be derived from a linear relation between the recession velocity and distance. Therefore, the distance between two events increases with time. Years later, in 1998, two independent collaborations discovered that the expansion of the Universe is accelerating [13, 14]. The cosmological principle and the expansion of the Universe constitute the fundamental basis of our description of the Universe through GR.

For a Universe undergoing accelerated expansion and obeying the cosmological principle, the most general solution to Eq. (1.1) is the *Friedmann-Lemaitre-Robertson-Walker* (FLRW) metric [35, 36, 37, 38] in spherical coordinates

$$\begin{aligned} \mathrm{d}s^2 &= -\mathrm{d}t^2 + a^2(t)\left(\frac{\mathrm{d}r^2}{1 - Kr^2} + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2\right)\right), \\ \mathrm{d}s^2 &= -\mathrm{d}t^2 + a^2(t)\left[\mathrm{d}\chi^2 + S_k^2(\chi)\mathrm{d}\Omega^2\right], \end{aligned} \tag{1.3}$$

[2]Although this principle is robust against observations, there is still on-going research testing small departures from it.

with the differential of solid angle $\mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2$. The function $a(t)$ is called scale factor and accounts for the cosmic expansion as function of cosmic time[3] t, which is normalized to $a = 1$ today. The radial coordinate χ is

$$\chi = \begin{cases} \sin^{-1}(r) & K = +1 \\ r & K = 0 \\ \sinh^{-1}(r) & K = -1 \end{cases}, \tag{1.4}$$

and K take one of the values $-1, 0, 1$ describing an open, flat or closed geometry, respectively. The value $K = 0$ corresponds to Euclidean geometry, and is strongly favoured by astronomical observations.

In order to describe the expansion of the Universe and study how $a(t)$ evolves, we need to determine the r.h.s. of Eq. (1.1). In the FLRW Universe, the energy-momentum tensor describes the matter content as a perfect fluid as seen by comoving observers, and is given by

$$T_{\mu\nu} = \begin{pmatrix} -\rho(t) & 0 & 0 & 0 \\ 0 & p(t) & 0 & 0 \\ 0 & 0 & p(t) & 0 \\ 0 & 0 & 0 & p(t) \end{pmatrix}, \tag{1.5}$$

where ρ and p refer to the total energy density and total pressure, respectively. We can derive the equations of motions in a FLRW Universe by solving the Einstein field equations for components $i, j = 0$ and spatial $i = j$ (in this metric, the other components vanish) in light of Eq. (1.5). The solutions are known as the Friedmann-equations

$$\left(\frac{\dot{a}}{a}\right)^2 \equiv H^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2}, \tag{1.6}$$

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

where the dot denotes derivative with respect to cosmic time t, and H is the Hubble function

$$H(t) = \frac{\dot{a}(t)}{a(t)}, \tag{1.8}$$

describing the expansion rate of the Universe. The key idea of GR is presented in these equations for the FLRW Universe: the evolution of the scale factor and hence the expansion is determined by the energy density ρ. It is common to denote the Hubble constant today as H_0, and define $H = 100\,h$ km Mpc^{-1}s^{-1} where h encodes our uncertainty on H. In order to learn about the dynamics of the energy components, we can take the covariant derivative of the energy-momentum tensor $\nabla_\mu T^\mu_\nu = 0$ or combine Eq. (1.6) with Eq. (1.7) to obtain

$$\frac{\partial \rho}{\partial t} + 3H(\rho + p) = 0, \tag{1.9}$$

[3]The cosmic time refers to the time measured by comoving observers that move with the cosmic flow.

which defines the continuity equation for a perfect fluid, where the pressure and energy density can be related through the equation of state

$$w = \frac{p}{\rho}. \tag{1.10}$$

Solving Eq. (1.9) tells us how the energy density of the components in the Universe evolves with time. For a perfect fluid with time-independent equation of state, integrating Eq. Eq. (1.9) leads to

$$\rho = \rho_0 a^{-3(1+w)}, \tag{1.11}$$

where ρ_0 is the present-day energy density for a certain specie. Among the components that play a role through the evolution of cosmic history, we find radiation (relativistic particles like photons), non-relativistic matter (baryonic matter, dark matter, massive neutrinos) and energy with negative pressure such as the cosmological constant. The energy density of these constituents dilutes with the expansion as

$$\rho = \begin{cases} \rho_r^0 a^{-4} & w = 1/3 \quad (\text{Radiation}) \\ \\ \rho_m^0 a^{-3} & w = 0 \quad (\text{Matter}) \\ \\ \rho_\Lambda^0 & w = -1 \quad (\text{Cosmological constant}). \end{cases} \tag{1.12}$$

In turn, inserting Eq. (1.11) into Eq. (1.6), and assuming the Universe is dominated by one of the energy species with equation of state w, the scale factor evolves with time as

$$\begin{cases} a \propto t^{\frac{2}{3}(1+w)} & w \neq -1 \\ \\ a \propto \mathrm{e}^{Ht} & w = -1. \end{cases} \tag{1.13}$$

Therefore, for $w = -1$ the scale factor evolves exponentially, which leads to an important remark: by combining both Friedmann Eqs. (1.6) and (1.7) we obtain the acceleration equation

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) = -\frac{4\pi G}{3}\rho(1 + 3w), \tag{1.14}$$

from which we conclude that cosmic acceleration $\ddot{a} > 0$ occurs if the dominant component exhibits an equation of state $w < -1/3$. The component that fullfills this condition is referred to as *dark energy*, whilst the particular case $w = -1$ is known as the cosmological constant. In light of the components listed above, the Friedmann equation (1.6) recast as

$$H^2 = \frac{8\pi G}{3}(\rho_m + \rho_r + \rho_\Lambda). \tag{1.15}$$

We can split the total density into these components, i.e. matter, radiation and Λ. To do so, it is convenient to define the relative abundances of each component through the dimensionless parameter

$$\Omega_{i,0} = \frac{\rho_{i,0}}{\rho_{cr}}, \qquad \rho_{cr} = \frac{3H_0^2}{8\pi G} = 1.88h^2 \times 10^{-29}\text{g cm}^{-3}, \tag{1.16}$$

where $\rho_{i,0}$ is the density of any particular component today $a = 1$, and ρ_{cr} is the average density in the Universe, also called *critical density*. In terms of the density parameters Ω, Eq. (1.6) reads as

$$\sum_i \Omega_i + \Omega_K + \Omega_\Lambda = 1, \tag{1.17}$$

where Ω_i corresponds to

$$\Omega_i \equiv \frac{8\pi G\rho_i}{3H^2}\,, \quad \Omega_\Lambda = \frac{\Lambda}{3H_0}, \tag{1.18}$$

and the normalized density parameter of curvature Ω_K

$$\Omega_K \equiv -\frac{K}{(aH)^2}\,. \tag{1.19}$$

The components of the Universe and their relative abundance define the dynamics of the expansion. They are determined by the cosmological parameters

- H_0: Expansion rate today,
- Ω_m: Energy density of total matter,
- Ω_K: Energy density of curvature,
- Ω_r: Energy density of radiation,
- w: Equation of state for dark energy,
- Ω_{DE}: Energy density of dark energy (or Ω_Λ when $w = -1$).

The goal of current research on cosmology is to obtain tight constraints to these parameters based on local observations and observations of photons from the early Universe. Knowing the cosmological parameters $\{\Omega_{0,i}, H_0, w\}$, we can reconstruct the expansion history of the Universe, which starts with an initial period of expansion called *inflation*, then a radiation-dominated era made of a thermal bath of relativistic particles (until $\sim$ 50.000 years after inflation [31]), followed by the matter-dominated era for 10 billion years. In the present epoch, dark energy plays a dominant role, and it is postulated as the source of the accelerated expansion. The latter, in turn, affects the measurement of distances in the Universe. Hence, in order to elucidate more about the properties of dark energy it is important to introduce the concept of distances in cosmology.

1.3 Redshift

As already mentioned, the observational evidence strongly supports an expanding Universe. This means that at early epochs, the distance between objects in the Universe was smaller than it is today. At the same footing, the electromagnetic waves emitted from distant galaxies are stretched due to the expansion, proportional to the scale factor. Therefore, the observed wavelength is larger than the emitted wavelength, leading to the lose of energy of the photon and making emission/absorption lines in the spectra to be shifted. This shift due to expansion is called *cosmological redshift*, or redshift hereinafter [4]. It is defined as

$$1 + z = \frac{\lambda_{obs}}{\lambda_{\rm e}} = \frac{a_{obs}}{a_{\rm e}}. \tag{1.20}$$

where $\lambda_{\rm e}$ stands for the emitted wavelength (as would be observed near the object in space and time) and λ_o the observed value. If we recall that $a(t_0) = 1$, the relation between scale factor and redshift is given by

[4]Concepts of blueshift or redshift are commonly associated to Doppler effect. Here, we have a special case as the redshift is due to the recession velocity of distant galaxies, driven by cosmic expansion.

$$1 + z = \frac{1}{a_{\rm e}}, \tag{1.21}$$

and present epochs corresponds to $z = 0$. With the definition of cosmological redshift, the Friedmann equation (1.15) reads as

$$H^2(z) = H_0^2 \left(\Omega_{m,0}(1+z)^3 + \Omega_{r,0}(1+z)^4 + \Omega_{K,0}(1+z)^2 + \Omega_{\Lambda,0}\right) \tag{1.22}$$

with

$$\Omega_{m,0} + \Omega_{r,0} + \Omega_{K,0} + \Omega_{DE,0} = 1. \tag{1.23}$$

1.4 Cosmic Distances

We can measure distances in terms of comoving coordinates, fixed in an expanding grid; and physical coordinates, whose distance increase with the expansion. Photons travel in null geodesics $\mathrm{d}s^2 = 0$, thus the metric reads as $\mathrm{d}r^2 = \mathrm{d}t^2/a(t)$ (recall we set speed of light $c = 1$). Then, the re-scaling of a physical distance by the scale factor gives the comoving coordinate. By taking the definition $\mathrm{d}t = \mathrm{d}a/(aH)$, we can write $\mathrm{d}r = \mathrm{d}a/(a^2H)$. These definitions lead to the derivation of *comoving distance*, which is the distance to an object whose light was emitted at t_1

$$\chi = \int \mathrm{d}r = \int \frac{\mathrm{d}t}{a(t)} = \int_{a(t_1)}^{a(t_0)=1} \frac{\mathrm{d}a'}{a'^2 H(a')}. \tag{1.24}$$

By manipulating through the equation (1.21), the comoving distance for a flat Universe reads as

$$\begin{aligned} \chi(z) &= \int_0^z \frac{\mathrm{d}z'}{H(z')} \\ &= \frac{1}{H_0} \int_0^z \frac{\mathrm{d}z'}{\sqrt{\Omega_m^0(1+z')^3 + \Omega_r^0(1+z')^4 + \Omega_{DE}^0}}. \end{aligned} \tag{1.25}$$

If we consider the distance that light travels since $t = 0$, and assuming the ideal case that light travels without further interactions, we obtain the *comoving horizon*

$$\eta = \int_0^t \frac{\mathrm{d}t'}{a(t')}, \tag{1.26}$$

which defines the maximum distance for causal connection, as no information could have travelled further than light in a comoving grid. It is important to note that we actually observe fluxes emitted by distant objects, and separation angles on the sky. Hence, we need to define two more concepts: *angular diameter distance* and *luminosity distance*.

1.4.1 Angular diameter distance

In astronomy, a common way to determine distances is to consider an object which physical size on the sky is l and the subtended angle is θ. If the angle θ is small, it is possible to estimate its angular diameter distance as

$$D_A \equiv \frac{l}{\theta}. \tag{1.27}$$

As the comoving size of the object is given by l/a (with a the scale factor), the comoving distance to that object χ will be

$$\chi = \frac{l/a}{\theta}, \tag{1.28}$$

and thus the comoving distance and angular diameter distance are related by the redshift/scale factor

$$D_a = \frac{\chi}{1+z}, \tag{1.29}$$

for a flat Universe. In the case of $K \neq 0$, the angular diameter distance generalizes to [30]

$$D_a(z) = \frac{1}{1+z}\frac{1}{H_0\sqrt{-\Omega_K}}\sin\left(\sqrt{-\Omega_K}H_0\chi\right). \tag{1.30}$$

1.4.2 Luminosity Distance

The luminosity L is related to the flux F through the luminosity distance D_L

$$f = \frac{L}{4\pi D_L^2}, \tag{1.31}$$

where L is the absolute luminosity of the emission source, and it emits in spherical shells of surface $4\pi D_L$. For a flat Universe, the projected surface is a sphere with area $A = 4\pi r^2$, where r is the radius $r = a\chi$. For an expanding Universe, this relation depends on the scale factor

$$f = \frac{L(a)}{4\pi\chi^2(a)}, \tag{1.32}$$

where the area is described by a comoving spherical shell with radius $\chi(a)$. Since the energy of the photon decreases by a factor of a due to expansion, $L(a)$ is also reduced by a. Another effect to take into account is the fact that the two photons emmited within an interval $\Delta t_{\rm e}$, will reach our detectors in an interval $\Delta t_0 = a\Delta t_{\rm e}$. Accordingly, their physical separation at arrival to the telescope is larger than the separation when emmited, due to the expansion. This produces

$$L_0 = L_{\rm e}/(1+z)^2, \tag{1.33}$$

$$f = \frac{L}{4\pi\chi^2(1+z)^2}, \tag{1.34}$$

and hence

$$D_L = D_A(z)(1+z)^2. \tag{1.35}$$

The luminosity distance is useful in situations when the intrinsic luminosity of the source is known, and the observed dimming can be used to measure the distance. These kind of sources are named *standard candles*, such as Supernovae Ia, whose light curve can be calibrated and standardized in order to unveil the cosmic expansion.

1.5 Density fluctuations and matter power spectrum

The cosmological principle is strongly supported by observations and only holds at very large scales. Below these scales, we observe that the Universe is rich of smaller structures that perturb such principle. Moreover, galaxies do not form randomly across space, instead they form in clusters and filaments that make up the so called *cosmic web*. Hence, structures in our Universe must emerge from initial perturbations, and consequently grow via gravitational instability. To describe a perturbed Universe that lead to structure formation, we consider departures from the FRLW metric using linear perturbation theory. A detailed and thorough derivation of the theory of linear perturbations and the solution to the Einstein-Boltzmann coupled equations is presented in [30, 33]. We refer the reader to the aforementioned references for further details.

Let us start by defining the density contrast. Perturbations are regions with overdensities of matter respect to the average amount of matter, and regions with underdensities of matter. The matter density contrast describes the density fluctuations at position $\mathbf{r}$ and time t with respect to the mean matter $\bar{\rho}_m(t)$ density of all space as:

$$\delta(\mathbf{r},t) = \frac{\delta\rho(\mathbf{r},t)}{\bar{\rho}_m(t)} = \frac{\rho_m(\mathbf{r},t) - \bar{\rho}_m(t)}{\bar{\rho_m}(t)}, \tag{1.36}$$

with

$$\begin{cases} \delta < 0 & \text{Underdense region} \\ \delta > 0 & \text{Overdense region.} \end{cases} \tag{1.37}$$

In linear theory, this perturbation is considered small $|\delta(\mathbf{r},t)| \ll 1$. The density contrast is a random variable and we want a statistical description of it. The first moment is the average, and it is given by

$$\langle \delta(\mathbf{r}) \rangle = 0, \tag{1.38}$$

where the angular brackets is the average over space, and is equal to 0 as we do not expect any preferential position in the sky. More information is contained in its second moment, the variance

$$\langle \delta(\mathbf{r})^2 \rangle = \sigma^2. \tag{1.39}$$

The two point correlation between under(over)dense regions in different positions in the sky also captures the variance information. In real-space, it is given by

$$\xi\left(\vec{x},\vec{y}\right) = \langle \delta\left(\vec{x}\right)\delta\left(\vec{y}\right)\rangle = \xi(|\vec{x}-\vec{y}|). \tag{1.40}$$

Following the cosmological principle, the correlation function ξ is not sensitive to the exact positions and hence only depends on the distance difference between both locations $|\vec{x}-\vec{y}|$. Due to this translational invariance, it is convenient to work with the Fourier modes of the density fluctuations. The Fourier transform $\delta(\vec{k})$ is connected to $\delta(\vec{x})$ in 3D as

$$\delta\left(\vec{x}\right) = \int \frac{\mathrm{d}^3k}{(2\pi)^3}\, \mathrm{e}^{i\vec{k}\cdot\vec{x}}\tilde{\delta}\big(\vec{k}\big)\,, \quad \tilde{\delta}(\vec{k}) = \int \mathrm{d}^3x\, \mathrm{e}^{-i\vec{k}\cdot\vec{x}}\delta\big(\vec{x}\big). \tag{1.41}$$

The Fourier transform of the two-point correlation function is

$$\left\langle \delta(\vec{k})\delta^*(\vec{k}\,')\right\rangle = \int \mathrm{d}^3x \int \mathrm{d}^3y\, \langle \delta\left(\vec{x}\right)\delta\left(\vec{y}\right)\rangle\, \mathrm{e}^{-i\vec{k}\cdot\vec{x}}\, \mathrm{e}^{i\vec{k}\cdot\vec{y}\,'}\,, \tag{1.42}$$

replacing Eq. (1.40) in Eq. (1.42) we obtain

$$\left\langle \delta(\vec{k})\delta^*(\vec{k}\,')\right\rangle = \int \mathrm{d}^3x \int \mathrm{d}^3y\, \xi\left(\vec{x}-\vec{y}\right)\, \mathrm{e}^{-i\vec{k}\cdot\vec{x}}\, \mathrm{e}^{i\vec{k}\cdot\vec{y}'}\,, \tag{1.43}$$

Defining the variable $\vec{w} = \vec{y} - \vec{x}$, the Fourier transform gives the power spectrum $P_\delta(\vec{k})$

$$\begin{aligned}\left\langle \delta(\vec{k})\delta^*(\vec{k}\,')\right\rangle &= \int \mathrm{d}^3x \int \mathrm{e}^{-i\left(\vec{k}-\vec{k}\,'\right)\cdot\vec{x}}\, \mathrm{d}^3w\, \xi\left(\vec{w}\right) \mathrm{e}^{i\vec{k}\,'\cdot\vec{w}} \\ &= (2\pi)^3\, \delta_D^3\left(\vec{k}-\vec{k}\,'\right) P_\delta\left(\vec{k}\right),\end{aligned} \tag{1.44}$$

where the Dirac delta function $\delta_D^3\left(\vec{k}-\vec{k}\,'\right)$ implies that different Fourier modes k are independent. Therefore, the power spectrum is

$$P_\delta\left(\vec{k}\,\right) \equiv \int \mathrm{d}^3x\, \xi\left(\vec{x}\right) \mathrm{e}^{-i\vec{k}\cdot\vec{x}}\,, \tag{1.45}$$

which is derived by considering the two point function, and hence it is connected to the variance. It tells us how the variance of the random field is distributed over spatial frequency [33]. As the field is isotropic, it only depends on the modulus $P_\delta\left(\vec{k}\,\right) = P_\delta(k)$. Physically, regions with high density contrast means regions with a stronger clustering of matter. The power spectrum will capture this effect by producing larger power in scales where these overdensities take place.
In practice, the power spectrum is computed on smoothed or filtered fields, where the smoothing kernel is depicted by $W(\vec{r})$[5]. The smoothed field is a convolution of the field by the kernel

$$\delta(\vec{x}) = \int \mathrm{d}^3r W(\vec{r})\delta(\vec{x}+\vec{r}), \tag{1.46}$$

and the variance is given by

$$\sigma_R^2 = \frac{1}{2\pi^2}\int P(k)|W(kR)|^2k^2\mathrm{d}k. \tag{1.47}$$

We can take advantage of this definition to compute the variance of density fluctuation in a smoothed field at a given scale. In this particular case, the kernel or window function is a spherical top-hat given by

$$W_R(r)\begin{cases} \frac{3}{4\pi R^3} & r < R \\ 0 & r > R. \end{cases} \tag{1.48}$$

Its Fourier transform corresponds to $\tilde{W}(kR) = 3j_1(kR)/(kR)$ where $j_1(kR)$ is the spherical Bessel function of the first kind (and order one). For a radius equal to the scale of massive galaxy clusters $R = 8\mathrm{Mpc/h}$ we define

$$\sigma_8^2 = \frac{1}{2\pi^2}\int P(k)\left[\frac{3j_1(kR)}{kR}\right]^2 k^2\mathrm{d}k. \tag{1.49}$$

[5]For example, the point spread function of the optical system.

the root mean square (rms) of the amplitude of matter perturbations smoothed over $8\,\mathrm{Mpc}/h$, at present time. The derived parameter σ_8^2 is also the normalization of the matter power spectrum and it is constrained by cosmological probes.
We can derive a theoretical expression for power spectrum of matter, which depends on cosmological parameters. To do so, we need to study the evolution of δ_m. But first, we can generally say that the matter power spectrum depends on initial conditions coming from the very early Universe through the *primordial power spectrum* $P_o(k)$, a *transfer function* $T(k)$ that links the initial conditions to the distribution of matter we observe today and a *growth function* $D(a)$ that depends on the redshift. The main quantity that encodes the physics behind perturbations is the transfer function. It describes the evolution of density perturbations $\delta(k)$ inside the horizon according to the dynamics during the radiation-dominated and matter-dominated eras. The expression for the matter power spectrum reads as

$$P_m(k,z) \propto P_o(k)T^2(k)D(a), \tag{1.50}$$

where the details about $T(k)$, $D(a)$ and $P_o(k)$ are explained below. To derive this expression, let us keep studying the density contrast. In the linear regime, it evolves according to the second order differential equation

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} = 4\pi G\bar{\rho}_m\delta, \tag{1.51}$$

which holds for sub-horizon scales. The right-hand side is the source of perturbation growth due to gravitational instability. The term $2H\delta$ in the left-hand side represents a friction term, in this case sourced by the expansion of space-time through the scale factor. The solution to this equation depends on the dominant matter specie in the Universe.

- For a radiation-dominated Universe, the contribution from the r.h.s. is negligible as the dominant component is radiation. Hence, the solution to $\ddot{\delta} + 2H\dot{\delta} = 0$ is

  $$\delta(t) = A + B\ln t, \tag{1.52}$$

  and perturbations will grow logarithmically, i.e. very slowly. This is reasonable as radiation pressure prevents matter to agglomerate.
- In the matter-dominated regime, the solution corresponds to

  $$\delta(t) \propto a(t), \tag{1.53}$$

  hence δ grows linearly to the scale factor.
- In the Λ-dominated regime, the solution is

  $$\delta(t) = A + B\exp(-2Ht), \tag{1.54}$$

  where the second term is a decaying solution that does not contribute as it falls quickly. The first term tells us that perturbations do not grow as dark energy suppress the formation of structures.

A general solution for a Universe filled with multiple ingredients includes a growing and decaying solution. We are interested in the growing modes, as decaying modes become negligible. The general solution reads as

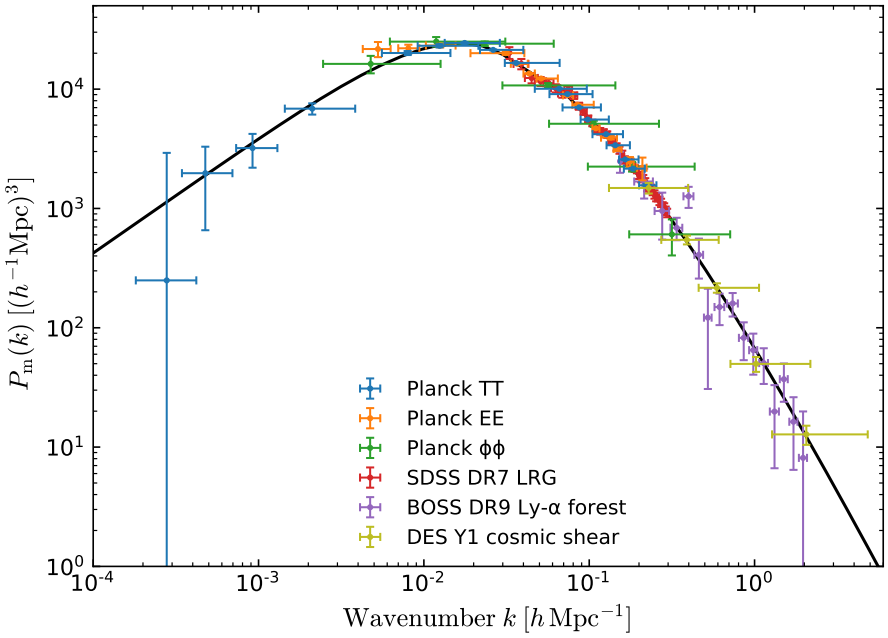


Figure 1.1: Linear matter power spectrum as a function of wavenumber k. Black solid line depicts the theory prediction given *Planck* 2018 cosmology at $z = 0$. Coloured points are compiled observations made by ground-based surveys collaborations such as SDSS and BOSS. Blue points are *Planck* 2018 results. Image original from [3].

$$\delta(\mathbf{r}, t) = D_+(t)\delta_+(\mathbf{r}) + D_-(t)\delta_-(\mathbf{r}) \tag{1.55}$$

where $D_+(t)$ is the scale-independent growth factor and modulates how the density contrast grows with the scale factor. This quantity is important as it describes the growth of structures mainly at late times. For a flat matter-dominated Universe, the growth factor is $D(a) = a$ [30]. When dark energy dominates, the growth decays. Moreover, it is defined as

$$D_+(a) = \frac{5\Omega_m}{2}\frac{H(a)}{2H_0}\int_0^a \frac{{\rm d}a'}{a'^3}\left(\frac{H_0}{H(a')}\right)^3 \tag{1.56}$$

considering dark energy in the form of a cosmological constant. Defining the growth factor in terms of the growth supression factor $g(a)$, the dimensionless matter power spectrum is

$$\Delta^2(k, a) = A_s\frac{4}{25}\frac{1}{\Omega_M^2}\left(\frac{k}{k_{piv}}\right)^{n_s-1}\left(\frac{k}{H_0}\right)^4 [ag(a)]^2T^2(k) \tag{1.57}$$

where A_s and n_s are cosmological parameters defined in the primordial power spectrum, and $k_{piv} = 0.05{\rm Mpc}^{-1}$. The key quantity that resumes the results from linear theory is the transfer function $T(k)$. Analytical expressions have been derived for $T(k)$, such as the Bardeen, Bond, Kaiser, and Szalay [39] fitting function

$$T(k) = \frac{\ln(1 + 2.34q)}{2.34q}\left[1 + 3.89q + (16.1q)^2 + (5.46q)^3 + (6.71q)^4\right]^{-1/4}, \tag{1.58}$$

where

$$q \equiv \frac{k}{\Omega_m h^2 Mpc}, \tag{1.59}$$

for ΛCDM. However, in other contexts numerical solutions are needed and thus it is common to run theory codes called *Boltzmann codes*, such as CAMB [40, 41] or CLASS [42] that output the transfer function and the matter power spectrum, among other quantities. These computer codes solve the coupled Einstein-Boltzmann equations for the evolution of perturbations given input cosmological parameters and can be extended to other cosmological models.

The equation (1.57) for the matter power spectrum succesfully describes the large-scale structure in linear theory, for scales > 10 Mpc. At small scales $k > 0.1$ Mpc^{-1}, non-linearities emerge which are better studied with N-body cosmological simulations [43]. Figure 1.1 shows the theoretical matter power spectrum in solid black line for a ΛCDM cosmology, and the measurements based on Planck, the Sloan Digital Sky Survey (SDSS), the Baryon Oscillation Spectroscopic Survey (BOSS) and the Dark Energy Survey (DES). We see a significant agreement between theory and data, which provides one of the many reasons that make the ΛCDM model a successful description of our Universe. We note that the power spectrum goes as $\sim k^{n_s}$ at large scales, and after the turnover it decays with k^{n_s-3}. The turnover happens at the matter-radiation equality with k_{eq} given by [31]

$$k_{\rm eq} = a_{\rm eq} H = H_0 \sqrt{\frac{2\Omega_m^2}{\Omega_r}} \sim 0.02 {\rm Mpc}^{-1}. \tag{1.60}$$

Hence, the matter power spectrum constraints cosmological parameters such as A_s, $\Omega_m h$ and σ_8, with the growth factor carrying information about the dark energy parameters Ω_Λ and w through its dependence on $H(a)$ in Eq. (1.56) (although it can also affect k_{eq}). The matter power spectrum itself can be considered a cosmological probe, as it is sensitive to many cosmological parameters and is one of the key quantities entering into the large-scale structure observables.

1.6 Cosmological probes

1.6.1 Cosmic Microwave Background

The discovery of the cosmic microwave background (CMB) at the early 1960s is a cornerstone of modern cosmology. The CMB observations provide the most precise constraints on cosmological parameters today, obtained from the the *Planck* satellite observations [2]. The study of the CMB started in 1950's, and its measurements resulted in the 1978, 2006 and 2019 Nobel prizes[6].

To understand the physics behind the CMB, we need to study the Universe back when it was 380.000 years old. At early times, the Universe was filled with a hot baryon-photon plasma characterized by photons and baryons tightly coupled. In this hot plasma, scattering

[6] 1978 Nobel prize was awarded to Arno Penzias and Robert Wilson, as they first detected this radiation. Later on, in 2006, John Mather and George Smoot became Nobel Prize laureates for their study of the CMB using the COBE satellite. In 2019, James Peebles obtained the Nobel prize for his work on a theoretical description of the CMB and the large-scale structure.

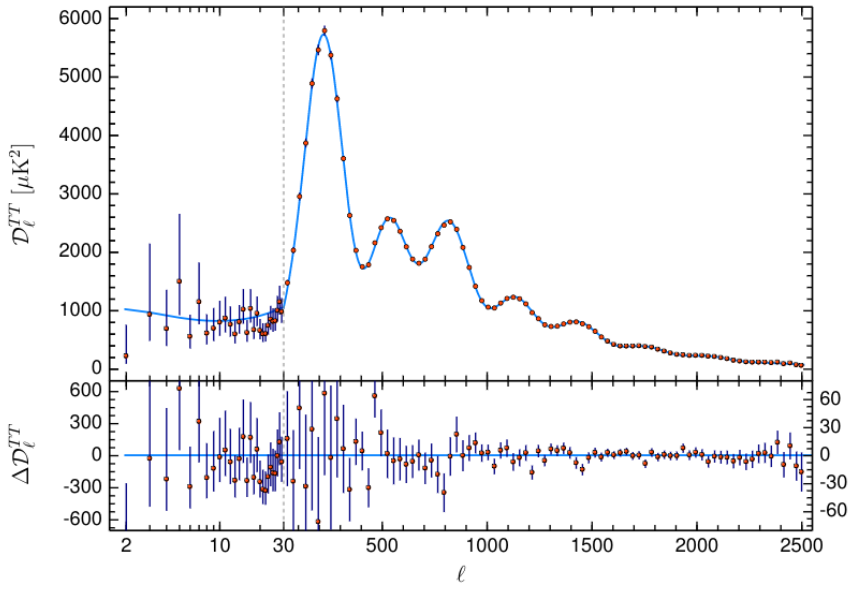


Figure 1.2: *Planck* 2018 CMB temperature power spectrum as a function of multipoles ℓ. Light-blue solid line is the ΛCDM fit to the data, whereas red points depict the data points of the CMB temperature power spectrum and error bars the 1σ Gaussian uncertainties. The lower panel shows residuals respect to a base-ΛCDM model, and the scatter in low-multipoles is due to cosmic variance (statistical uncertainties due to the large angle correlations). Taken from [2].

of photons and electrons via Thomson scattering, and electrons and protons coupled through Coulomb interactions make up a Universe dominated by this one-fluid. The mean free path of the photons was very short, and thus the universe was still opaque. Meanwhile, the Universe continued expanding and cooling.

The interplay between the radiation pressure due to photons, and gravity compressing the plasma generated baryon-photon acoustic waves, where gravity pulling comes from dark matter in the potential wells. At regions where acoustic waves go through the compression phase, the photons become hotter than plasma average temperature, and, conversely, regions undergoing the rarefaction phase, photons get colder. At $z \simeq 1275$, expansion reduces the interaction rate to the point that protons and electrons combine into neutral hydrogen, a process called *recombination*. Photons decoupled from baryons and, at $z \simeq 1090$, they free-stream through space. The Universe rapidly transition from an opaque plasma to become transparent, and the last interaction between photons and baryons make up the last scattering surface[7]. The temperature of the photons at the moment of decoupling was $T \approx 3000$ K. As Universe expands, the temperature of radiation decreases with the scale factor $T(t) = T_0(a = 1)/a(t)$, and thus we observe this relic radiation today in the microwave regime with temperature $T \sim 2.725K$. The decoupling produces that the acoustic waves stop oscillating and got frozen, with matter leaving a density profile of a characteristic scale, imprinted in the matter distribution we see today.

[7]It is important to note that none of these events happen suddenly, rather they lasted for some thousands of years. Recombination is a gradual process, and it is thought to happen from $z \sim 1450$, when 90% of Hydrogen was still ionized, to $z \sim 1060$. Similar situation occurs for the last scattering, as some photons scattered $\Delta z \sim 100$ later.

The Cosmic Background Explorer, also known as COBE (1992), confirmed for the first time that the CMB follows a blackbody spectrum [44], as predicted by the theory. The satellites *Wilkinson Microwave Anisotropy Probe* (WMAP) and more recently *Planck* characterized the CMB properties with further details. The most precise full-sky measurements of the CMB radiation by *Planck* report that CMB average temperature (the monopole) is uniform across the sky, and its value corresponds to

$$T_0 = 2.7255 \pm 0.0006\,\mathrm{K}, \tag{1.61}$$

with tiny fluctuations of order

$$\frac{\Delta T}{T} \simeq 10^{-5}. \tag{1.62}$$

described by the quadrupole moment of the CMB[8].

The effects of the acoustic waves are imprinted in the CMB anisotropies, as illustrated in Fig.1.2. The main observables we extract from the CMB photons observations are the angular power spectrum of temperature anisotropies and the CMB lensing power spectrum. The physics behind the CMB is encoded in the peaks of the angular power spectrum, as shown in Fig. 1.2. The horizontal axis ℓ are the multipoles, with low-ℓ tracing large scales in the sky, and high-ℓ describing small scales. The position ℓ of the first peak is sensitive to spatial curvature, and provides the constraints $\Omega_K = -0.044^{+0.018}_{-0.015}$ consistent with a flat Universe. On the other hand, recalling the acoustic waves, CMB provides rich information on $\Omega_c h^2$ and $\Omega_b h^2$. The constraints are $\Omega_c h^2 = 0.11933 \pm 0.00091$ and $\Omega_b h^2 = 0.02242 \pm 0.00014$, respectively. Hence, the CMB observation directly probes the existence of dark matter, its value, and the non-baryonic nature it poses, as it interacts very weakly with baryons in the acoustic waves. Moreover, the combination of CMB temperature power spectrum and CMB lensing gives tight constraints on matter clustering σ_8 and total matter Ω_m, and better constraints on dark energy can be provided combining CMB and type Ia Supernovae data.
It is worth noting that more experiments are being designed to study the CMB radiation, such as the LiteBIRD satellite[9] [46, 47] and Simons Observatory [48] in the Atacama desert in Chile. Most recent results of the CMB lensing were reported by the Atacama Cosmology Telescope (ACT) collaboration in 2023 [49], whose telescope is also located in Chile and aims to study the CMB itself, galaxy clusters, and provide better constraints in combination with galaxy surveys data [50].

1.6.2 Supernovae Ia (SN Ia)

Consider a binary system containing at least one compact white dwarf star. The white dwarf can interact with its companion star, whether as accreting material from the other, or by merging with other white dwarf. When one of these situations occurs, there will be a point when the white dwarf mass exceeds the Chandrasekhar mass $\sim 1.4\mathrm{M}_\odot$. As a result, the white dwarf will undergo thermonuclear collapse and then explode into a Supernovae Ia (SN Ia). Spectroscopically, SN Ia are detected by the presence of strong silicon lines

[8]The dipole contribution comes from the motion of the Earth respect to the CMB rest frame [45]

[9]LiteBIRD stands for *Lite (Light) satellite for the studies of B-mode polarization and Inflation from cosmic background Radiation Detection*

and no hydrogen in their spectrum. In terms of photometry, SN Ia exhibits a characteristic light curve, with a maximum value and then a rapid decay of its brightness. However, light curves of SN Ia can scatter in their shape, including the peak brightness value, and can be reddened due to intervening dust. In order to promote SN Ia as standard candle, a standarisation of the light curve is needed. In 1993, Phillips [51] observed a relation between the peak luminosity and the initial decline rate in the B band based on observations of the *Tololo Supernova Program* [52]. This effect was later tested and confirmed by studying 32 Supernovaes Ia from the Calán-Tololo survey [52]. In order to correct for this effect, many methods have been proposed, such as the stretch factor [53] or the Δm_{15} correction [54]. All these efforts, including corrections by reddening and host-galaxy contamination [55], allowed to promote SN Ia as *standard candles*. Back in the 1990s, the high-redshift supernova surveys *Supernova Cosmology Project* [14] and *High-z Supernovae Search Team* [13] measured the apparent magnitude of 34 supernovae. Their work ended up in one of the major discoveries in cosmology (whose impact led to the 2011 Nobel Prize). They found that the Universe is undergoing a phase of accelerated expansion, and this acceleration can be explained by assuming an Universe dominated by some exotic fluid, called later as *dark energy* [56]. To understand how SN Ia can be used to constrain cosmology, we need to define a few astronomical concepts. The apparent magnitude is a measure of the apparent brightness of an object as observed in the sky, and it depends on the flux f as

$$m \equiv -2.5\log_{10}\frac{f}{f_x}. \tag{1.63}$$

On the other hand, the absolute magnitude is related to intrinsic luminosity through

$$M \equiv -2.5\log_{10}\frac{L}{L_x}. \tag{1.64}$$

where the subscript x represent some reference value given the photometric filter. The values M, L are assumed to be approximately constant for each object that belongs to the class of *standard candle*. The difference between the apparent magnitude m and absolute magnitude M defines the *distance modulus* μ:

$$\mu \equiv m - M = 2.5\log_{10}\left(\frac{L}{f}\right) \tag{1.65}$$

If we recall the equation 1.31 that relates flux and luminosity, the distance modulus reads as:

$$\begin{aligned}\mu(z_i) \equiv m - M &= 5\log_{10}\left(\frac{D_L(z_i)}{10\text{PC}}\right) \\ &= 5\log_{10}\left(\frac{D_L(z_i)}{1\text{PC}}\right) + 25\end{aligned} \tag{1.66}$$

$$m - M = 5\log_{10}(D_L) + 25. \tag{1.67}$$

Equation 1.67 basically relates the apparent magnitude to the luminosity distance of an object, which in turn depends on the cosmological parameters. Both parameters H_0 and M are degenerated, which means we can not extract information on H_0 without knowing something about M. Using Cepheids stars as calibrators, the absolute magnitude value is

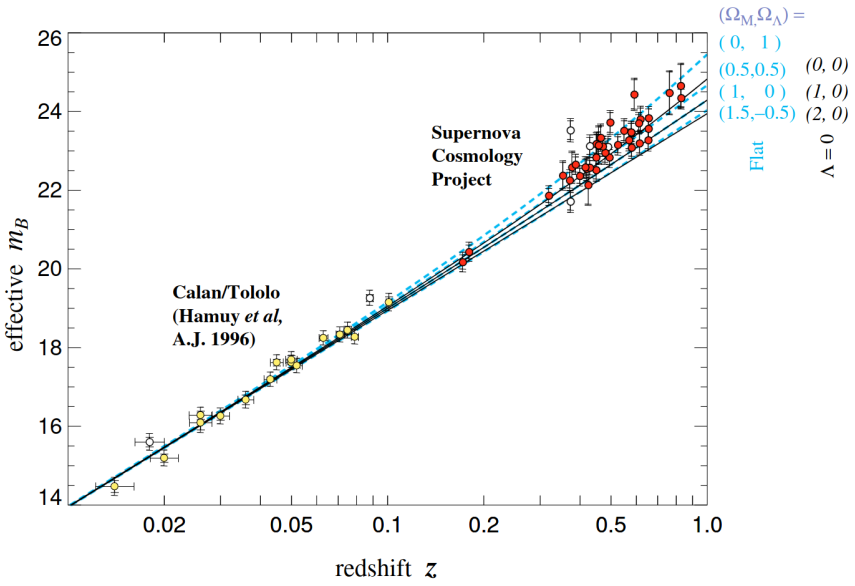


Figure 1.3: The Hubble diagram for high-redshift Supernovae Ia from the *Supernova Cosmology Project* (in red points) and low-redshift data from the Calán/Tololo Supernova Survey (in yellow). Black solid lines are the theoretical predictions for $m_B(z)$ for cosmological models $(\Omega_\Lambda, \Omega_m) = (0, 0)$ (top), $(\Omega_\Lambda, \Omega_m) = (0, 1)$ (middle) and $(\Omega_\Lambda, \Omega_m) = (0, 2)$ (bottom). Dashed light-blue curves are fits based on flat cosmological models $(\Omega_\Lambda, \Omega_m) = (1, 0)$ (top), $(\Omega_\Lambda, \Omega_m) = (0.5, 0.5)$ (second from top), $(\Omega_\Lambda, \Omega_m) = (0, 1)$ (second from bottom) and $(\Omega_\Lambda, \Omega_m) = (-0.5, 1.5)$ (bottom). Error bars are 1σ Gaussian uncertainties. Original plot from Perlmutter *et al.* [4].

estimated to be $M \sim -19.5$ in the B-band. In Fig. 1.3 we see the type Ia Supernovae m_B data from [4], which includes the low-z measurements from the Calán/Tololo survey [54]. In dashed light-blue lines we see the fit of different cosmological models varying Ω_m and Ω_Λ. They found significant evidence of $\Lambda \neq 0$, with best-fit value $\Omega_m = 0.26 \pm 0.08$ and $\Omega_m + \Omega_\Lambda$ for a flat cosmology. Hence, they provide significant evidence of dark energy. In order to determine spatial curvature, an analysis combining CMB+SNIa is needed.

1.6.3 Cosmic Chronometers

Cosmic Chronometers (CC) [57, 5] corresponds to a novel method to reconstruct the expansion history of the Universe by measuring the age difference between *passively-evolving* galaxies. CC does not rely on any integrated quantity, neither on early physics assumptions nor distance ladder calibration. With the minimal assumption of the FLRW Universe, it is able to constraint cosmological parameters through

$$H(z) = \frac{1}{a}\frac{\mathrm{d}a}{\mathrm{d}t} = -\frac{1}{(1+z)}\frac{\Delta z}{\Delta t}. \tag{1.68}$$

where the ratio $\Delta z/\Delta t$ ($\mathrm{d}z/\mathrm{d}t$) is obtained from the spectroscopy of a selected sample of galaxies.
The method procceed as follows. First, an homogeneous and synchronised sample of galaxies in some redshift range is selected, for which tracers of age can be found. The most well-suited

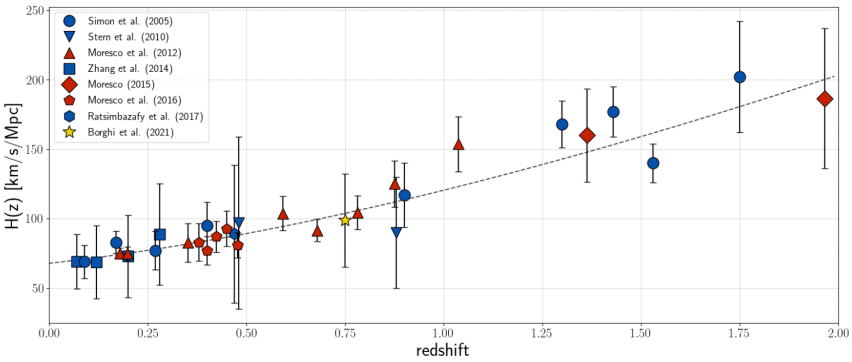


Figure 1.4: Compilation of $H(z)$ values. Each color represent a different differential ages estimation method. Dashed black line is the prediction for a flat ΛCDM model based on *Planck* cosmology. Error bars are 1σ uncertainties. Figure original from [5]

objects are massive and passive galaxies, with old stellar populations. Any hint of ongoing star formation must be carefully addressed as it can lead to biased results, hence the selection of the sample implies analyzing the spectral features of galaxies. Based on models of stellar evolution and chemical composition of stars, it is possible to derive their differential age. These massive and passive galaxies build up their mass during short timescales and at high redshift, exhausting most of their gas reservoir very early in their life and hence evolving very passively as seen from today. In general, measuring dt is challenging and there are many methods to estimate the differential age dt, all based on characterizations of the spectrum. A review of the methods can be found in [5]. For instance, one approach consists in using the spectral line 4000 Å, which is correlated to the age of old stellar populations [58]. This line is a break in the continuum due to absorption lines, whose amplitude is also correlated with metal abundance. Once the metallicity is determined, the dt between two galaxies can be obtained from the difference in their 4000 Å amplitude.

Figure 1.4 shows a compilation of CC $H(z)$ data measured by different collaborations. The potential of $H(z)$ data to provide constraints on H_0, w and Ω_Λ has led to forecasts on simulated data based on the Euclid satellite and the ATLAS future space-mission. Figure 1.5 illustrates results on the $\Omega_m - H_0$, $\Omega_\Lambda - H_0$ and $w - H_0$ parameter planes. As this probe is very sensitive to H_0, it can shed light on the H_0 tension and say something about the current cosmological paradigm.

1.7 ΛCDM model

So far, we have revisited three cosmological probes that are in excellent statistical agreement with the theory, at the expense of the inclusion of three exotic ingredients: dark energy, dark matter and inflation (initial conditions). These components are three, purely phenomenological, pillars of the model. However, they lack an underlying physical theory that explains their nature. Current experiments aim to unveil more about the fundamental physics behind these components, with the ΛCDM model as a simple framework from which we can test departures from it. Moreover, two cosmological parameters tensions have motivated the exploration of extensions to the base ΛCDM model as well.

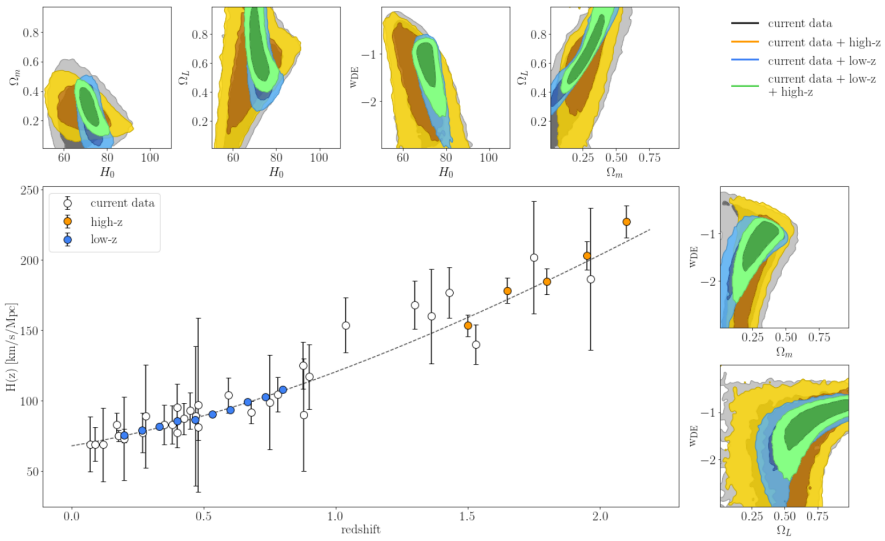


Figure 1.5: Forecast of CC real+mock data. Current data are the compiled H(z) values, depicted in white scatter points. Blue and yellow points represent data and error bars at low redshift (mainly based on SDSS re-analysis) and high redshift, respectively. High-redshift points are simulated data based on Euclid survey and Astrophysics Telescope for Large Area Spectroscopy (ATLAS) probe space-mission. Figure original from [5].

1.7.1 Inflation

Inflation is a period of accelerated exponential expansion at the very early stages of the Universe [59, 60]. In its simplest form, it consists in fluctuations of a single and quantum scalar-field, which seeds the formation of structures that make up our Universe. Hence, we can work on the inverse problem and study cosmic inflation by analyzing the distribution of matter and radiation in the Universe. As reviewed in the previous sections, the standard model of cosmology encompasses differential equations for the dynamics of the Universe whose initial conditions were not originally provided. Moreover, observations of the CMB show that the Universe is spatially flat with high-precision, and it is statistically uniform on large scales (despite the tiny fluctuations). These observations can only be explained by fine-tuned initial conditions. These fine-tunings are known as the *flatness problem* and *horizon problem* [60, 61, 62]. Back in the 1980's, a number of authors [63, 61, 62] introduced inflation for the first time. They showed that mechanism of an accelerated expansion at the very early stages of the Universe, where primordial fluctuations were in causal contact on subhorizon scales, provide the solutions to such problems.
Equally important, inflation successfully provides the initial conditions for structure formation. Primordial fluctuations evolve through cosmic history, and once they enter the horizon, seed the matter perturbations that evolve in the structures observed today. The primordial fluctuations are described by the primordial power spectrum given by

$$P(k) = A_s \left(\frac{k}{k_*}\right)^{n_s - 1}, \tag{1.69}$$

where k_* is a pivot scale, n_s is the spectral index parameter and tells us whether the power spectrum is scale dependent; and A_s is the amplitude of primordial fluctuations.
Inflation imprints its features into the large-scale structure observables, as the primordial

power spectrum is directly connected to the matter power spectrum. It also predicts primordial gravitational waves, which emerge from tensor perturbations in the metric. The primordial tensor modes of inflation induce B-modes in the CMB polarization, and thus the detection of such CMB B-modes would be considered as a strong evidence that inflation occurred [47].

1.7.2 Cold dark matter (CDM)

We have compelling evidence that there is some missing mass in the total matter budget in the Universe. This missing mass component is non-luminous, and hence it is called *dark matter* [64]. Cosmological probes have shown that dark makes up around 85% of the total matter content in the Universe, and today it encompass about 1/4 of the total energy budget.
Probes of dark matter range from small to large scales in the Universe. The first strong evidence of dark matter was obtained by measuring the galaxy rotation curves. Vera C. Rubin [65] showed that the velocity of stars at the outskirts of spiral galaxies deviate from what is predicted by Newtonian gravity

$$v(r) = \sqrt{\frac{GM(r)}{r}}. \tag{1.70}$$

Equation (1.70) tells us that the velocity of a star is expected to behave as $v(r) \approx r^{-1/2}$, however, velocity of stars is roughly constant $v(r) \approx$ const. including at the outskirts. If we assume Newtonian physics to hold, we need the presence of extra matter (which is not seen) in order to match the flat rotation curve. Observations of clusters of galaxies and gravitational lensing also points to the missing mass problem, hence supporting the inclusion of dark matter to the cosmological model.
The strongest evidence of dark matter nowadays is the large-scale structure and the cosmic microwave background radiation. Both probes emerge from an inhomogeneous Universe, where overdense regions are mainly composed by dark matter which drives the structure formation in the Universe through gravitational clustering, as seen in Eq. (1.51) and the matter power spectrum. The dark matter density $\Omega_c h^2$ is tightly constrained by the peaks location and heights of the CMB angular power spectrum. This probe accurately infers that the density of baryonic matter is way smaller than that of cold dark matter $\Omega_b \ll \Omega_c$ and thus dark matter is very likely to be non-baryonic in nature.
The extensive efforts up to date have enabled the characterization of the dark matter mechanism, although its microscopic nature is still unknown. Observations support that dark matter acts as a non-relativistic (hence, it is cold) perfect fluid, with very small pressure $w \approx 0, c_s \approx 0$ and interacting extremely weakly with standard model particles. The most popular candidate are weakly interacting massive particles WIMPs [66], such that they weakly interact with baryons. Thermal relics particle from the early Universe may fulfill this condition and become a WIMP candidate. In order to elucidate the fundamental physics behind DM, the efforts must come from many fronts, including particles physics, galaxy surveys and Milky Way astronomical observations. Dark matter imprints its distinct properties in different environments, and thus the multi-experiments approach is an imperative in the roadmap for the next years of DM searches [67].

1.7.3 Dark energy

What is the physical source that drives the acceleration? This answer is still one of the most intriguing questions in cosmology. The ΛCDM model is determined by dark energy in the form of cosmological constant Λ. Physically, Λ represents vacuum energy density due to the fluctuations sourced by the creation/annihilation of particle-antiparticle pairs in vacuum. The implementation of Λ in Eq. (1.1) takes advantage of the general covariance of Einstein field equations, which allows to add an extra term to this equation. Actually, it was first introduced by Albert Einstein in 1917 to obtain static solutions [68], and later re-introduced after the discovery of the accelerated expansion. The energy density and pressure of Λ are

$$\rho_\Lambda = \frac{\Lambda}{8\pi G} \quad P_\Lambda = -\rho_\Lambda, \tag{1.71}$$

respectively, where we take $w = -1$. This equation of state defines the mechanism that can explain accelerated expansion, as probed early in this chapter. However, the energy density required to drive the accelerated expansion is about $\sim 10^{123}$ smaller than the energy scale predicted by particle physics.

Acceleration can also emerge from a general form of dark energy for which $w < -1/3$ holds. To distinguish between dark energy mechanisms with time-independent energy density, we can consider w as a free parameter. This is often referred to as the wCDM model. If flatness is assumed, the two parameters Ω_m and w can fully characterize the mechanism. Another widely considered model adds another parameter that allows dark energy to vary with time, the so called w_0w_aCDM [69] model, where the equation of state is parametrized as

$$w(z) = w_0 + w_a\left(\frac{z}{1+z}\right), \tag{1.72}$$

and the constrained parameters are Ω_m, w_0 and w_a. For a time-varying equation of state $w(z)$, the normalized energy density is $\rho_{DE}(z)/\rho_{DE}(0)$ and the Hubble function corresponds to

$$H^2(z) = H_0^2\left[\Omega_{m,0}(1+z)^3 + \Omega_{r,0}(1+z)^4 + \Omega_{DE,0}\exp\left(\int_0^z \frac{3(1+w(z'))}{1+z'}\mathrm{d}z'\right)\right]. \tag{1.73}$$

In spite of the nature of dark energy, the measurements of $w(z)$ or $\rho_{DE}(z)$ can be regarded as a phenomenological description of cosmic expansion, as it allows to reconstruct the expansion history through measurements of the luminisity distance (which in turns reconstruct $H(z)/H_0$ and hence w). We can summarize the dark energy parametrizations as

$$\begin{cases} w = -1 & \Lambda, \\ w \neq -1 & w\text{CDM}, \\ w(a) = w_0 + w_a(1-a) & w_0w_a\text{CDM}, \\ \frac{\rho_{DE}(z)}{\rho_{DE}(0)} = \exp\left(3\int_0^z (1+w(z')\mathrm{dln}(1+z'))\right) & \text{XCDM}. \end{cases} \tag{1.74}$$

The dynamical approach can be generalized introducing a spatially homogeneous and dynamical scalar field $\phi(t)$ in the form of a perfect fluid with density and pressure

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi), \tag{1.75}$$

$$P_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi)\,, \tag{1.76}$$

respectively. The field evolves according to the Klein-Gordon equation of motion $\ddot{\phi} + 3H\dot{\phi} + \mathrm{d}V(\phi)/\mathrm{d}\phi = 0$. The equation of state parameter for such scalar field is

$$w_\phi = \frac{P_\phi}{\rho_\phi} = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)}. \tag{1.77}$$

In order to obtain $w < -1/3$, it is required that $\dot{\phi}^2 < V(\phi)$, which means that the scalar-field potential must be shallow enough in order to ϕ slowly vary (or roll) along the potential [70]. These kind of models are referred to as *Quintessence* models and cosmologists propose models for $V(\phi)$ that fulfills the slow-roll condition. As an alternative, modifications of the l.h.s. of Eq. (1.1) are also considered. Up to date, all proposed mechanisms have their own pitfalls and fine-tunings, including the cosmological constant as already explained in this section. For a detailed and meticulous revision of dark energy models, we refer the reader to [70].

1.7.4 Cosmological parameters and model tensions

Cosmological analysis relies on high-dimensional likelihoods. In this context, it is common to encounter parameters degeneracies, which means that the variation of multiple parameters in specific manners can simulate the impact of altering a single parameter. For instance, Ω_K and Ω_Λ are degenerated in the CMB [30]. As multiprobes analysis can break parameter degeneracies and lead to tighter constraints [71], we utilize a variety of datasets in our cosmological analysis. However, different probes have led to discrepant best fits on parameters S_8 and H_0. We proceed to present these tensions in this last part of the chapter.

We have measured the H_0 parameter since the groundbreaking discovery of the expansion by Edwin Hubble in 1929 [34]. He first inferred that the Hubble constant is about $H_0 \sim 500$ km $\mathrm{Mpc}^{-1}s^{-1}$. Decades after, the H_0 constraints ranged from 50 km $\mathrm{Mpc}^{-1}\mathrm{s}^{-1}$ to 100 km $\mathrm{Mpc}^{-1}\mathrm{s}^{-1}$, with variations mainly arising due to statistical errors and calibrations [31]. Nowadays, the inferred H_0 value span a narrower range of values, from ∼67 to ∼74 km $\mathrm{Mpc}^{-1}s^{-1}$. Nonetheless, given the precision achieved with current experiments, a tension on the expansion rate today H_0 has emerged, being established as a major puzzle aimed to be solved nowadays [72]. Observations based on the Cosmic Microwave Background power spectrum point to $H_0 = 67.27 \pm 0.60$ km $\mathrm{Mpc}^{-1}\mathrm{s}^{-1}$, whereas local observations employing the distance ladder approach report $H_0 = 73.5 \pm 1.1$ km $\mathrm{Mpc}^{-1}s^{-1}$ [73], in $\sim 4\sigma$ tension with the CMB. The H_0 values are

$$H_0 = \begin{cases} 67.27 \pm 0.6 & \textit{Planck} \quad \text{CMB} \\ \\ 67.9 \pm 1.5 & \text{ACT CMB} \end{cases} \tag{1.78}$$

$$H_0 = 73.5 \pm 1.1 \quad \text{Cepheids-SNIa.} \tag{1.79}$$

A wide range of low-redshift probes also show disagreement with the CMB constraints at the 4σ level [74, 75, 76]. Among the solutions, several papers have considered departures from

the standard model, such as early dark energy models [77] or dark sector interactions [78]. On the other hand, the calibration using Cepheids stars has led extensively discussion. For instance, in [79] the authors study the impact of metallicity on the period-luminosity relation and analize the recalibration of Cepheids data. This led to $H_0 = 70.6 \pm 3.3$ km Mpc^{-1}s^{-1}, in 1σ tension with *Planck* CMB data.
Up to date, calibrations and systematic errors are still being revisited and no concensus has been reached whether the H_0 tension is a statistical fluke or a hint to new physics [80].

At the level of cosmological perturbations, the S_8 tension has emerged as another matter of scrutinity of the ΛCDM model. Although not as statistically significant as the H_0 tension, the tension has pursuingly been revisited by many independent collaborations.
In the recent years, Stage-III weak lensing surveys such as Dark Energy Survey (DES) [81, 82], Kilo Degree Survey (KiDS) [83], and the Subaru Hyper Suprime-Cam (HSC) [22, 27] have achieved great precision in constraining cosmological parameters Ω_m and σ_8 by performing two-point statistics analysis of galaxy shapes at low-redshift ($z \leq 1.5$). Moreover, current research report constraints on the growth parameter combination $S_8 \equiv \sigma_8 * (\sqrt{\Omega_m/0.3})$. Such surveys have revealed that the inferred S_8 value obtained from weak lensing analysis is in tension with the value reported by the *Planck* CMB power spectrum [2]. The reported values are the following

$$S_8 = \begin{cases} 0.759^{+0.024}_{-0.021} & \text{KiDS-1000 weak lensing} \\ 0.772^{+0.018}_{-0.017} & \text{DES Y3 weak lensing} \\ 0.800^{+0.029}_{-0.028} & \text{HSC Y1 weak lensing} \end{cases} \tag{1.80}$$

$$S_8 = 0.828 \pm 0.016 \quad \textit{Planck} \ \text{CMB}. \tag{1.81}$$

In other words, the inferred amplitude of fluctuations in the low-redshift regime is lower than the value inferred by early time CMB data. KiDS is in a $\sim 2.5\sigma$ level tension with *Planck*, whereas DES Y3 exhibits $\sim 2.4\sigma$ tension.
This has lead intensive investigation on many approaches intended to relieve the S_8 tension [20, 84, 85, 86]. One of them consists in improving the theoretical modelling of the weak lensing signal and understanding the systematical effects that may impact the analysis from catalogues to cosmology. It has been shown [20] that the low S_8 amplitude constraints comes mainly from non-linear scales. Recent works have explored whether the *Planck* result can be reconciled with weak lensing S_8 constraints by altering the matter power spectrum on non-linear scales using DES Y3 and KiDS data. They found that the baryonic feedback[10] may alleviate the tension, however the amplitude of the feedback needed to match *Planck* results is larger than the provided by current state of the art hydrodynamical simulations.
In [87], the authors infer cosmological parameters using DESY3+KiDS-1000 data jointly and found 1.7σ tension with *Planck* for certain modelling choices. Other large-scale structure probes such as cluster counts and redshift space distortions agree with significant lower S_8

[10]Baryonic feedback refers to a wide range of astrophysical processes such as stellar winds, star formation, AGN feedback, among others, that can reshape gas inside haloes and alter the clustering at small scales. This thesis extensively review this effect in Chapter 5.

[75]. Moreover, we are preparing for the Stage-IV data releases such as LSST and Euclid galaxy shapes catalogues that will bring precision cosmology to an absolutely new level and will shed new light on the significance of the S_8 tension. Both surveys will resolve small scales with better precision, hence will contribute significantly to the discussion. Hence, it is still too soon to discard the possibility that the tension may be solved by new physics, such as dark matter candidates whose mechanism can contribute to the suppression of clustering at small scales.

Chapter 2

Elements of statistics and data analysis

Statistics play a key role in data-driven cosmology. The interplay between parametric cosmological models that describe our Universe and astronomical data is mediated by statistical tools that summarize the information contained in data and provide credible values for the model parameters. Moreover, to design an experiment investigating the nature of dark matter and dark energy, statistical methods and metrics are employed to evaluate the future performance of such an experiment. Nevertheless, the complexity of the theoretical models and the precision of observations are continuously increasing, requiring more refined and sophisticated data analysis techniques to handle and decode the physical information in the data. The latter has been made possible in recent years by the increase in computational power.
On the other hand, large-scale structure observables rely on the study of random fields. The cosmic density field of the late Universe was seeded by the primordial density perturbations of the inflaton field, which is quantum in nature and mostly predicted to be a Gaussian random field. Therefore, it is important to note that we cannot access the *true* density field; instead, we observe a random realization of it. This implies that the sophisticated tools used to analyze cosmological data must be incorporated within a framework that includes elements of statistics.
In summary, many aspects of statistics and data analysis come into play. Let us start this chapter by studying random variables and later discuss descriptive statistics in general.

2.1 Random variables and basics of probability

A random variable, denoted by A, is a parameter that exhibits variation as a result of chance. It is drawn from a probability distribution $P(A)$, which is defined over the sample space $\mathcal{S}$: the entire set of possible outcomes. Random variables can be interpreted as events. As such, it is possible to define the probability $P(A, B)$ that two events A and B occur simultaneously and the conditional occurrence of event A given that event B already happened, denoted as $P(A|B)$. Probability theory is founded on the Kolmogorov axioms, [88]:

1. P(A) ≥ 0,
2. P($\mathcal{S}$)=1,

3. $P(A_1 \cup A_2 \cup ... \cup A_n) = P(A_1) + P(A_2) + .. + P(A_n)$ for mutually exclusive events $\{A_j\}$.

From these axioms, we can derive the law for a joint event $P(A, B)$ as $P(A|B)P(B) = P(B|A)P(B)$. This reduces to Bayes' theorem [89]

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \tag{2.1}$$

Bayes' theorem is an important result widely applied in all fields of science, and it is obtained as a simple derivation that follows from probability distribution manipulations. Before studying this theorem in more detail, let us define a few concepts of descriptive statistics.

2.1.1 Moments

Probability distributions can be defined by their moments. The n^{th} moment m_n corresponds to the expectation value defined as

$$m_n = \langle x^n \rangle = \int_{-\infty}^{\infty} x^n P(x) \mathrm{d}x, \tag{2.2}$$

from which the mean μ is

$$\mu = \langle x \rangle = \int_{-\infty}^{\infty} x P(x) \mathrm{d}x. \tag{2.3}$$

We can also define the central moments, which are computed based on deviations from the mean rather than with respect to zero, as

$$m_n = \langle (x - \langle x \rangle)^n \rangle = \int_{-\infty}^{\infty} (x - \mu)^n P(x) \mathrm{d}x. \tag{2.4}$$

The second moment of the distribution $P(x)$ is the variance

$$\mathrm{Var(x)} \equiv \sigma^2 = \langle (x - \mu)^2 \rangle = \int_{-\infty}^{\infty} (x - \mu)^2 P(x) \mathrm{d}x, \tag{2.5}$$

which measures the spread of the distribution around the mean. Naturally, $P(x)$ is not necessarily a symmetric distribution. Hence, higher order moments are needed to capture the asymmetry and other features of the distribution. Third and fourth order moments are the skewness S and kurtosis K [90], defined as

$$S \equiv \left\langle \left(\frac{x - \mu}{\sigma} \right)^3 \right\rangle, \tag{2.6}$$

$$K \equiv \left\langle \left(\frac{x - \mu}{\sigma} \right)^4 \right\rangle, \tag{2.7}$$

and both scaled by σ. For this reason, they are also called *standardized central moments*. To measure the asymmetry of a probability distribution, we compute the skewness S. If S takes a positive value, the distribution exhibits longer right tails and the mean shifts from

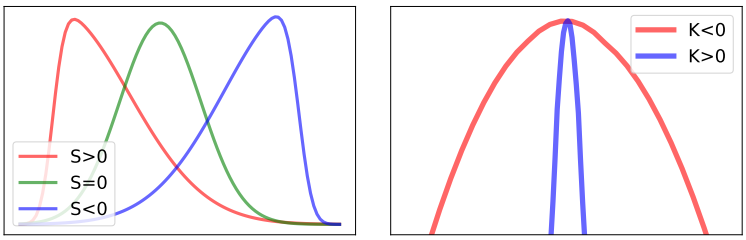


Figure 2.1: *Left:* Example of skewness for three different distributions. Red curve shows positive skewness, blue curve negative skewness and the green curve depicts the Gaussian distribution with vanishing skewness. *Right:* Two distributions showing the effect of high (red) and low kurtosis (blue).

the median. On the contrary, negative skewness means longer left tail and the mass of the distribution shifts to the right. Kurtosis quantifies the extent to which the tails of a distribution are heavier (high kurtosis) or lighter (low kurtosis) than those of a normal distribution. For instance, the presence of outliers (sample points that deviate from the 3σ region around the peak of the distribution) affect the tails of the distribution and therefore, since $(x-\mu)/\sigma$ is raised to the fourth power, will impact the value of Kurtosis.

As noted previously, moments allow us to characterize probability distributions. In general, when even moments exist and odd moments vanish, the distribution is symmetric with respect to the mean. However, cases can also emerge where the moments of the probability distribution do not converge.

2.1.2 Gaussian distribution

The Gaussian distribution is one of the most implemented distributions to describe nature. The reason lies in the *central limit theorem* which basically states that any distribution with finite variance will converge to a normal distribution as the sample size increases [91].
The univariate Gaussian distribution is

$$P(X|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-1}{2}\left(\frac{X-\mu}{\sigma}\right)^2\right]. \tag{2.8}$$

where μ is the mean (first moment) and σ^2 is the variance (second moment). Hence, the distribution is parameterized by the two parameters $\{\mu, \sigma^2\}$ only, with higher-order moments equal to 0. For this reason, it is common to denote a Gaussian distribution as $\mathcal{N}(\mu, \sigma^2)$.
For multivariate random variables, X turns into a n length vector denoted as $\mathbf{X}$, with the Gaussian distribution given by

$$P(\mathbf{X}) = \frac{1}{(2\pi)^{n/2}|\mathrm{det}C|^{1/2}} \exp\left[\frac{-1}{2}(\mathbf{X}-\boldsymbol{\mu})^T C^{-1}(\mathbf{X}-\boldsymbol{\mu})\right], \tag{2.9}$$

where

$$C = \langle(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T\rangle \tag{2.10}$$

is the covariance matrix and has shape $n \times n$.
In many situations, random variables follow non-Gaussian distributions (deviations from

Gaussianity can be easily detected by computing skewness and kurtosis). This situation can complicate probability distribution manipulations, such as computing marginals or implementing an accurate distribution of the data. For this reason, methods that "Gaussianize" the distribution have emerged [92, 93, 94]. Some examples are the method of data compression [95, 96], or the application of transformations to data that reshape the probability density nearly to a Gaussian shape [93].

2.1.3 χ^2- distribution

The Chi-squared or χ^2-distribution is a continuous probability distribution of common use in statistical inference and also in cosmology.
Consider n independent random variables z_i where $z_i \sim \mathcal{N}(0,1)$. The sum of their squares

$$x = \sum_{i=1}^{n} z_i^2, \tag{2.11}$$

follows the χ^2-distribution

$$\chi^2(x,k) = \frac{x^{\frac{k}{2}-1}\mathrm{e}^{-x/2}}{2^{k/2}\Gamma(\frac{k}{2})}, \tag{2.12}$$

where k is the number of degrees of freedom (number of random variables n).
In 1900, Karl Pearson proposed the χ^2-test [97] for hypothesis testing. It states that given a set of independent observations $\{X_i\}$ sampled from some distribution, then the χ^2-test

$$\chi^2 = \sum_i \left(\frac{X_i - \mu_i}{\sigma_i}\right)^2, \tag{2.13}$$

follows a Chi-squared distribution.

2.1.4 Poisson distribution

The Poisson distribution is the probability of a number of events x occur in a given interval of time and space, hence describing point-like events in nature. It has special interest for observational astronomy as it describes photon noise and contributes to the covariance of large-scale structure observables.
The distribution of a discrete random variable x is given by

$$P(x = n) = \frac{\lambda^n \mathrm{e}^{-\lambda}}{n!}, \tag{2.14}$$

where e is Euler's number, n the number of occurrences and λ a model parameter in the real domain which determines the shape of the distribution and represent the average rate [98].

2.2 Statistical inference

In order to provide a statistical description of the properties of the Universe, we need to introduce a physical model described in terms of a set of model parameters. With the

aim of determining the physical model, we first need to estimate the value of the model parameters. Of course, this need to be done based on the observations of nature, i.e. the observed distribution of galaxies in redshift surveys, the Hubble diagram of galaxies, or the temperature fluctuations of the Cosmic Microwave Background radiation. Therefore, we need to go beyond summarizing the observed data through descriptive statistics and the basics of random variables, and somehow connect the random data with the random model parameters. Statistical inference is the framework that accomplishes this, and in this section we present an overview of how it is implemented in cosmology. We aim to answer the following questions:

- How to estimate such model parameters?
- How well the model fits the observed data?
- What is the allowed range of the free parameters?
- How do we incorporate our prior beliefs on the free parameters in the inference algorithm?.

There are many ways to tackle the aforementioned questions. For convenience, let us first introduce the following notation:

- In this chapter, we denote by $\mathcal{M}(\boldsymbol{\theta})$ the parametric model described by the set of free parameters $\boldsymbol{\theta}$ (say, the model parameters of ΛCDM).
- $\mathbf{X} = \{X_1, X_2 \dots X_N\}$ denotes a realization of the (observed) data.

2.2.1 Likelihood

In cosmology, the data vector $\mathbf{X}$ can be a set of observables such as the Type-Ia supernovae distance modulus μ, the correlation functions of galaxy positions, etc, all derived from raw astronomical observations. An effective physical model $\mathcal{M}(\boldsymbol{\theta})$ provides theoretical predictions for such observables, once $\boldsymbol{\theta}$ are specified. To determine $\boldsymbol{\theta}$, we want to compare the observed data with the theoretical predictions and establish which set of free model parameters $\boldsymbol{\theta}$ minimize the discrepancy between the two. This analysis is done by invoking the likelihood function or forward model given by

$$\mathcal{L} = P(\mathbf{X}|\boldsymbol{\theta}, \mathcal{M}), \tag{2.15}$$

which is the probability of measuring the data given a set of parameters $\boldsymbol{\theta}$ assumed to be true. In other words, it describes how frequently nature (as described by the proposed physical model) generates data $\mathbf{X}$.
In most cases, the likelihood function is modeled as a multivariate Gaussian distribution,

$$\mathcal{L}(\mathbf{X}|\boldsymbol{\theta}, \mathcal{M}) = (2\pi)^{\frac{-N}{2}} |C|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{X} - \boldsymbol{\mu}(\boldsymbol{\theta}, \mathcal{M}))^t C^{-1} (\mathbf{X} - \boldsymbol{\mu}(\boldsymbol{\theta}, \mathcal{M}))\right], \tag{2.16}$$

where $\boldsymbol{\mu}(\boldsymbol{\theta}, \mathcal{M})$ is the (theoretical) model prediction and C^{-1} the data precision matrix. It is worth noting that the true likelihood function is unknown, but can be estimated analytically or by sampling the distribution using a set of simulations. We typically assume a Gaussian likelihood due to the central limit theorem, but in many situations this does not hold [99].

The parameter estimation is carried out by comparing data and predictions through the maximum likelihood estimation (MLE) method. MLE consists in choosing the value of parameters $\boldsymbol{\theta}$ that would have been most likely for generating the observed data $\mathbf{X}$. This means we aim to maximize $\mathcal{L}$ with respect to $\boldsymbol{\theta}$.

2.2.2 Bayes' Theorem

In the field of statistical inference, two primary approaches exist: frequentist statistics and Bayesian statistics. The key distinction between these approaches lies in how they interpret probabilities. In the frequentist interpretation, a probability is the frequency with which an event occurs in N number of independent trials, in the limit $N \Rightarrow \infty$. Particle physics is widely regarded as a frequentist science because its theoretical foundations rely on event rates, transition probabilities, branching fraction, etc. In brief, it is built based on the fact that experiments can be repeated a finite number of times. However, there are situations where repeating the experiment is unfeasible (for example, the Trinity test of the nuclear bomb, or any event that can only be experienced once), thus making the Bayesian approach a natural choice. As we have only one Universe to observe, cosmology inherently favors the Bayesian reasoning, where probability is taken as a *degree of belief*. It is important to note that both approaches agree on Bayes' theorem as a mathematical result, which stems from simple derivations based on probability axioms. However, for Bayesians, Bayes' theorem assumes the fundamental role in statistical inference. Under our current notation, Bayes' theorem reads as

$$P(\boldsymbol{\theta}|\mathbf{x}) = \frac{P(\mathbf{x}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\mathbf{x})}. \tag{2.17}$$

The probabilities presented in this equation correspond to

- $P(\mathbf{x}|\boldsymbol{\theta})$ expresses the probability of observing/measuring the data, given a set of parameters $\boldsymbol{\theta}$ which we assume to be true. This is the definition of the likelihood function denoted as $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta})$ in Section 2.2.1.
- $P(\boldsymbol{\theta})$ is the prior distribution of the parameters $\boldsymbol{\theta}$. This probability incorporates our knowledge about the model before the data have been gathered. It incorporates external information, such as knowledge from previous experiments or constraints imposed by the physical theory.
- P($\mathbf{x}$) is called the evidence. It describes the probability of measuring the data $\mathbf{x}$ given all possible outcomes of parameters $\boldsymbol{\theta}$.
- $P(\boldsymbol{\theta}|\mathbf{x})$ is the posterior probability that the model is true, given the data and the prior beliefs on $\boldsymbol{\theta}$.

The key aspect of Bayes' theorem as a inference framework lies in its ability to combine different data sets, such as incorporating the information obtained from one data set as prior information. In summary, it allows to update the prior knowledge in light of new data, making Bayesian thinking a natural choice for cosmology, a physical theory driven by the discoveries of newer experiments."

2.2.3 Prior

One of the reasons of debate between frequentist and Bayesians is the prior. The prior is a fundamental ingredient in Bayesian statistics, and it is not possible to get rid of it. However, specifying the prior belief as a probability distribution over $\boldsymbol{\theta}$ is a non-trivial task. Therefore, it is natural to ask how do you select the prior distribution. We hope for the data to be informative, which means the likelihood will dominate over the prior. However, this is not always the situation [100].
Among the classes of priors, it is customary to find *flat priors*, where the parameters $\boldsymbol{\theta}$ are bounded within specific lower and upper limits determined by the underlying theory. In such cases, the prior probability can be modelled as a uniform distribution, which assigns equal probability to all values within the defined range. For instance, a common example is assuming Ω_m to lie within the interval $0 < \Omega_m < 1$. These priors are sometimes referred as uninformative priors, however there is debate on whether the uniform probability distribution is in all cases uninformative [101]. It is also common to adopt Gaussian priors when the posterior distributions of θ obtained from previous experiments result in a Gaussian distribution. Another important distinction for priors is conjugate versus non-conjugate priors. If the posterior and prior probability distributions belong to the same family of distribution, the prior corresponds to a conjugate prior.
For instance, a Gaussian prior convolved with a Gaussian likelihood produces a Gaussian posterior, and thus the Gaussian prior is a conjugate prior [102].

2.2.4 Sampling

Cosmology aims to infer the model parameters $\boldsymbol{\theta}$ based on the available data, leading to the quest for obtaining the posterior distribution of these parameters. However, in numerous scenarios, the posterior distribution of the parameters is non-Gaussian, making analytical solutions infeasible. Consequently, resorting to computational methods becomes necessary to approximate the posterior distribution. A commonly employed approach for this purpose is Markov Chain Monte Carlo (MCMC) sampling methods.

2.2.5 Monte Carlo Markov Chain

As in a Monte Carlo integration, the MCMC algorithm samples the parameter space to approximate the posterior distribution. A Markov chain represents a series of random states $\boldsymbol{\Theta}$, where each state depends only on its immediately preceding state $\boldsymbol{\Theta} - 1$ (and not on earlier ones). Incorporating Markov chains into the approach renders it a powerful tool for efficiently sampling posterior distributions. The Metropolis-Hastings algorithm [103, 104] is the simplest efficient mechanism employed to transition from one state to another. It work as follows [31]

- Initialize the algorithm at a given point $\boldsymbol{\Phi_0}$.
- Define a proposal distribution $q(\boldsymbol{\Phi_0})$ (usually a multi-variate Gaussian).
- Iterate: from i=0 to i=N_{mcmc}
  - Propose a new state by drawing a random point $\boldsymbol{\Phi_0^*} \sim q(\boldsymbol{\Phi_0})$
  - Evaluate the ratio of the posterior at the proposed state and the current state $R = P(\boldsymbol{\Phi_0^*})/P(\boldsymbol{\Phi_0})$

- If $R > 1$, accept $\mathbf{\Phi_0}$ and store it.
- If $R < 1$, draw $\alpha \sim Uniform[0, 1]$. If $R > \alpha$, accept $\mathbf{\Phi_0}$ and store it. If If $R < \alpha$, reject proposed state.

Comparing R with a random point from a uniform distribution is the crucial factor for why the Metropolis-Hastings algorithm is successful. This particular step enables the walker to effectively explore the parameter space and prevents it from getting trapped in regions of local maximum: at these local maximum, no neighboring points have a lower probability, making it unfeasible to keep exploring the space of parameters. When comparing with a random point, there are chances to jump into another state, even if the walker is trapped in a local maximum. When the chain converges, optimal performance is achieved when the acceptance rate of the chain (acceptance-to-rejection fraction) is approximately 0.3. A high acceptance rate, around ~ 0.9, indicates that the sampler is taking steps that are too small, resulting in the walkers not efficiently exploring all regions in the parameter space. On the other hand, a low acceptance rate, approximately ~ 0.05, implies that steps are extremely large, causing the sampler to easily leave the global region of highest probability. The acceptance rate could be far from optimal typically in high-dimensional posterior distributions or when the distribution is multi-modal. In such cases, the sampler may not reach convergence in a reasonable running time, or the chain may continuously get stuck. Therefore, alternative algorithms are available, such as Hamiltonian Monte Carlo [105] or Nested sampling [106] which can outperform Metropolis-Hastings.

2.3 Machine Learning

In this thesis, our objective is not only to *infer* an underlying physical model from data but also to *learn* structures in the data. To achieve this, we delve into machine learning methods, which have emerged as crucial tools in astronomy in recent years. Significant strides have been achieved across various domains of data analysis in the field of astronomy, particularly in the realm of cosmology. The volumes of data obtained from ground-based telescopes and satellites have increased drastically. We will move from 1 TB[1] of data per night (as in DES [107]) to 15TB in Rubin-LSST. Therefore, the amount of data is no longer an important constraining factor; however, the statistical analysis of such data can become burdensome unless we come up with an efficient approach to process and summarize this wealth of information.

To maximize the scientific-return from large volumes of precise data, machine learning has emerged as a powerful tool in cosmology. Some examples of successful implementations of machine learning methods in cosmology includes photometric redshift estimation [108, 109], emulators of summary statistics [110, 111, 112, 18, 19, 113], deep learning for transient alerts [114, 115, 116] (with accurate detection of Type-Ia Supernovae), among others. In this thesis I focus on supervised methods, which are characterized by the inclusion of the desired solutions in the training data-set, commonly referred as labeled examples. Some supervised machine learning algorithm include *Support Vector Machines (SVMs), Decision Trees* and *Neural networks* [117]. Firstly, it is convenient to provide a summary of some key concepts in machine learning:

[1] Terabyte

- **Features**: Features (often called *attributes*) refer to a property of an observation. Let us consider the data-set arranged as a matrix $\mathcal{D}$ with dimensions $N \times M$. Each entry $\mathcal{D}_i$ with $i = 1 \dots N$ will be a M-length vector, where M stands for the number of features. If we aim to train a machine learning model that classifies observations into two categories, babies and adults, each entry will represent an observation (either of a baby or an adult), and the M features can be attributes such as age, height, etc. It is crucial that these features carry informative content about the category they represent.
- **Model parameters and hyper-parameters**: The former refers to trainable parameters of the model. These are constantly being updated by the algorithm in each epoch of training until the learning converges to the optimal values. On the other hand, hyperparameters are chosen by hand before the training process, and thus they are not learned during training. These parameters include the loss function (and some of its parameters), the number of training epochs, and other model-specific parameters. Hyperparameters may vary between different machine learning methods, and depending on the model's performance, the designer may consider tuning the hyperparameter values.
- **Training-Validation-Test sets**: In order to train a machine learning method, we first randomly split the data-set in three sets: training, validation and test. The training set typically covers 70-80% of the sample and is used to learn the model parameters that best map from input to output. Once the training has finished, we obtain our first approach of the machine learning model. The validation set is then utilized to evaluate the model derived from the training process. If the performance on the validation set is poor, the training process is repeated. We iterate between training and validation until we achieve the optimal model parameters. Finally, we perform the final model evaluation using the test set (also known as cross-validation). If there are two competing models, we utilize the test set to determine which one exhibits the best score. It is crucial that the three sets do not overlap and are representative of the overall distribution of features.
- **Loss function**: The performance of the machine learning model is evaluated using the loss function, which assesses the error between the prediction and the label.

A successful supervised learning model will be capable of learning the structure contained in the features of the labeled outputs and generalizing to unseen examples (for example, those present in the test set). However, these models can fail due to overfitting, which means that the model learns all details and fine structures contained in the training set (in the worst scenario, it fits the noise in the data). As a result, it lacks the ability to generalize to features that deviate from what it has learned [118]. For successful machine learning training, a good dataset should include enough samples relevant to the problem, exhibit representativity, and include informative features.
Also, it is important to distinguish between two common science cases: regression algorithms and classification. Neural networks are a class of ML method that can frequently outperform other ML algorithms in both cases, so we first introduce artificial neural networks.

2.3.1 Artificial neural network

A neural network (NN) corresponds to a non-linear model that maps input data $\mathbf{x}$, containing M features, to the output $\mathbf{y}$, which approximates the prediction of the target's true value τ

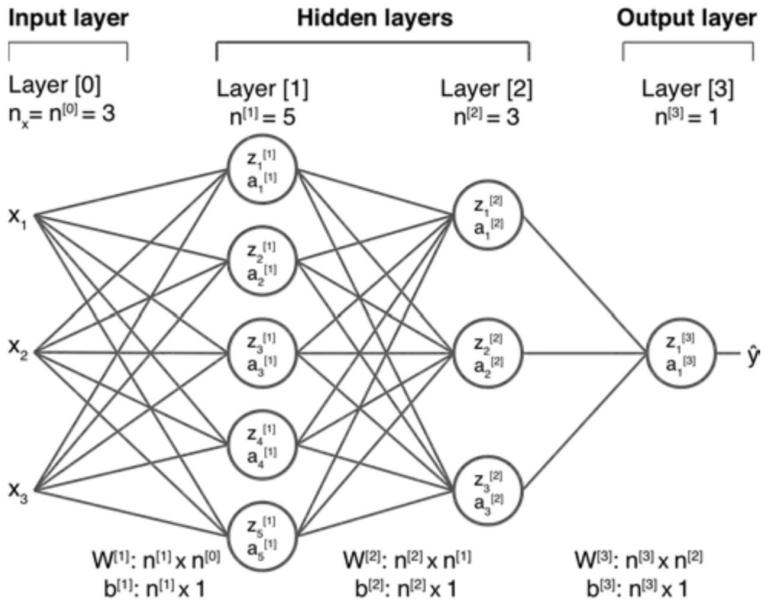


Figure 2.2: Multilayer perceptron with 1 input layer, 2 hidden layers and 1 output layer. Input layer consists in 3 neurons, first hidden layer contains 5 neurons and second hidden layer has 3 neurons. Each neuron contains the original input $z_i^{N_l}$ which results from the weighted sum. Then, the activation function is applied to $z_i^{N_l}$ and outputs the result $a_i^{N_l}$ which is fed to the next layer. Figure taken from [6].

[119]. The specific configuration of a neural network is called its *architecture*.
In a nutshell, a neural network operates as follows: The architecture comprises an input layer, which takes in the data $\mathbf{x}$ into the system. A layer is composed of a finite number of neurons, with which the algorithm performs a weighted sum of the inputs. An activation function is applied to the weighted sum, and this non-linear result is then fed into the next layer. Intermediate layers, known as *hidden layers*, repeat the weighted sum until reaching the last layer, which outputs the desired outcome.
The simplest neural network is called a *perceptron* [120] and consists of one layer with a single neuron. It is suitable for simple linear binary classification problems but lacks the flexibility required for more complex problems. Nowadays, we rely on multilayer perceptrons (a set of interconnected perceptrons) shown in Fig. 2.2, which incorporate hidden layers and employ differentiable activation functions that provide the non-linearity feature to neural network models.
The non-linear model is parameterized by the set of trainable parameters called *weights* w, with one weight associated to one neuron. The weighted sum reads as

$$a_i = \varphi\left(\sum_i^L w_{ki} x_k\right). \tag{2.18}$$

It is worth mentioning that the activation function is chosen before the training process. Therefore, the number of layers, number of neurons and activation functions are *hyperparameters*.
To determine the weights, we utilize the loss function $\mathcal{L}$. This function is computed at the end of the feed-forward process (i.e., when all the weighted sums are performed from the input layer until the output layer), and it measures the discrepancy between the predicted outputs

τ and the targets $\mathbf{t}$. Then, the neural network propagates this error through a backward pass, from the output layer to the input layer. This process, known as backpropagation, aims to find the weights that minimize the loss function, utilizing an optimization method called gradient descent [119]. In this algorithm, the weights are iteratively updated by following

$$w^{new} = w^{old} - \lambda \frac{\mathrm{d}\mathcal{L}}{\mathrm{d}w}, \tag{2.19}$$

where the derivative illustrates how the algorithm follows the gradient of the loss function in order to find the global minimum. The hyperparameter λ, known as the learning rate, determines the extent to which we update the weights to explore the loss potential and discover the global minimum. If λ is too small, the algorithm will explore the loss function through small steps, leading to a longer time to find the minimum. Conversely, if λ is too large, the algorithm may take excessively large steps, potentially causing it to miss the optimal weights during iterative updates. For a more detailed overview of artificial neural networks, see [121].

2.3.2 Gaussian Process

Gaussian Process is a supervised machine learning method to reconstruct an unknown function (without the assumption of a parametric model) of a stochastic process defined as $y = f(t) + \epsilon$. It can work as an interpolation method or regression model given data (y_i, t_i) with $i = 1...N$, and provide predictions on new points (y_i^*, t_i^*) with the corresponding uncertainties. The errors on the observations are represented by ϵ which follows a Gaussian distribution $\mathcal{N}(0, \sigma_y^2)$.
The Gaussian process method consist in a correlated set of random variables sampled from a joint Gaussian distribution $\mathcal{N}(\mu, K)$:

$$\begin{bmatrix} f(t_1) \\ f(t_2) \\ \vdots \\ f(t_N) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu(t_1) \\ \mu(t_2) \\ \vdots \\ \mu(t_N) \end{bmatrix}, \begin{bmatrix} k(t_1,t_1) & k(t_1,t_2) & \dots \\ k(t_2,t_1) & k(t_2,t_2) & \\ \vdots & \vdots & \dots \\ k(t_N,t_1) & \dots & k(t_N,t_N) \end{bmatrix}, \right) \tag{2.20}$$

where $k(t_i, t_j)$ is the kernel function and defines the covariance at two locations t_i and t_j; and $\mu(t_i)$ is the expected value at the location t_i. It is usually assumed that the kernel corresponds to

$$k(t_i, t_j) = \sigma_f^2 \exp\left[\frac{-(t_i - t_j)^2}{2l^2}\right], \tag{2.21}$$

where σ_f and l are hyperparameters. This equation determines similarity between two observations and tells us that two random variables that are close-by in location $t_i \approx t_j$, will present high covariance. If we sample $f(t_i)$ and it results in a large value, then we expect that a value $f(t_j)$ close to it, will be a large value as well.This property is important as data will provide guidance of the most likely $f(t_*)$ for a new point t_*. The parameters σ_f and l

are learned. We introduce the Gaussian noise of our observations into the modelling, as we aim to predict y_* instead of $f(t_*)$. The joint distribution including a new observation is

$$\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu \\ \mu_* \end{bmatrix}, \begin{bmatrix} K + \sigma^2 I_n & K_*^T \\ K_* & K_{**} \end{bmatrix} \right). \tag{2.22}$$

As we are working in the framework of machine learning, what we actually want is the conditional probability of a new observation t^* (as present in the test-set) given training data y, t. We can derive this expression in a Bayesian framework

$$p(y_*|y) \sim \mathcal{N}(\bar{y^*}, cov(y^*)) \tag{2.23}$$

where the mean of the posterior distribution and the covariance are

$$\bar{y^*} = \mu(t_*) + K(t^*, t)[K(t^*, t) + C]^{-1}(\mathbf{y} - \mu(t)) \tag{2.24}$$

$$cov(y^*) = K(t^*, t^*) - K(t^*, t)[K(t^*, t) + C]^{-1}K(t, t^*) \tag{2.25}$$

with $C_{n\times n} = \sigma_y^2 I_{n\times n}$.
It is important to train the GP model using different kernels. Among the possibilities, the more implemented kernels in astronomy are:

- Kernel: Radial Basis Function (RBF)

$$k(t_i, t_j) = \sigma_f \exp\left(\frac{-(t_i - t_j)^2}{2l^2} \right) \tag{2.26}$$

- Kernel: Matérn $\nu = 5/2$

$$k(t_i, t_j) = \sigma_f \exp\left(\frac{-\sqrt{5}|t_i - t_j|}{l} \right) \left(1 + \frac{-\sqrt{5}|t_i - t_j|}{l} + \frac{5(t_i - t_j)^2}{3l^2} \right) \tag{2.27}$$

More details on Gaussian processes and some derivations of the quantities described above can be found at [122, 123].

Chapter 3

Cosmic shear

The perturbed Universe consists of gravitational potentials, seeded by tiny density fluctuations in the early Universe, whose dominant component is dark matter. The pull of gravity in these dark matter overdensities attracts baryonic matter to clump. Baryonic gas inside the potential wells cools down, leading to galaxy formation, and clustering of galaxies at large scales. As dark matter can not be directly observed, our only chance to trace the underlying dark matter distribution and study the properties of the large-scale structure is through the observation of galaxies at different wavelengths [124]. However, this is not an easy task as galaxy formation and evolution occurs at large time-scales, it depends on interactions with the matter distribution and the underlying physical processes are still not entirely characterized on a range of scales. Therefore, the clustering of luminous matter can trace the dark matter web in a way that is not completely understood, so we say it is a biased tracer. It is desirable to have a tracer in place, which is sensitive to the total matter density (instead of luminous matter only) and thus trace the cosmic web on large scales in an unbiased way. This tracer corresponds to the gravitational lensing of galaxies. Gravitational lensing affects the apparent shapes and position of objects in the Universe, due to the light ray being bent around the curvature induced by the intervening matter in the path of light. This effect is well described in General Relativity and it was actually predicted by several physicists in the classical Newtonian limit years before GR [125].

Depending on the mass of the intervening structures, and the relative distances between the background objects, the deflecting mass and the observer, gravitational lensing can be classified in three types of lensing: strong lensing, microlensing and weak lensing. In this thesis, we focus on the weak gravitational lensing of galaxies by the large-scale structure in the Universe, also called *cosmic shear* [126].

Cosmic shear corresponds to slight distortions and magnifications of galaxy shapes. It is a weak signal and, as such, it can not be detected by observing one single background galaxy. Instead, statistical correlations among thousands of galaxy shapes are needed in order to measure the cosmic shear signal. Since its discovery in 2000, cosmic shear has led to stringent cosmological constraints, and was considered in the Dark Energy task force [127] as one of the key cosmological probes to study the properties of dark energy.

This chapter presents an overview of the gravitational lensing formalism, its statistical description and the systematics that can affect the lensing signal in the weak lensing regime.

Many notes and definitions are based on the review articles [126, 128] and textbooks [30, 31].

3.1 Basics of weak gravitational lensing

Light from distant objects propagates along null geodesics from its origin to our telecopes, being continuously deflected by inhomogeneities in space-time. Gravitational lensing is the distortion and magnification observed in the images of objects as a result of the deflection of light. The massive object causing the perturbation is called *lens*, whilst the source emitting light is called the *source*. As gravitational lensing is an effect due to the lens gravitational potentials, let us start by considering the weakly perturbed metric under a flat FLRW background

$$\mathrm{d}s^2 = -\left(1+\frac{2\Psi}{c^2}\right)c^2\mathrm{d}t^2 + a^2(t)\left(1-\frac{2\Phi}{c^2}\right)\mathrm{d}r^2, \tag{3.1}$$

where, as in Chapter 1, $a(t)$ is the scale factor and t is cosmic time. The perturbed metric incorporates the Bardeen potentials $\{\Theta,\Phi\}$ that describe perturbations in the weak field limit $\Psi, \Phi \ll c^2$. In this particular case, we explicitly write the speed of light c for convenience.

As light propagates in null geodesics $\mathrm{d}s = 0$, the effective speed of photons in the presence of Φ is $c' = c\left(1+\frac{2\Phi}{c^2}\right)$ at first order, which implies that the potentials Ψ and Φ act as a medium with effective refraction index $n = 1-\frac{2\Phi}{c^2}$, altering the trajectory of light. Hence, by applying Fermat's principle which states that light rays will follow the path that extremizes $\delta t = 0$, we can find the deflection angle [129]

$$\tilde{\boldsymbol{\alpha}} = \frac{-2}{c^2}\int \boldsymbol{\nabla}_\perp \Phi \mathrm{d}r, \tag{3.2}$$

where the integration variable $\mathrm{d}r$ is the radial path along the light ray trajectory and the 2D gradient operator $\boldsymbol{\nabla}_\perp$. Since the typical deflection angles are small, it is sufficient to work in the *Born approximation*, i.e. the perturbed light path can be approximated by the unperturbed one, and thus we can integrate over the unperturbed trajectory. In what follows, we first introduce the lensing formalism in the *thin-lens approximation*, which states that the lens mass distribution is thin compared to the distance between the observer, lens and source. This means that we can approximate the lens distribution to a planar distribution of matter, often referred to as the *lens plane*. It is also common to define the source to lie in the *source plane*. In practice, the deflection angle $\tilde{\boldsymbol{\alpha}}$ induces a shift in the apparent angular position $\boldsymbol{\theta}$ of the observed source due to the deflecting mass. The original angular position, denoted by $\boldsymbol{\beta}$, can be recovered by subtracting the amount of lens deflection $\boldsymbol{\alpha}$ to the observed position. The latter is the definition of the *lens equation* given by

$$D_s\boldsymbol{\beta} = D_s\boldsymbol{\theta} - D_{ls}\boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{3.3}$$

where all angular positions are 2D vectors with respect to the optical axis (the line of sight). The distances are the angular diameter distances to the source D_s and between the lens and the source D_{ls}. On the other hand, $\tilde{\boldsymbol{\alpha}}$ and $\boldsymbol{\alpha}$ only differ in the location from where the

deflection is being measured. The former is the amount of shift as seen from the lens position, and the latter is the observed shift from Earth. Both are related through the distance ratio

$$\boldsymbol{\alpha} = \frac{D_{ls}}{D_s}\tilde{\boldsymbol{\alpha}}. \tag{3.4}$$

We are interested in the reduced deflection angle which can be written in terms of the *lensing potential*

$$\phi(\theta) = \frac{2}{c^2}\frac{D_{ls}}{D_{\mathrm{d}}D_s}\int_0^{\chi} \mathrm{d}z\Phi(D_{\mathrm{d}}\boldsymbol{\theta}, z), \tag{3.5}$$

which traces the properties of the gravitational lens, and can be thought of as a scaled projection of the three-dimensional potential Φ. Hence, the reduced deflection angle Eq. (3.4) can be denoted by $\boldsymbol{\alpha} = \boldsymbol{\nabla}_{\boldsymbol{\theta}}\phi$ and thus the following relations can be obtained

$$\boldsymbol{\nabla}\cdot\boldsymbol{\alpha} = \boldsymbol{\nabla}^2\phi = 2\kappa(\boldsymbol{\theta}), \tag{3.6}$$

where κ, called *convergence*, defines the projected surface mass density

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(D_l\boldsymbol{\theta})}{\Sigma_{crit}}, \tag{3.7}$$

scaled by the geometrical factor that characterizes the lens system

$$\Sigma_{crit} = \frac{c^2}{4\pi G}\frac{D_s}{D_l D_{ls}}, \quad \Sigma(D_l\boldsymbol{\theta}) = \int \rho(D_{\mathrm{d}}\boldsymbol{\theta}, z)\mathrm{d}z. \tag{3.8}$$

As the lens mass distribution can be approximated by the lens plane, the surface density takes the simple form of Eq. (3.8), which integrates the three-dimensional density profile ρ along the radial coordinate z. Hence, the projected surface mass density κ sources the inhomogeneites encoded in the potential Φ that bends the emitted light from a distant source.
In order to understand the distortions due to lensing, we can build a local mapping between the lensed coordinates and unlensed coordinates by means of the Jacobian $\mathcal{A}_{ij} = \partial\beta_i/\partial\theta_j$. The elements of the Jacobian define the distortion matrix in terms of the lensing potential as

$$\mathcal{A}_{ij} = \frac{\partial\beta_i}{\partial\theta_j} = \delta_{ij} - \frac{\partial^2\phi}{\partial\theta_i\partial\theta_j}. \tag{3.9}$$

This matrix accounts for two situations: given a lens mass distribution, the Jacobian describes the curvature effect induced by the lens through second derivatives of ϕ, and in the absence of lensing, the matrix simply reduces to the delta Kroenecker δ_{ij}, i.e. the identity matrix.
We can manipulate the distortion matrix and write the derivates of ϕ in terms of convergence κ and the shear γ

$$\mathcal{A}_{ij} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}, \tag{3.10}$$

$$\kappa = \frac{1}{2}\left(\partial_1\partial_1 + \partial_2\partial_2\right)\phi = \frac{1}{2}\nabla^2\phi, \quad \gamma_1 = \frac{1}{2}(\partial_1\partial_1 - \partial_2\partial_2)\phi, \quad \gamma_2 = \partial_1\partial_2\phi, \tag{3.11}$$

where $\{\gamma_1,\ \gamma_2\}$ are the shear components in $\gamma = \gamma_1 + i\gamma_2 = |\gamma|\exp(2i\varphi)$ describing the distortion in the shape of the source. In the following, results holds in the weak lensing regime, $\kappa, \gamma \ll 1$. In Fourier space, the relations defined above are

$$\tilde{\kappa}(\boldsymbol{\ell}) = \left(\frac{\ell_1^2 - \ell_2^2}{\ell_1^2 + \ell_2^2}\right)\tilde{\gamma}_1(\boldsymbol{\ell}) + 2\left(\frac{\ell_1\ell_2}{\ell_1^2 + \ell_2^2}\right)\tilde{\gamma}_2(\boldsymbol{\ell}), \tag{3.12}$$

where $\boldsymbol{\ell} = (\ell_1, \ell_2)$ is the Fourier counterpart of the angular positions. With this expression, we can convert the shear field into a convergence field.
To gain physical insights of convergence and shear, we can split Eq. (3.10) into an isotropic and anisotropic trace-free part as

$$\mathcal{A}_{ij} = (1-\kappa)\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} - \gamma\begin{pmatrix}\cos(2\phi) & \sin(2\phi)\\ \sin(2\phi) & -\cos(2\phi)\end{pmatrix}. \tag{3.13}$$

Equation 3.13 explains the effect of lensing into more detail. The shear is an anisotropic streching of the shape, turning a circular shape into an elliptical form. The convergence in Eq. (3.13) is an isotropic contribution that only alters the size of the image. This effect is called *magnificaton* and is quantified by the inverse of the determinant of $\mathcal{A}$

$$\mu \equiv \left|\frac{\partial\theta}{\partial\beta}\right| = \frac{1}{\det\mathcal{A}} = \frac{1}{(1-\kappa)^2 - |\gamma|} \approx 1 + 2\kappa, \tag{3.14}$$

at first order. This expression tells us that magnification is determined by the convergence with no contributions from the shear.
Let us consider a circular source under the lens effect. Its shape describes an ellipse due to lensing, with semi-major axis a and semi-minor axes b. These values are proportional to the eigenvalues of the inverse $\mathcal{A}^{-1}$

$$\lambda_+ = \frac{1-\kappa+|\gamma|}{\det\mathcal{A}} = \frac{1}{1-\kappa-|\gamma|}, \quad \lambda_- = \frac{1-\kappa-|\gamma|}{\det\mathcal{A}} = \frac{1}{1-\kappa+|\gamma|}. \tag{3.15}$$

Therefore, the ellipticity of that source is

$$\epsilon \equiv \frac{a-b}{a+b} = \frac{\lambda_+ - \lambda_-}{\lambda_+ + \lambda_-} = \frac{\gamma}{1-\kappa}, \tag{3.16}$$

where the last equality defines the reduced shear

$$g = \frac{\gamma}{1-\kappa}. \tag{3.17}$$

As the factor $(1-\kappa)$ does not contribute to shape distortion, what we can actually measure is the reduced shear. In cosmology, we are interested in the lensing of galaxy shapes. As galaxies are not intrinsically circular objects, the observed ellipticity is a contribution from the intrinsic ellipticity and the reduced shear

$$\epsilon \approx \epsilon_s + g. \tag{3.18}$$

In practice, the typical value of shear for lensing induced by the large scale structure in the Universe is approximately 100 times smaller than the intrinsic ellipticity. Therefore, in order to detect the weak lensing signal g, an average over the ellipticities ϵ of hundreds of galaxies is needed (g is a small quantity, same as κ and γ). We can estimate g with the central assumption that the galaxy shapes are randomly oriented and hence $\langle \epsilon_s \rangle = 0$,

$$\langle \epsilon \rangle \approx g. \tag{3.19}$$

In reality, the intrinsic ellipticities follow a normal distribution with $\sigma_\epsilon \approx 0.2$. This contribution is called shape noise. The signal-to-noise ratio for a lens system in the weak lensing regime is defined in terms of the shape noise and the number of galaxies N

$$\frac{S}{N} = \frac{\gamma\sqrt{N}}{\sigma_\epsilon}, \tag{3.20}$$

which automatically explains why galaxy surveys aim to observe a large number of galaxies in order to reduce the noise.

3.1.1 Generalization to lensing by the large-scale structure

In what follows, instead of considering the thin lens system, we study the weak gravitational lensing due to the large scale structure. In this context, the lens is all the intervening matter along the line of sight, and hence the distance to the lens is a continuous quantity to integrate over. Therefore, Eq. (3.5) generalises to comoving distances inside the integral instead of a pre-factor of fixed distances to the lens-source system. The lens potential generalizes to

$$\phi(\theta) = \frac{2}{c^2}\int_0^\chi \mathrm{d}\chi' \frac{\chi - \chi'}{\chi}\Phi(\chi'\boldsymbol{\theta}, \chi'), \tag{3.21}$$

where χ' is the comoving distance to the lens. As one of the central quantities in lensing is the convergence κ, it is possible to replace the Poisson equation and the mean matter density,

$$\nabla^2\Phi = \frac{4\pi G}{c^2}\bar{\rho}_m a^2 \delta, \quad \bar{\rho}_m = \frac{3H_0^2}{8\pi G}\Omega_m a^{-3}, \tag{3.22}$$

into $\nabla^2\phi/2$, to derive the convergence for an extended lens or multiple sources. The convergence results in

$$\kappa(\boldsymbol{\theta}, \chi) = \frac{3H_0^2\Omega_m}{2c^2}\int_0^\chi \mathrm{d}\chi' \, \frac{\chi'(\chi - \chi')}{a(\chi')\chi}\delta(\chi'\boldsymbol{\theta}, \chi'). \tag{3.23}$$

Again, from Eq. (3.23) it is clear that convergence describes the projected density along the line of sight, weighted by the geometrical factor $\chi'(\chi - \chi')/(a(\chi')\chi)$ involving the distances to the elements of the system. Also, it is seen that the convergence directly depends on the inhomogeneities in the Universe, and therefore is a powerful tool to study the perturbed Universe and provide a statistical description of the clustering of structures. For a sample of source galaxies, the mean convergence can be obtained as

$$\kappa(\boldsymbol{\theta}) = \int_0^{\chi_{lim}} \mathrm{d}\chi \, n_\chi(\chi) \, \kappa(\boldsymbol{\theta}, \chi), \tag{3.24}$$

where $n_\chi(\chi)$ denotes the distribution of galaxies as a function of the comoving distance, and χ_{lim} is the comoving distance to the farthest galaxy. The probability distribution of $n_\chi(\chi)$ can be computed in terms of redshift through the equality $n(\chi)\mathrm{d}\chi = n(z)\mathrm{d}z$, with $\int n(\chi)\mathrm{d}\chi = 1$. Inserting Eq. (3.24) into (3.23), the convergence recast into

$$\kappa(\boldsymbol{\theta}) = \frac{3H_0^2\Omega_m}{2c^2}\int_0^{\chi_{lim}} \mathrm{d}\chi g(\chi)\chi\frac{\delta(\chi\boldsymbol{\theta},\chi)}{a(\chi)}, \tag{3.25}$$

where

$$g(\chi) = \int_\chi^{\chi_s} \mathrm{d}\chi' n(\chi')\frac{\chi'-\chi}{\chi}, \tag{3.26}$$

is the lensing efficiency, which characterizes the lensing strength at a given comoving distance. From Eq. (3.25), it becomes clearer that weak lensing measurements are sensitive to the cosmological parameters defined in Chapter 1. This is not only due to the direct dependence on $H_0^2\Omega_m$, but also through the comoving distances and the density contrast.
In order to make use of the lensing convergence to extract the cosmological information encoded in weak lensing data, a statistical description of κ is needed. As explained in Chapter 1, we cannot access the exact position of the density contrast in the Universe. In the same way, the truth underlying lensing effects induced by the density field in a given direction are also unknown. What we know in practice, is that the measured image distortion of two nearby galaxies will be similar, as the photons emitted by both galaxies go through nearby structures in their journey towards us. On the contrary, the photons from two galaxies far away of each other will go through different structures, and hence the observed shape distortions may be very different. This indicates that the galaxy shapes are correlated. As the mean density contrast is equal to 0, the mean convergence also does vanish $\langle\delta\rangle = \langle\kappa\rangle = 0$. Therefore, the lensing effect is encoded in the two point correlation function

$$\xi(|\varphi|) = \langle\kappa(\boldsymbol{\theta})\kappa(\boldsymbol{\theta}+\boldsymbol{\varphi})\rangle, \tag{3.27}$$

where $\boldsymbol{\varphi}$ is the angular separation vector between two objects and the two-point function only depends on the modulus $|\varphi|$ due to statistical isotropy. In some occasions, it is useful to work with the *angular power spectrum*

$$C_\kappa(\ell) = \int \mathrm{d}^2\varphi\xi(\varphi)\mathrm{e}^{-i\vec{\ell}'\vec{\varphi}}, \tag{3.28}$$

which is the Fourier-transform of Eq. (3.27). In this expression, $\vec{\ell}'$ is the 2D wave vector, the Fourier conjugate of the separation angle $\vec{\varphi}$.
We can recast the angular power spectrum of convergence (3.28) into an integral of z or χ. First, it is important to note that the lensing effect is an integral over a 2D projection of the 3D mass fluctuations along the line of sight. It is possible define a 2D quantity $Q(\theta)$ as a projection of a 3D function by means of a weight function W. This is known as the *Limber approximation*

$$Q(\theta) = \int_0^{\chi_s} W(\chi)y(\chi\boldsymbol{\theta},\chi), \tag{3.29}$$

where $Q(\theta)$ is the weight function along the radial direction, and $y(\chi\boldsymbol{\theta},\chi)$ is defined in three-dimensions. The Limber's approximation gives an expression for the convergence angular power spectrum in terms of the 3D matter power spectrum through the weight function

$$W_\kappa(\chi) = \frac{3H_0^2\Omega_m}{2c^2}\frac{g(\chi)\chi}{a(\chi)}, \tag{3.30}$$

with the convergence power spectrum of cosmic shear given by

$$C_\kappa(\ell) = \frac{9H_0^4\Omega_m}{4c^4}\int_0^{\chi_s} \mathrm{d}\chi \frac{g^2(\chi)\chi}{a^2(\chi)} P_\delta\left(k = \frac{\ell}{\chi}, \chi\right). \tag{3.31}$$

Cosmic shear traces the matter distribution and it directly depends on cosmology through the prefactor of Eq. (3.31), distances and the matter power spectrum. The latter can be obtained from theory given a set of cosmological parameters, by means of a Boltzmann solver or through analytical expressions of the transfer function. The convergence power spectrum depends on the shape of the matter power spectrum, and hence it can provide tight constraints on the combination of parameters $\Omega_m^2\sigma_8^2$, as demonstrated in [130]. Therefore, equation (3.31) illustrates why cosmic shear is a powerful tool to study the growth of cosmic structures and reveal information about the properties of dark energy.

3.2 Systematics

Cosmic shear distorts the galaxies shapes very weakly, and therefore many processes that degrade the quality of the images can distort the galaxy shapes and dilute the lensing signal. In these situations, the estimation of shear from galaxy images may include other effects apart from lensing, which in turn may lead to biased results. The large-scale structure analysis incorporates many methods to account for these various processes, also called *systematic errors*, and hence survey strategies are designed to mitigate these effects [124].

One observational source of systematic error corresponds to the response of the telescope optics and the atmospheric effects, which are modelled by the Point Spread Function (PSF) [124, 131]. The image is always a convolution of the real galaxy image and the PSF. Hence, the modelling of the PSF properties is important in order to assess reliable measurements of galaxy shapes. Moreover, if the images exhibit an overlap of multiple galaxies, the galaxy shape measurement could also be degraded. This effect is called *blending* and can be a major issue in the upcoming galaxy surveys. Based on the Hyper Suprime-Cam wide survey realistic simulations, the authors [132] detected 58% objects in some form of blending configuration, thus a strategy to mitigate this effect is an imperative for galaxy surveys.

Another central aspect of weak lensing analysis is the estimation of redshifts. As reviewed in this chapter, the cosmic shear analysis implies integrals along the line of sight that depends on the distribution of the redshift of source galaxies $n(z)$. Accurate cosmological constraints depend on accurate estimations of $n(z)$, mainly obtained from photometric redshifts derived from flux measurements in photometric bands. Poor blending mitigation strategies can also lead to wrong photometric redshift estimates [133] and degrade cosmological constraints.

Many other systematic errors take place, and Stage-III weak lensing surveys have provided a deeper comprehension of the systematic effects in weak lensing measurements and analysis. Two important aspects that will highly impact the upcoming galaxy surveys are astrophysical processes that can contaminate the cosmic shear signal, and the modelling of the non-linear evolution of the matter density field. In the past years, it has been shown that astrophysical processes associated with baryons, such as star formation, magnetic fields, and the impact of

jets and winds that emanate from Active Galactic Nucle (AGN) impact the agglomeration of material on small physical scales [7]. Therefore, the matter power spectrum is re-shaped by such effects, especially at non-linear scales. An accurate theoretical modelling of the matter power spectrum involves thus a detailed study of the impact of these astrophysical phenomena. These astrophysical processes are often referred to as *baryonic feedback* or *baryonic processes*, and are the main topic of the following section.

3.2.1 Baryonic feedback

Baryonic processes such as radiative cooling, AGN feedback, magnetic fields, star formation and radiative cooling modify the distribution of matter inside haloes, reshaping the gravitational potentials where cosmic shear signal takes place [7, 134, 135]. Consequently, baryons suppress the clustering of matter from intermediate to small scales, although the amplitude of suppression is not precisely determined and scatters around a wide range of values [7]. In Fig. 3.1, each solid coloured line shows the impact of baryons on the matter power spectrum at $z = 0$, with the levels of suppression oscillation betwen 10%-30%. Unfortunately, these feedback mechanisms encompass an extensive list of effects which lack a completely theoretical modelling, and thus they are hard to address analytically. As it is expected that these effects can propagate into the cosmological parameter inference and may bias the parameter constraints, different approaches have been adopted by the community in order to mitigate these biases. Stage-III surveys have applied severe small-scale cuts to the two-point statistics in real and harmonic space to mitigate potential biases due to the unmodelled baryonic feedback [136, 137, 22], but such choice remove cosmological information, degrades the cosmological constraints and hence will prevent us of unleashing the statistical power of Stage-IV surveys. Many other strategies adopted to date rely on hydrodynamical simulations. These simulations include gravity (as in N-body simulations) but also incorporate fluid dynamics that account for the baryonic physics that happens in the Universe. By providing gravity-only simulation runs as counterparts with the same cosmology and initial conditions, it is possible to quantify the impact of baryons on the clustering of structures. A fast and accurate approach is the *baryonification* modelling [138, 139]. It consists in displacing particle outputs of N-body simulations that can mimic the effect of baryons inside haloes, producing modified density profiles of haloes. It has been shown that the method reproduces the power spectrum based on hydrodynamical simulations at 2% accuracy, but also is in good agreement with observations [140]. Moreover, the authors found that the baryonic strength is mainly determined by the maximum radius of gas ejection and the slope of the gas profile. They also show that baryons produce a 10-20% suppression of the weak-lensing power spectrum at scales scales $\ell > 500$, however, this method has not been sufficiently accurate to predict other statistics, such as the peak counts [141]. The authors in [142] present the halofit model, a matter power spectrum fitting function based on the halo model that provides predictions for the non-linear scales. The calibration is based on N-body simulations at various cosmological parameters, however, in [143] it was found that this model predicts $\sim 5\%$ less suppression at $k \sim 1h$ Mpc^{-1} compared to recent high-resolution N-body simulation. Therefore, an optimized approach was proposed in [144], with the inclusion of variations of the dark energy equation of state. Regarding the effect of baryons, they found that baryon cooling and the effect of massive neutrinos are important at small scales, with the baryon cooling enhancing the clustering at $k \sim 10h$ Mpc^{-1}. More recently, in [145] the authors provide an analytic fitting formula for the matter power spectrum that describes the effect of baryonic feedback

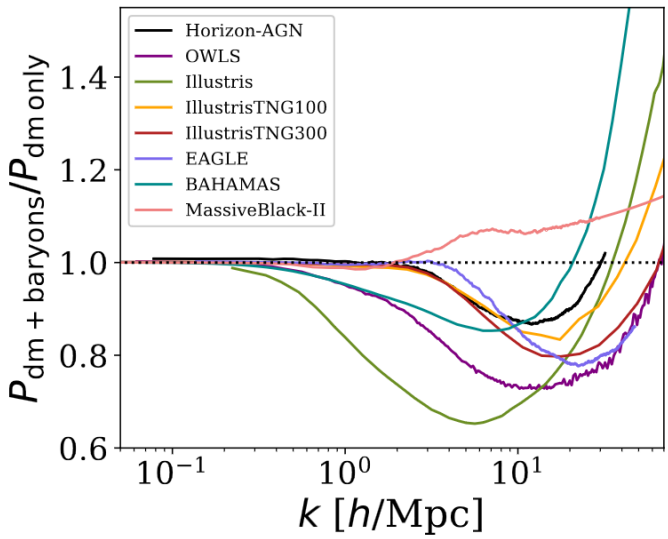


Figure 3.1: Ratio of the matter power spectrum at $z = 0$ as measured in hydrodynamical simulations, compared to the gravity-only runs. Each coloured solid line represent one different simulation. Simulations predict different amplitudes and scale-dependent impact of baryons. Original plot from [7].

in the range $0 < z < 1.5$ and scales $k < 100h/\text{Mpc}$, with baryons being modelled according to the Over-Whelmingly Large (OWL) hydrodynamical simulation.
By analyzing at which scales these baryonic feedback mechanisms can impact the weak lensing statistics, we can determine to what extent the effects of baryonic feedback can contribute to the understanding of the inferred low S_8, as explained in Chapter 1. Also, including baryons in the modelling helps to decide appropriate scale cuts for the tomographic redshift bins in future galaxy survey analysis.

3.3 Non-Gaussian statistics

The two-point statistics in real and harmonic space provide a full characterisation of a Gaussian random field. Moreover, the weak lensing surveys have placed tight cosmological constraints based on the two-point statistics: the angular correlation function or angular power spectrum of galaxy shapes. However, the non-linear evolution of structures in late times introduces non-Gaussianities in the matter density field. In that sense, the traditional two-point statistics cannot completely capture all the information encoded in the density field traced by weak lensing data. This acquires special care in light of the upcoming weak lensing surveys such as LSST and Euclid, as these surveys will deliver precise measurements of the small scales. In order to fully exploit the weak lensing data in all scales, we need statistical tools beyond the second-order statistical measurements. Various groups have introduced the *non-Gaussian statistics*, a set of statistical tools very sensitive to the non-Gaussian information in the cosmic density field that have shown great statistical power in constraining cosmological parameters.
Non-Gaussian statistics [146] include the bispectrum [147], the peak counts (or simply peaks) [148, 149, 21], minimum counts (or *minima*)[21], the probability distribution function (PDF)

[150, 151, 152], Minkowski functionals [153], among others. All of them have shown their potential to capture information beyond the one encoded in the angular power spectrum alone. Furthermore, non-Gaussian statistics are sensitive to systematics effects in different ways, therefore, a detailed analysis of its sensitivity to these systematic effects could guide decisions for constructing future pipelines in weak lensing.

3.3.1 Bispectrum

An alternative to the power spectrum is the bispectrum. As the power spectrum contains the information of the two-point correlation function in Fourier space, the bispectrum corresponds to the Fourier transform of the three-point correlation function [154]. It is lowest order statistic beyond the power spectrum that is sensitive to the non-Gaussian information of random fields, and hence its signal vanishes for Gaussian random fields.
The bispectrum measures three-point correlations of the convergence defined on in triangle configurations in Fourier space [155]. As a complete bispectrum calculation is computationally prohibitive, recent works have employed the binned bispectrum estimator [147], which corresponds to the sum over products of a set of filtered lensing maps. Since the first measurements, there has been significant work to understand the bispectrum constraning power and how different systematics can impact the bispectrum measurements. In [156], they found that the bispectrum and power spectrum contains similar information on parameters $\{\Omega_m, \sigma_8, w_0, w_a\}$ for lensing tomography analysis. In combination with CMB and SNIa data, the bispectrum can provide tight constraints on dark energy parameters. In a more recents work [147], it was shown that the bispectrum can provide 30%, 13% and 57% improved contraints on the neutrino mass M_ν, Ω_m and A_s for a LSST-like survey.

3.3.2 Peaks and Minimum counts

Peak statistic has been studied in the literature extensively and hence positioned as one of the most popular non-Gaussian statistic to study the weak lensing field [157, 158, 149]. Peaks are computed by counting the local maxima of a shear or convergence map, imprinted by massive structures along the line of sight. These allow us to extract the information from the non-linear regime, being complementary to the information from the variance of the field. In a 2D map, the peaks correspond to pixel values that are higher than their surrounding eight pixels. Conversely, minimum counts correspond to pixels whose values are lower than their eight neighbors. Physically, minima trace the underdense region of the convergence field, and hence can provide complementary information to the peak counts [21].
The first application of peak counts on real data was done by [148], for CFHTLenS data. The authors found that the combination of the power spectrum with the peak counts reduces the area of parameter countours by a factor ~ 2 compared to the power spectrum alone, and concluded that the peak counts are less sensitive than the power spectrum to systematic effects. In another work [159], the authors found that cosmological constraints based on the power spectrum and, separately, on the peak counts, are biased by magnification systematics in different directions, and hence a combined analysis can mitigate the impact of this systematic in the final results. More recently, [157] performed an analysis using the peak counts for DES-Y1 real data and found a 20% improvement in precision by combining the peak counts and the two-point function compared to the 2PCF alone.
Regarding the minimum count analysis, [21] shows that when it is applied to convergence

maps, the constraning power is slightly less than the peaks counts, although a combination of peak counts and minimum counts can outperform the angular power spectrum in constraning A_s by a 68% for a LSST-like survey.

3.3.3 Probability distribution function

The probability distribution function is a promising tool to access the non-Gaussian information of the weak lensing field, with applications in the convergence κ field. It is sensitive to cosmological parameters [151] and can provide precise constraints on the neutrino mass, Ω_m, σ_8 and w_0. In practice, the convergence PDF is built upon histogramming pixels of normalized convergence maps. In [152], the authors propose a theoretical model for the PDF based on the halo model with accurate predictions in the high PDF tail, whereas the authors in [150] make use of large deviation statistics to model the PDF at mildly no linear scales. Even though both approaches have found success at different scales, emulators based on interpolation methods that are trained on N-body numerical simulations have provided accurate theoretical predictions as well [160]. In [151], it is shown that the non-linear growth produces non-Gaussianities in the PDF. They also show that PDF shrinks the errors on neutrino mass and σ_8 by 20%, compared to the angular power spectrum alone in a tomographic analysis for a LSST-like survey. Also, it has been shown that it is capable of shrinking the errors in 27% for a Euclid-like survey compared to the power spectrum alone.

3.3.4 Minkowsi Functionals

Minkowski functionals (MF) are a set of mathematical descriptors of a field; invariants under rotations and translations. To characterise a D-dimensional random field, a set of $D+1$ minkowski functionals can be obtained [161]. For a two-dimensional map, three MF exist:

$$V_0(\nu) = \int_{\Sigma_\nu} \frac{1}{A} \mathrm{d}a \tag{3.32}$$

$$V_1(\nu) = \int_{\partial\Sigma_\nu} \frac{1}{4A} \mathrm{d}l \tag{3.33}$$

$$V_2(\nu) = \int_{\partial\Sigma_\nu} \frac{1}{2\pi A} \mathrm{d}l. \tag{3.34}$$

where $\mathrm{d}a$ is the area element, and $\mathrm{d}l$ the line element. MFs are obtained by means of excursion sets Σ_ν (with boundaries $\partial\Sigma_\nu$), i.e. regions where the pixel values exceed a threshold level ν. In this sense, the excursion set corresponds to a set of simple disjoint regions, or multiply connected regions above the threshold value. The MF V_0 in Eq. (3.32) is the area covered by the excursion set and V_1 defines the length of the boundary, all at a given value ν. Finally, the MF V_2 defines the connectivity of the field, and it is obtained as the difference between the total number of connected regions $> \nu$ and the number of connected regions $< \nu$. In practice, many measurements of the MFs can be obtained by varying ν. In [153, 162], it is confirmed that MFs can capture non-Gaussian information in the κ weak lensing field. The authors in [153] show that MFs alone can outperform the constraints with respect to the power spectrum alone by a factor of four for a LSST-like survey, whereas in [162] the MFs can also outperform power spectrum for constraining the dark energy equation of state.

Chapter 4

A Bayesian solution to neural-net assisted inference

4.1 Overview

Cosmology currently finds itself in the fortunate state of collecting ever more, and in particular, ever more precise data. Upcoming Stage-IV surveys such as Euclid [163], Rubin-LSST [15], and SKA[164] will provide data vectors with multiple million data points, yielding sub-percent precision on numerous cosmological parameters. Like-wise, the trend of collecting vaster, more precise data is repeated in all other disciplines of astronomy that also enjoy a progress in instrumentation.
In order to prepare for the arrival of real data, we need to adjust pipelines according to analysis choices. These choices come from iteratively testing the mitigation of a wide range of systematic effects (such as baryonic feedback), optimize survey strategies, forecasting theoretical models beyond ΛCDM, testing the performance of cosmological probes combinations, among others. To do so, simulated likelihood analyses are performed repeatedly in order to investigate the aforementioned science cases, which itself implies a large computational cost. Furthermore, the likelihood analysis for the cosmic shear power spectrum and other large-scale structure probes[1] entails the execution of accurate but slow Boltzmann codes[2] such as `CAMB` or `CLASS` to compute the theoretical predictions. The running of such Boltzmann solvers in the sampling of cosmological parameters posteriors is known to become the bottleneck for Stage-IV surveys analysis. On the other hand, the increasing precision of data demands an equal increase in the precision of theoretical models. All these situations pose computational challenges for parameter inference and hence precise and faster numerical methods are needed. In the past decade, methods based on machine learning have become the cornerstone of data-driven cosmology. In particular, the training of machine learning models that emulate the outputs of Boltzmann solvers have been considered as a compelling solution to reduce the CPU hours needed to perform inference. As artificial neural networks (hereafter, neural nets) are – once trained – compellingly fast, astronomy lately sees a surge of neural-net based emulation of theoretical models and likelihoods. We refer to those cases as

[1] And the cosmic microwave background anisotropies power spectra.
[2] Besides the fact that such codes do not provide accurate model predictions in the non-linear regime.

neural-net assisted inference as the neural net is directly involve in the likelihood evaluations. [18] provide neural net predictions for half a million data points of weak lensing, galaxy clustering and their cross-spectrum of a Euclid-like survey (sky coverage and tomographic bins), computing the posterior of seven cosmological parameters when also emulating the theoretical predictions for a Planck-like [165] observation of the cosmic microwave background. [17] provide a neural net emulator of the matter power spectrum and CMB power spectra (TT, EE and ET), with accurate predictions in the 5σ region around *Planck* 2018 [2] best fit, and also facilitating the incorporation of nuisance parameters. A different approach is adopted in [166], where trained neural nets are used to replace expensive integrations of perturbation equations inside Boltzmann solvers, instead of emulating the power spectra at once. Many other approaches have been proposed, by means of different machine learning techniques or more complex cosmological models [167, 111, 113].

As reviewed in Chapter 2, any neural net is by design an approximator, which means it learns some target function during the training process, and output approximations of such function as accurate as possible. Therefore, a far unsolved problem in neural-net assisted inference is how to avoid that the neural net's approximation error will bias the parameter inference. It is vital to have a clean solution for propagating its error and safe-guard the inference in the era of precision cosmology.

This chapter presents, for the first time, a Bayesian solution to propagating a neural net's approximation error and thereby debiasing parameter inference. The solution consists in a new posterior distribution for the parameters, assisted by trained neural nets. In this work, the neural nets predict the theoretical linear matter power spectrum in the likelihood evaluations. Crucially, our solution is expected to work for a wide range of successfully and rather unsuccessfully trained neural nets. As presented in our paper Grandón et al. [24], our method provides the missing element to judge the accuracy of a posterior if it cannot be computed based on an infinitely accurate theory code.

The Bayesian solution to propagating biases and approximation errors in a neural-net assisted inference is presented in Sect. 4.2. The main results of this work are Eq. (4.7) and Eq. (4.8), which give the adapted posterior function that replaces the posterior yielded from a non-approximate code. Eq. (4.7) is the special case of Eq. (4.1) if the data follow a Gaussian sampling distribution and the neural net's errors can be approximated as Gaussian too, as described in Sect. 4.3. Sect. 4.4 compares this adapted posterior to the 'naive' posterior gained from wrongly assuming the neural net has an infinite precision. We demonstrate two cases of varying neural net training success. Our results are discussed in Sect. 4.5.

4.2 Bayesian posterior for neural-net assisted inference

In this section, we assume that a neural net has been trained to *learn* the mapping from an input data to some target observable. Hereafter, we refer to the input data as *input parameters* or simply *parameters*, as in cosmology, it is very common that the input layer of the neural net consists of cosmological parameters. In this context, the target observable can be any parametric function that we know is cosmology-dependent, such as the comoving distance, the growth function, the CMB power spectrum or the matter power spectrum, among others.

We denote the neural net's output predictions of the target observable as $\boldsymbol{\mu}_{\rm NN}(\boldsymbol{p})$, and we assume that the neural net has undergone validation by testing its predictions on an inde-

pendent validation set that contains a number V of samples of parameters $\boldsymbol{p}$.
The neural net's approximation error can then be evaluated at each sample of this validation set, and we define the discrepancy between the neural net prediction and the error-free theoretical prediction $\boldsymbol{\mu}$ at $\boldsymbol{p}$ to be $\boldsymbol{\Delta}(\boldsymbol{p}) = \boldsymbol{\mu}(\boldsymbol{p}) - \boldsymbol{\mu}_{\rm NN}(\boldsymbol{p})$. However, in the parameter space between the validation points, the approximation error will remain unknown. Hence, as $\boldsymbol{\Delta}(\boldsymbol{p})$ is unknown for all but a set of points in parameter space, we treat it as a random variable in the Bayesian sense. Therefore, we shall treat $\boldsymbol{\Delta}(\boldsymbol{p})$ as a random variable.
As reviewed in Chapter 2, any random variable is described by a probability distribution. We denote the distribution as $\boldsymbol{\Delta}(\boldsymbol{p}) \sim \mathcal{D}(\boldsymbol{\Delta}(\boldsymbol{p})|\cdot)$, where the placeholder $\cdot$ denotes variables that specify the distribution further, such as the mean and the variance if $\mathcal{D}$ is a Gaussian distribution. In the case of poor neural net training, the neural net will exhibit large approximation errors that strongly vary with the values of $\boldsymbol{p}$. For instance, there might be a region in parameter space where the neural net successfully approximates the parametric theory explanation of the data, while other regions may result in poor neural net performance with significantly large approximation errors. In this case, we would have $\mathcal{D}(\boldsymbol{\Delta}(\boldsymbol{p})|\boldsymbol{p}, \cdot)$, indicating that the distribution of the neural net approximation errors explicitly depends on the cosmological parameters themselves. A Bayesian solution to propagating the neural net's errors would then still be possible as long as $\mathcal{D}$ can be accurately estimated. One approach to achieve this is by histogramming the approximation errors over different patches of the parameter range. Another solution to a spatially-varying distribution $\mathcal{D}$ would be to model it with a Gaussian process, see e.g. [110], though they find that the posterior using a Gaussian process model differs more from the true posterior more than if the error were ignored all together. On the other hand, using a Gaussian process for the sake of modelling spatially varying approximation errors necessitates that hyperparameters must be set. To avoid such hyper-parameters, we hence prefer to treat the approximation error as random also at the validation points. This is a conservative choice as it treats the approximation error at the validation points with the same uncertainty as in between the validation points.
We hence adopt the approximation that for all values of $\boldsymbol{p}$, the neural net's errors follow the same distribution, namely $\boldsymbol{\Delta} \sim \mathcal{D}(\boldsymbol{\Delta}|\cdot)$, where the placeholder $\cdot$ now excludes the parameters $\boldsymbol{p}$. In numerical experiments of reasonably trained nets we found it to hold with excellent precision. In summary, our notation proceeds as follows

- $\boldsymbol{p}$: Model parameters,
- $\boldsymbol{\mu}_{\rm NN}(\boldsymbol{p})$: Neural net prediction,
- $\boldsymbol{\mu}(\boldsymbol{p})$: Theory prediction,
- $\boldsymbol{\Delta}(\boldsymbol{p})$: Neural net approximation error,
- $\mathcal{D}(\boldsymbol{\Delta}|\cdot)$: Probability distribution of the approximation error.

4.2.1 Marginalizing the neural net error and bias

We assume the neural net provides predictions for n data points, and we denote the scientific data by $\boldsymbol{x}$. The likelihood of the data $\mathcal{L}(\boldsymbol{x}|\boldsymbol{p}, \boldsymbol{\Delta})$ is dependent on the neural net's approximation error $\boldsymbol{\Delta}$: if the neural net's prediction is erroneous, other values of the parameters $\boldsymbol{p}$ might fit the data better in order to compensate for the neural net's error. This would automatically lead to a bias of the cosmological parameters constraints. Such a bias in the inferred parameters is the key worry we aim to eliminate by accounting for the neural net's

error.
We also assume that the prior probability of the parameters $\pi(\boldsymbol{p})$ is independent of the prior probability of approximation errors $\pi(\boldsymbol{\Delta})$. This will hold if the neural net was trained according to best practice, in particular if it was provided with 'sufficiently' many training points in each region of the prior range. In brief, the training has to guarantee that $\pi(\boldsymbol{p}, \boldsymbol{\Delta})$ factorizes. Counterexamples can be engineered. Namely $\pi(\boldsymbol{p}, \boldsymbol{\Delta})$ will not factorize when the neural was trained on highly unrepresentative samples of $\boldsymbol{p}$, for example having no training samples in certain subregions, or if $\boldsymbol{\Delta}$ particularly deviates from average in certain region of parameter space. Propagating the neural net's prediction error then implies that we must marginalize $\boldsymbol{\Delta}$. The posterior of parameters inferred from an imperfect neural net prediction is hence the compound of the posterior of vanishing neural net error with the distribution of the error

$$\begin{aligned}\mathcal{P}(\boldsymbol{p}|\boldsymbol{x}) &= \int \mathcal{P}(\boldsymbol{p}, \boldsymbol{\Delta}|\boldsymbol{x})\mathrm{d}^n\Delta \\ &= \int_{-\infty}^{+\infty} \mathcal{P}(\boldsymbol{p}|\boldsymbol{\Delta}, \boldsymbol{x})\mathcal{D}(\boldsymbol{\Delta}|\cdot)\mathrm{d}^n\Delta \\ &\propto \int_{-\infty}^{+\infty} \mathcal{L}(\boldsymbol{x}|\boldsymbol{p}, \boldsymbol{\Delta})\pi(\boldsymbol{p})\mathcal{D}(\boldsymbol{\Delta}|\cdot)\mathrm{d}^n\Delta,\end{aligned} \tag{4.1}$$

where the proportionality sign arises by not having written out the Bayesian evidence. Eq. (4.1) describes that parameter inference can be safeguarded against a neural net's approximation error by specifying the data's likelihood as a function of physical parameters and the neural net's approximation error. A prior is set on the parameters, and it is compounded with the distribution $\mathcal{D}(\boldsymbol{\Delta}|\cdot)$. The posterior of physical parameters is hence the marginal over the neural net's approximation error.
We shall now specialize Eq. (4.1) to the most common case of Gaussian data.

4.2.2 Specialization to Gaussian data

In this section we provide a closed-form expression for the posterior of marginalized neural net error for the case of Gaussianly distributed data. In line with this assumption, the likelihood now reads

$$\mathcal{L}(\boldsymbol{x}|\boldsymbol{p}) \propto \exp\left(-\frac{1}{2}[\boldsymbol{x} - \boldsymbol{\mu}(\boldsymbol{p})]^T \mathsf{C}^{-1}[\boldsymbol{x} - \boldsymbol{\mu}(\boldsymbol{p})]\right), \tag{4.2}$$

where $\boldsymbol{\mu}$ is the error-free prediction of the data, evaluated as a function of the parameters $\boldsymbol{p}$, and C^{-1} is the inverse covariance matrix of the data.
Defining $\boldsymbol{\xi}(\boldsymbol{p}) = \boldsymbol{x} - \boldsymbol{\mu}(\boldsymbol{p})$, the likelihood with neural net error is

$$\mathcal{L}(\boldsymbol{x}|\boldsymbol{p}, \boldsymbol{\Delta}) \propto \exp\left(-\frac{1}{2}[\boldsymbol{\xi}(\boldsymbol{p}) - \boldsymbol{\Delta}]^T \mathsf{C}^{-1}[\boldsymbol{\xi}(\boldsymbol{p}) - \boldsymbol{\Delta}]\right). \tag{4.3}$$

We define the mean of the neural net approximation error as

$$\bar{\boldsymbol{\Delta}} = \frac{1}{V}\sum_{v=1}^{V} \boldsymbol{\Delta}_v, \tag{4.4}$$

where v runs over the samples of the validation set. Accordingly, the covariance matrix of the neural net's prediction error is

$$\Sigma = \frac{1}{V-1}\sum_{v=1}^{V}(\boldsymbol{\Delta}_v - \bar{\boldsymbol{\Delta}})(\boldsymbol{\Delta}_v - \bar{\boldsymbol{\Delta}})^T. \tag{4.5}$$

For an excellently trained neural net, the mean $\bar{\boldsymbol{\Delta}}$ will vanish, meaning the neural net provides an unbiased prediction of the theory function over the entire parameter range. On the other hand, the covariance matrix Σ will contain large variances and covariances for a neural net whose training did not yield very accurate predictions of the target function. The more training samples were seen, and the smaller the loss during training, the smaller the elements of Σ will become. The Gaussian distribution of $\boldsymbol{\Delta}$ then is

$$\mathcal{D}(\boldsymbol{\Delta}) \propto \exp\left(-\frac{1}{2}[\boldsymbol{\Delta} - \bar{\boldsymbol{\Delta}}]^T\Sigma^{-1}[\boldsymbol{\Delta} - \bar{\boldsymbol{\Delta}}]\right). \tag{4.6}$$

In the limit $\Sigma \to 0$, this distribution tends to Dirac's delta distribution, correctly reflecting that no errors are folded into the posterior of parameters.
As the compound of a Gaussian with another Gaussian yields a third Gaussian, inserting Eq. (4.3) and Eq. (4.6) into Eq. (4.1) yields

$$\boxed{\mathcal{P}(\boldsymbol{p}|\boldsymbol{x}) \propto \exp\left(-\frac{1}{2}[\boldsymbol{\xi}(\boldsymbol{p}) - \bar{\boldsymbol{\Delta}}]^T\Psi[\boldsymbol{\xi}(\boldsymbol{p}) - \bar{\boldsymbol{\Delta}}]\right)\pi(\boldsymbol{p}),} \tag{4.7}$$

where

$$\Psi = (\mathsf{C} + \Sigma)^{-1}, \tag{4.8}$$

and $\bar{\boldsymbol{\Delta}}$ is given by Eq. (4.4). Equation (4.7) is the main result of this work, which is widely applicable with minimum effort. For a neural net that produces approximately Gaussian-distributed inaccuracies in its predictions, Eq. (4.7) propagates the net's uncertainties, and effectively removes neural net biases in the inference of cosmological parameters.

4.3 Demonstration of parameter inference

We demonstrate the inference of cosmological parameters H_0, Ω_c and Ω_b when the theory prediction for the data has been replaced by the neural net output prediction. We will highlight two cases of how (un)successful neural net training deforms the posterior of inferred parameters. This analysis is performed using the public code `emcee` [168], with the chains visualized using GetDist [169].
For our demonstration, we simulated a cosmological analysis of the linear matter power spectrum $P(k)$, where k is the wavenumber. Our data range spans the range $k \in [0.0033\,h/\mathrm{Mpc}, 0.4\,h/\mathrm{Mpc}]$. Originally, the theory predictions for the matter power spectrum would come from a Boltzmann solver or an analytic fitting function, as discussed in Chapter 1. For this case, the notation defined before corresponds to

- $\boldsymbol{p}$: Cosmological parameters $\{H_0, \Omega_\mathrm{c}, \Omega_\mathrm{b}\}$,
- $\boldsymbol{\mu}(\boldsymbol{p})$: `CAMB` prediction for the matter power spectrum ,

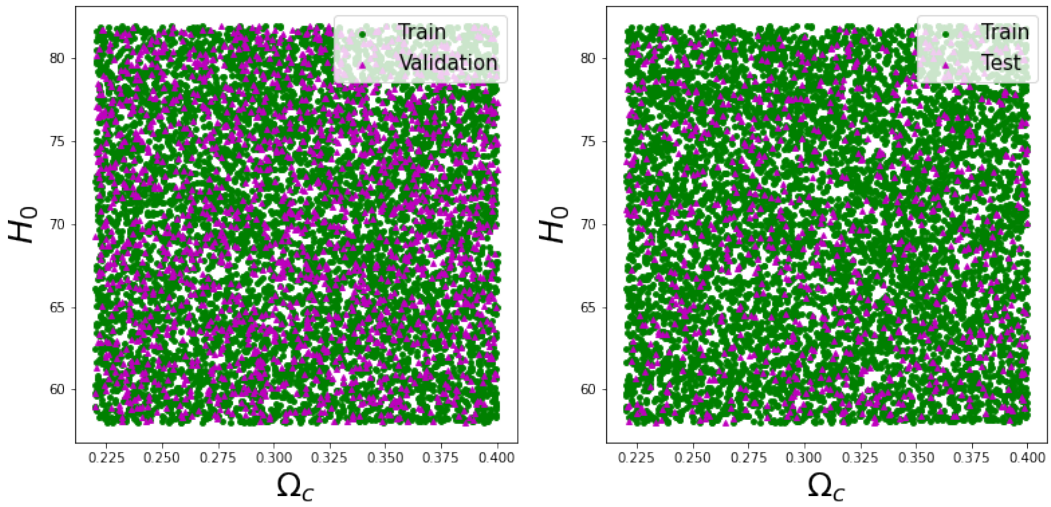


Figure 4.1: The training, validation and test set samples distributed across parameter space.

	Prior ranges	Fiducial cosmology
Ω_c	$\mathcal{U}$[0.22,0.4]	0.269
Ω_b	$\mathcal{U}$[0.02,0.06]	0.0494
H_0	$\mathcal{U}$[58,82]	66.8

Table 4.1: Prior ranges and input cosmology for Ω_c, Ω_b and H_0.

- $\boldsymbol{\mu}_\mathrm{NN}(\boldsymbol{p})$: Neural net prediction for the matter power spectrum.

In order to train a neural net, we generate 10,000 power spectra with `CAMB` by randomly drawing from the uniform prior given in Table 4.1 for the parameters H_0, Ω_c and Ω_b, which are the Hubble constant, the density parameter of cold dark matter, and the density parameter of baryons, respectively. The fiducial cosmology depicted in Table 4.1 is used to generate the data $\boldsymbol{x}$, which is explained in more detail later in Sec. 4.4. We split the `CAMB` predictions into batches of 70% for training, 20% for the validation set and 10% for test. These batches are used to find the best architecture for the neural nets, with the test set serving to evaluate the performance of the neural net on an unseen set of points. The training, validation and test points in parameter space are shown in Fig. 4.1. As we can see, all points in each batch span the parameter space equally. This means that the neural net training is performed in all regions of the parameter space and that validation of the neural net architecture is tested in all representative regions. We also segmented the power spectra into two wavenumber ranges of $k \in [0.0033\,h/\mathrm{Mpc},\ 0.0444\,h/\mathrm{Mpc}]$ and $k \in [0.044\,h/\mathrm{Mpc},\ 0.4\,h/\mathrm{Mpc}]$, and trained a neural net for each k-range individually. The segmentation into k-ranges is not necessary, but advantageous because it reduces the layer sizes and number of neurons, and hence reducing the neural net trainable parameters.

For the training we engineered two cases. One is a successfully/accurately trained neural net whose error on the test set is at the percent level on average over k. The second case is an

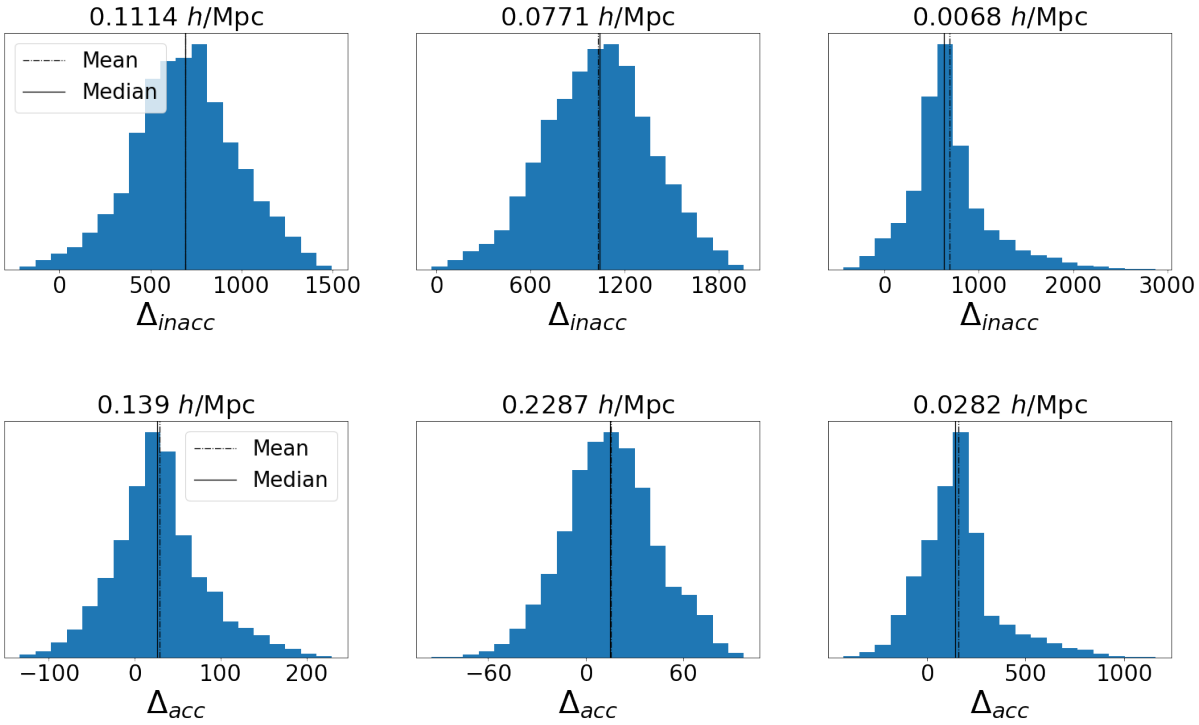


Figure 4.2: Histograms of the approximation error $\boldsymbol{\Delta}$ at specific wave number k values obtained for all the points in parameter space in the fourth set $\mathbf{V}$. Upper panel are results for the inaccurate neural net, and lower panel for the accurate model. From left to right, the two first histograms represent the general tendency of our histograms, where gaussianity condition is fullfilled. The third column shows histograms where the distributions slightly deviate from gaussianity.

unsuccessfully/inaccurately trained neural net where the prediction error is up to 40%. Such a big error would clearly bias the posterior of cosmological parameters unless it is corrected for.

To proceed with the Bayesian error propagation, we generate a fourth set $\mathbf{V}$ including new 10,000 power spectra at random $\boldsymbol{p}$ from the same prior range. We found it advantageous if the fourth set includes more points than the test set. For each point v in $\mathbf{V}$ it is possible to obtain the neural net approximation error $\boldsymbol{\Delta}_v(\boldsymbol{p}) = \boldsymbol{\mu}(\boldsymbol{p}) - \boldsymbol{\mu}_{\rm NN}(\boldsymbol{p})$ that accounts for the error in the prediction, where v runs from 1 to 10,000.

In this derivation, as stated in Section 4.2.2, we assume $\boldsymbol{\Delta}$ is a random variable that follows a Gaussian distribution, so it is crucial to check this assumption by computing histograms for different k. Fig.4.2 shows histograms for the two models at different k. As a very simple test, the mean and the median are displayed as we expect both match for a normally distributed sample. For both upper and lower panels of Fig. (4.2), the first two figures from left to right show cases where Gaussianity assumption is completely fullfilled. The third column shows handpicked examples where the Gaussian approximation is stretched, but are very rare cases and no more extreme cases were observed. We also computed skewness and curtosis and ensured that no bimodal structure occurs in the histograms. By computing these higher-order moments and applying Kolmogorov-Smirnov test for all histograms of k, we found that our

assumption of Gaussianity is indeed supported for both models. In particular, there were no bimodal distributions or heavy tails for the distribution $\boldsymbol{\Delta}$, which would most quickly invalidate the approximation of $\boldsymbol{\Delta}$ being Gaussianly distributed. On the other hand, we check that $\mathcal{D}(\boldsymbol{\Delta})$ does not depend on parameters by doing a sub-volumes analysis. By sub-volume we imply small regions of the parameter space, where we study if Gaussianity also holds and if the mean $\boldsymbol{\Delta}$ in these sub-volumes falls well within the bulk in the general $\mathcal{D}(\boldsymbol{\Delta})$ distribution shown in Fig. 4.2. This is shown in Fig. 4.3, where most of the sub-volumes mean fall into the 2σ region. Also, this holds for both the accurate and the inaccurate neural net. Hence, $\mathcal{D}(\boldsymbol{\Delta})$ does not depend on cosmological parameters which fulfills our assumption.

As presented in Eq. (4.7), our adapted posterior accounts for the bias and the approximation error by defining $\bar{\boldsymbol{\Delta}}$ and $\boldsymbol{\Sigma}$. From $\boldsymbol{\Delta}_v(\boldsymbol{p})$, we apply Eq. (4.4) and Eq. (4.5) to obtain $\{\bar{\boldsymbol{\Delta}}_{\rm acc},\boldsymbol{\Sigma}_{\rm acc}\}$ and $\{\bar{\boldsymbol{\Delta}}_{\rm inacc},\boldsymbol{\Sigma}_{\rm inacc}\}$ for the accurate and inaccurate models, respectively. In this demonstration, the quantities $\boldsymbol{\Delta}$ and $\boldsymbol{\Sigma}$ do not vary either using 8,000, 9,000 or 10,000 points of V, but we choose to use the largest number possible to provide a good approximation of the distribution for $\boldsymbol{\Delta}$. Also, the sub-volumes analysis shows that the approximation error is constant throughout parameter space. Hence, we can safely state that $\bar{\boldsymbol{\Delta}}_{\rm acc}$ and $\bar{\boldsymbol{\Delta}}_{\rm inacc}$ represent the approximation error of the neural net's for this particular case. The resulting $\bar{\boldsymbol{\Delta}}_{\rm acc}$ and $\bar{\boldsymbol{\Delta}}_{\rm inacc}$ are displayed in Fig. 4.4, and compared to the power spectra at the fiducial cosmology in Table 4.1. The lower panel shows the fractional difference that goes up to 5% compared to the input power spectra for the accurate neural net, and 40% for the inaccurate model. Hence, we expect the bias of the best fit of parameters is small or even negligible when comparing the resulting posteriors with and without error propagation for the accurate model. In Fig. 4.5 we see the variance and covariances for both models, with the unsuccessfully trained model display larger (co)variances, hence impacting strongly the covariance matrix of the adapted posterior.

4.4 Results

Once all the components of Eq. (4.7) are derived, we constrain cosmological parameters for three scenarios: when the theory prediction is computed by `CAMB`, and when the prediction is done with the accurate neural net model, and finally with the inaccurate model. The last two parameter inference cases are referred to as *neural-net based inference*. For both, we consider the cases when the posterior does and does not propagate the neural net's error, where the former is the adapted posterior proposed in Eq. (4.7) and the latter the traditional inference approach, i.e. a naive plug-in of the neural net into the standard Gaussian likelihood. We generated synthetic matter power spectra as our data vector $\boldsymbol{x}$ by a random draw from a Gaussian distribution with mean the fiducial cosmology: $H_0 = 66.8$ km s^{-1} Mpc^{-1}, $\Omega_{\rm cdm} = 0.269, \Omega_{\rm b} = 0.0494$; and a covariance matrix C as the square of the fiducial power spectrum. The mean is displayed as a black solid line in Fig. 4.4, where the synthetic data scatters around.

The results for the parameter inference in all the scenarios are shown in Fig. 4.6 and Fig. 4.7. In all figures, orange contours stand for the posteriors obtained from the inaccurate neural net, purple posteriors for the accurate net, and light blue contours for the inference with `CAMB`. The input parameters are indicated in all corner plots by dashed black lines.

- In Fig. 4.6, we depict the 'naive' posteriors, yielded by simply replacing the (taken as

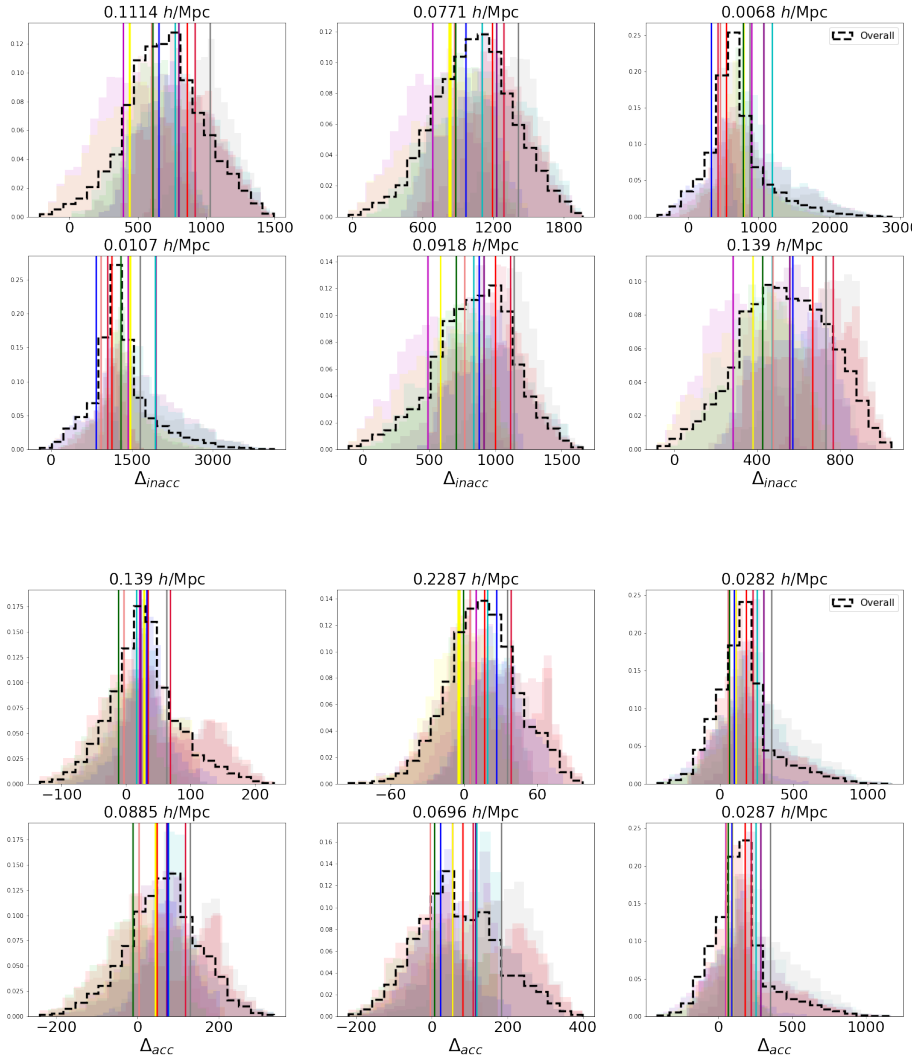


Figure 4.3: Sub-volume histograms for different wavenumber k, where $\boldsymbol{\Delta}$ is obtained with the innacurate and accurate neural net. The overall distributon over all parameter space is depicted in a dashed-black contour. Shaded colors in the background represent the obtained histograms for different chosen sub-volumes, where the means are shown in vertical coloured lines. We conclude that the neural net approximation error does not directly depend on the sub-volume and hence not directly on the cosmological parameters themselves.

infinitely accurate) `CAMB` mean $\boldsymbol{\mu}(\boldsymbol{p})$, or the neural net's prediction, neither propagating the neural net's bias nor its approximation error around the true mean. Orange contour show a noticeable non-zero bias in the best fit, while the purple contour mostly match the `CAMB` blue contour. Hence, Fig. 4.6 demonstrates that parameter inference will incur biases if the neural net's approximation error is not propagated.

- Accordingly, Fig. 4.7 demonstrates the main results of this work, namely the adapted-posterior that propagates the neural nets errors into the final parameter inference. For both neural nets, the posterior of parameters is widened up as a consequence of the covariance matrix of the neural net approximation error shown in Fig. 4.5. As this

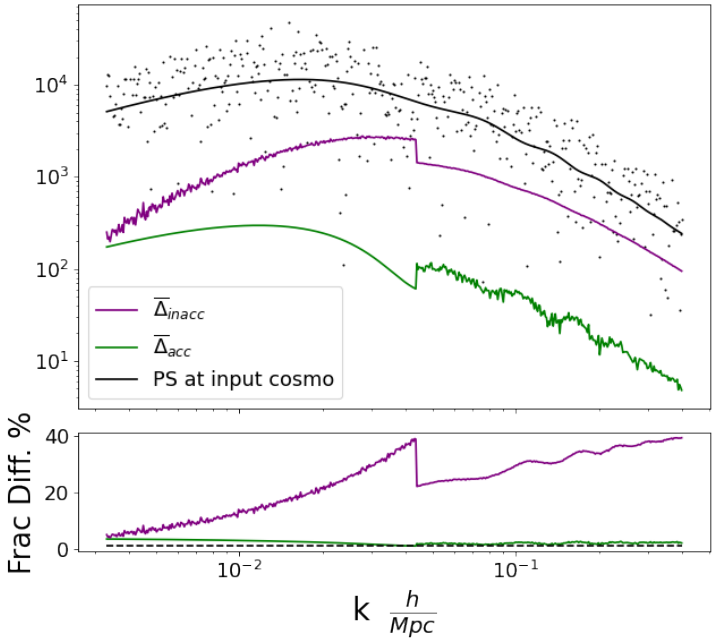


Figure 4.4: The prediction errors for the two neural nets are displayed and compared to the matter power spectra at the input cosmology. The prediction error for the accurate model $\bar{\mathbf{\Delta}}_{\mathrm{acc}}$ is shown in green and the prediction error $\bar{\mathbf{\Delta}}_{\mathrm{inacc}}$ for the inaccurate model in purple. The scatter plot represents the synthetic matter power spectra used for the inference.

covariance matrix adds to the original covariance matrix of the data, the uncertainty on the parameter constraints increases. This is much more prominent for the inaccurately trained neural net. Also seen is that the subtraction of the neural net's bias leads to a debiasing of the posterior: The fiducial cosmology is included in the posterior after debiasing, but is excluded beforehand. This evidences that our Eq. (4.7) is a simple but powerful method for safeguarding parameter inference against the approximative nature of neural nets.

4.5 Discussion

In this work we derived an adapted posterior that propagates the error of the predictions of neural-net based inference. The adapted posterior results in Eq. (4.7) and (4.8), and implies a modified precision matrix and a bias correction when the error in the neural net prediction is non-negligible. We investigate the impact of the neural net prediction error on a cosmological analysis for constraining parameters Ω_{c}, Ω_{b} and H_0. The method is easy to implement when the prediction errors follow a Gaussian distribution, which must be checked. It is worth noting that for this approximation, Gaussianity holds even if a few examples present mild longer tails. However, non-negligible higher moments can emerge in cases where the distribution is bimodal or where heavy tails are present. In those cases, it is crucial to either extend the

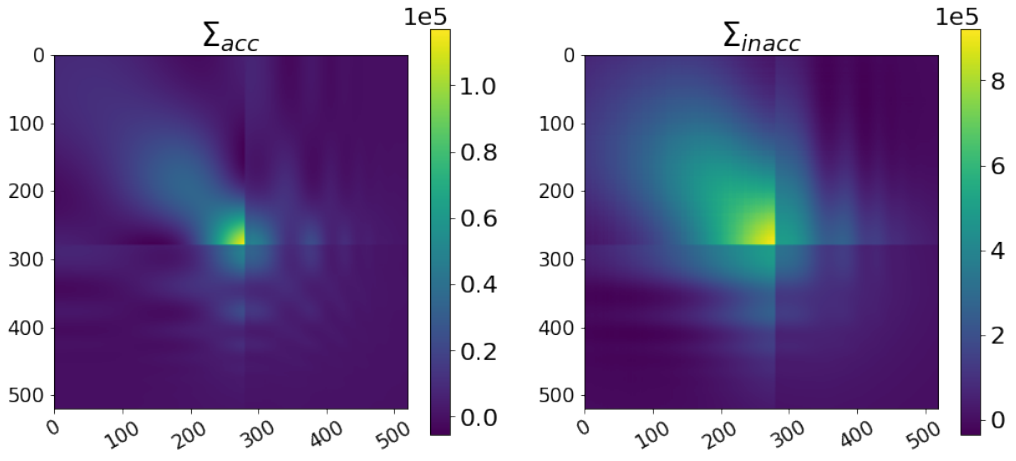


Figure 4.5: Left: the covariance matrix $\Sigma_{\rm acc}$. Right: $\Sigma_{\rm inacc}$. We see that the badly trained neural net causes not only a higher bias as seen in Fig. (4.4) but additionally also more correlated errors in its prediction.

analysis and propose a new analytic form for $\mathcal{D}(\mathbf{\Delta})$, or to improve the neural net training, in order to achieve Gaussian approximation errors. In this analysis, we found that Gaussianity holds in most of the trained models, making the method simple to apply. In Section 4.4 we show that error propagation can restore the input cosmology in the best-fit even if the neural net is inaccurate.

As many works in the literature had shown the advantages of accelerated neural-net based inference in the context of future Stage-IV cosmological surveys, we consider it critical to also include error propagation for neural nets inference. In the era of precision cosmology, where many efforts are still being done on the control of astrophysical or instrumentation systematic effects, it is also necessary to consider and characterise the biases that our modern statistical methods propagates into the inference.

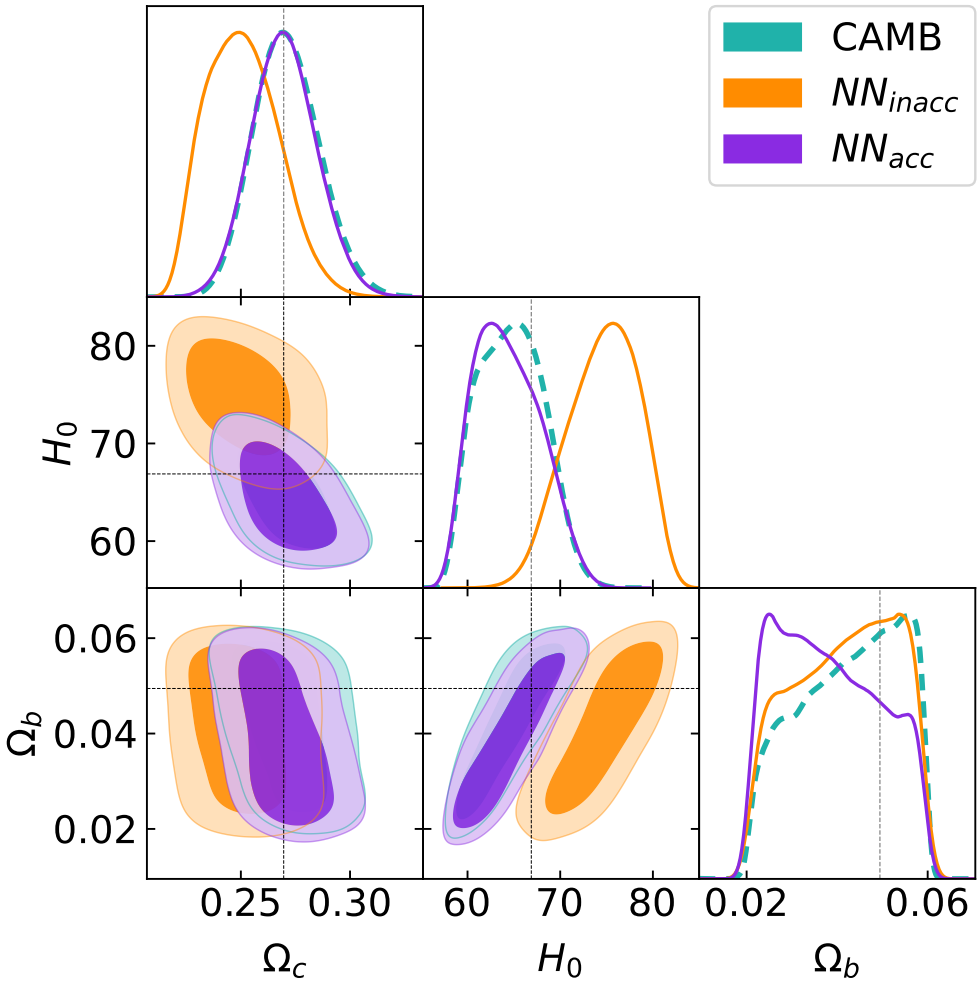


Figure 4.6: Marginal posterior contours obtained for neural-net based inference and normal Boltzmann-code based inference for synthetic matter power spectra data. The contours contain 68% and 90% posterior volume.

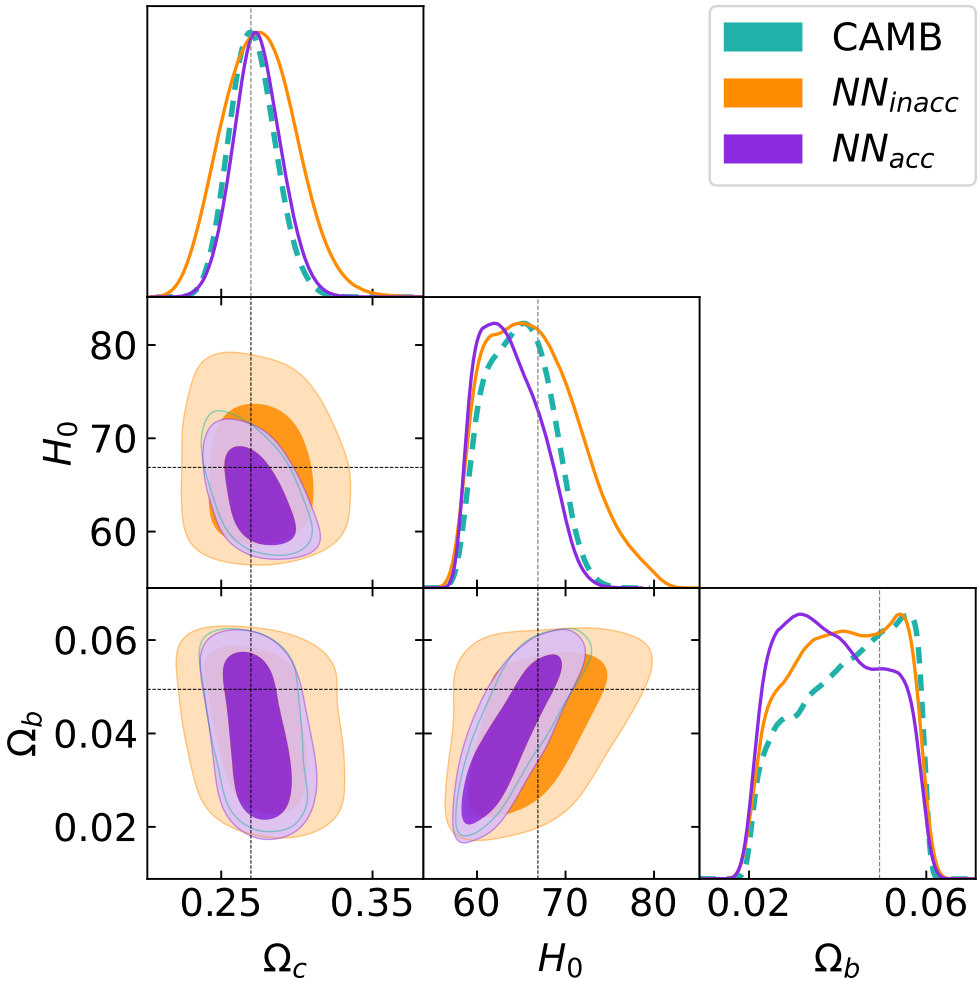


Figure 4.7: Marginal posterior contours obtained for neural-net based inference implementing the proposed adapted posterior. `CAMB` results are also displayed for comparison. The contours contain 68% and 90% posterior volume.

Chapter 5

The impact of baryons on weak lensing non-Gaussian statistics

5.1 Overview

Weak gravitational lensing (WL), i.e., small distortions of the observed shapes of galaxies due to the large-scale distribution of matter, has become a powerful probe for testing the current cosmological paradigm. It is sensitive to the total density of matter in the Universe, and its tomographic analysis allows us to trace the evolution of structure growth, affected by the accelerated expansion of the Universe and the clustering of matter.
The two-point correlation function or angular power spectrum estimators of the weak lensing signal have shown to be particularly sensitive to the matter clustering amplitude $S_8 = \sigma_8(\sqrt{\Omega_m/0.3})$. In recent years, the results obtained by the on-going weak lensing surveys have revealed a tension between their inferred value of S_8, and the value derived by *Planck* data. One approach adopted to explore this tension consists of understanding the systematic and astrophysical effects that may impact the cosmological constraints.
We already reviewed in Chapter 3 that baryonic processes imprint specific signatures on the small scales of the LSS, which can potentially be detected by non-Gaussian statistics. Therefore, exploiting statistical tools beyond the angular power spectrum that capture the non-Gaussian information at small scales presents an opportunity to deepen our understanding of the effects of baryonic processes on the LSS and cosmological parameters. This is a challenge to be targeted by the forthcoming Stage-IV surveys, such as the Vera Rubin Observatory LSST [15] and Euclid [163], where the precision of data demands a thorough study of the impact of baryons on the relevant scales of the weak lensing signal. Moreover, systematic effects might impact the statistical tools differently, and hence it is important to consider more than one statistical tool to study the weak lensing data. [21] investigate the impact of baryons on the cosmological parameters neutrino mass sum, Ω_m and A_s utilizing the weak lensing peaks and minimum counts for LSST-like weak lensing data at a single redshift. They found that a 0.5σ bias from the minimum counts alone, and up to 4σ bias in the inferred parameters from peak counts alone. [170] also study the impact of baryons from different cosmological hydrodynamic simulations on non-Gaussian statistics for noiseless weak lensing convergence maps, finding that baryons can modify the measured statistics up to 10% for

$z = 1$. However, they do not propagate the impact of baryons into cosmological inference. More recently, [160] found $< 0.4\,\sigma$ bias on the inferred S_8 from the weak lensing PDF when including baryonic feedback, however they apply aggressive scale cuts which were decided before analyzing the HSC real data.
In this chapter, we employ the cosmic shear power spectrum and non-Gaussian (NG) statistics such as peak counts, minimum counts, and probability distribution function (PDF) to investigate the influence of baryons on the inferred cosmological parameter S_8 using the Subaru Hyper Suprime-Cam (HSC) Year-1 weak lensing data. We often refer to all the statistics employed as *summary statistics*.
As presented in Grandón et al. [25], we first investigate the baryonic feedback effect on the summary statistics and on the inferred S_8 based on HSC-Y1 weak lensing convergence mocks, employing a data-vector contamination strategy. The baryonic feedback implementation is based on cosmological hydrodynamical simulations IllustrisTNG and BAHAMAS. We employ different analysis choices to determine the extent to which baryons can influence the summary statistics and introduce biases into the constraints on the amplitude of matter clustering S_8. In this way, we can analyze the scale cuts needed for different baryonic feedback scenarios. Finally, we conduct parameter inference using HSC-Y1 real data and compare the results with those derived from our mock-based analysis.

5.2 Subaru Hyper Suprime-Cam (HSC) survey

The HSC first-year data corresponds a catalog of galaxy shape measurements covering in total 136.9 deg^2 of the sky, splitted in six disjoint patches called: `XMM`, `GAMA09H`, `WIDE12H`, `GAMA15H`, `VVDS`, and `HECTOMAP` [171, 172, 27]. The year 1 data release consists in observations from March 2014 to April 2016, using the Subaru Hyper Suprime-Cam in five broadbands, *grizy* [173]. The redshift distribution of source galaxies in the Y1 catalog is shown in Fig. 5.1. These redshifts are determined in [8] based on the HSC five broadband photometry by using the machine learning-based method `MLZ`. For our analysis, we restrict the source redshift range to $0.3 < z_{\rm best} < 1.5$, where $z_{\rm best}$ is the best-fit photometric redshift determined by [8]. The resulting effective number density of galaxies per square degree is 17 arcmin^{-2}, after employing magnitude and redshift cuts. In this work, we consider four photometric redshift bins, with bin edges [0.3, 0.6, 0.9, 1.2, and 1.5]. However, we exclude the highest redshift bin from our analysis due to indications of either unmodelled effects in our mocks or systematic issues in the real data such as a mis-calibration of redshift in this bin [22]. This decision ensure a more conservative approach to our analysis. The redshift ranges and galaxy number density of each tomographic bin are summarized in Table 5.1.
Cosmological constraints based on HSC-Y1 measurements have already been released for a ΛCDM model, with an inferred amplitude of matter clustering $S_8 = 0.804^{+0.032}_{-0.029}$ based on the cosmic shear two-point correlation functions analysis [10], and $S_8 = 0.78^{+0.030}_{-0.033}$ based on the cosmic shear power spectrum [9].
As the deepest Stage-III galaxy weak lensing survey conducted to date, analyzing NG statistics of HSC data provides a great avenue to delve into the modeling and impact of systematic effects. This serves as a crucial stepping stone towards the upcoming Rubin LSST, enabling us to enhance our understanding of systematic challenges of future weak lensing surveys.

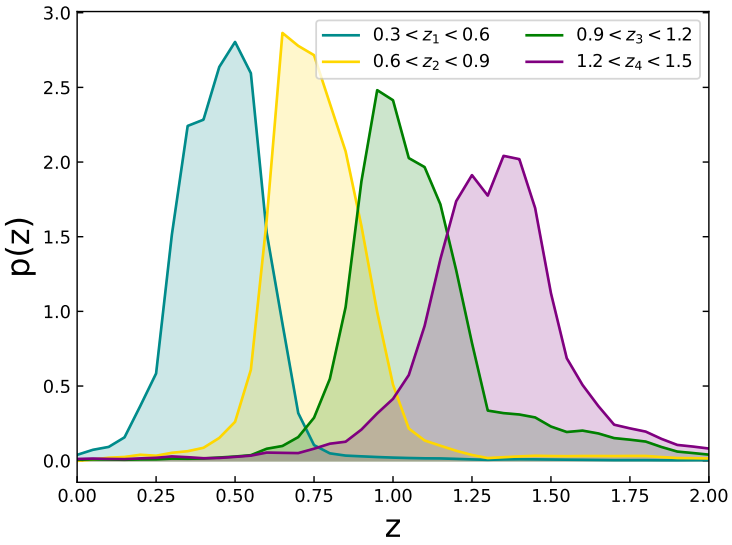


Figure 5.1: Redshift distribution for the four tomographic bins, based on the photometric redshifts estimates by [8] for HSC-Y1 data.

z-range	$n_g^{\rm eff}[\rm arcmin^{-2}]$
$0.3 < z < 0.6$	5.14
$0.6 < z < 0.9$	5.23
$0.9 < z < 1.2$	3.99
$1.2 < z < 1.5$	2.33

Table 5.1: Properties of individual tomographic bins. The $z-$range refers to the photometric redshift range that defines the four HSC tomographic bins, and the effective number density $n_g^{\rm eff}$ assumes the definitions of [1]. Note that the fiducial sample in our analysis is the first 3 bins of the HSC sample. The information of the highest photo-z bin is shown in grey to indicate it is not part of the fiducial analysis.

5.3 Data and simulations

In our analysis, we rely on a set of HSC Y1-like weak lensing convergence maps (also called κ maps or *mass maps*) that have been reconstructed from cosmological N-body and hydrodynamical simulations. Upon these maps, we build the analysis pipeline, which includes the likelihood and the inclusion of baryonic feeback through a contaminated data vector. In the following section, we provide an overview of the simulations.

5.3.1 N-body simulations

To perform the parameter inference, we utilize convergence maps based on two sets of N-body simulations.

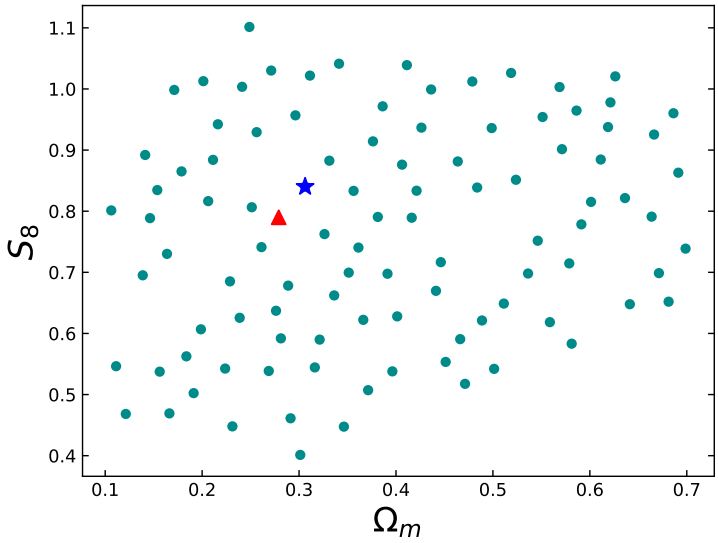


Figure 5.2: The distribution of cosmological models covered by the simulations. Cyan points illustrate the 100 models of the cosmo-varied simulations, with the blue-star depicting the model we employed for our mock data vector in the likelihood for all the statistics. The red triangle represents the cosmology of the fiducial maps.

- **Fiducial maps:** To construct the covariance matrix, we utilize 2268 realizations of HSC-Y1 weak lensing mocks at the Wilkinson Microwave Anisotropy Probe (WMAP) nine-year cosmology : $\Omega_{\rm b} = 0.046$, $\Omega_{\rm m} = 0.279$, $\Omega_\Lambda = 0.721$, $h = 0.7$, $\sigma_8 = 0.82$, and $n_s = 0.97$ [174]. These mocks are generated from 108 full-sky weak lensing maps presented in [175] and are specifically designed to include the associated HSC Y1 specifics, such as shape noise levels, variations in the lensing weight, image calibration uncertainties, the spatial inhomogeneity of source galaxies, redshift distribution uncertainties, and survey geometry. The method of the mock production has been presented in [26].
- **Cosmo-varied maps:** To perform cosmological inference, we need the predictions for the summary statistics for arbitrary cosmologies. For this purpose, we adopt the convergence maps based on the N-body simulations introduced in [176], which consists of ray-tracing mocks for 100 different cosmologies on the $\Omega_m - \sigma_8$ parameter space, with 50 realizations each model under the flat-sky approximation. The cosmological parameters span the ranges $\Omega_m \in [0.1, 0.7]$ and $\sigma_8 \in [0.4, 1.4]$, and considers the Hubble parameter fixed at $h = 0.6727$, the baryon density $\Omega_b h^2 = 0.02225$, and the spectral index $n_{\rm s} = 0.9645$. Fig. 5.2 shows the distribution of the simulations in the Ω_m- S_8 plane, also depicting the fiducial maps cosmology as a red triangle in parameter space. For a complete description of the simulation implementation and the mock production process, we refer the reader to [177, 26].

5.3.2 Hydrodynamical simulations

As we lack a complete analytical modelling of the wide range of baryonic processes that impact the distribution of matter, a state-of-the-art alternative is to employ hydrodynamical simulations, where baryonic effects are mimicked by sub-grid models. As presented in Chapter

3 and reviewed with detail in [7] , there is a large scatter around the amplitude of suppression in the power spectrum due to baryonic feedback modelled by subgrid recipes. Hence, in this work we not only rely on one hydrodynamical simulation, instead, we utilize BAryons and haloes of MAssive Systems (BAHAMAS, [178]) and IllustrisTNG [170] for weak lensing maps. Along with the fundamental differences between each other in terms of box sizes, cosmology, or mass resolution, each implementation of baryonic feedback considers specific calibration strategies that aim to model the processes accurately. Therefore, by using the two hydrodynamical simulations, we can explore and compare to what extent different baryonic process implementations and strengths can impact our cosmological constraints.

- **BAHAMAS:** The BAHAMAS suite is a 400 h^{-1}Mpc box size simulation at the WMAP 9-yr cosmology [179]. It provides modelling of stellar and AGN feedback, aiming to reproduce the observed present-day amplitude of gas fraction–halo mass relation of galaxy groups and clusters, and galaxy stellar mass function. The sub-grid prescription also incorporates star formation, stellar evolution, and radiative cooling that match a wide range of observable quantities. The box size is larger compared to other hydrodynamic simulations because it has been designed to study the effects of baryons on large scales for cosmology, suitable for the upcoming galaxy surveys analysis [180]. BAHAMAS also incorporates two additional large-volume simulation runs with the aim of varying the strength of AGN feedback. In one of them, the AGN heat temperature is increased by 0.2 dex, and in the other the AGN temperature is decreased by 0.2 dex. Hereafter, we refer to the +0.2 dex and -0.2 dex cases as high-AGN model and low-AGN model, respectively. The reference model from which the AGN heat temperature is shifted is referred to as fiducial-AGN model. The purpose of varying this temperature is to encompass the scatter observed in the gas fraction for galaxy groups in X-ray observations. We utilize weak lensing convergence maps for 60 source redshifts up to $z_s = 3$ based on BAHAMAS simulations. This set of maps includes 10.000 realizations per redshift and per model, each of them covering an area of 5×5 deg^2 of the sky and containing 340^2 pixels.
- **IllustrisTNG:** The TNG300-1 simulations are a set of cosmological and large-scale hydrodynamical simulations of box-size 205 h^{-1} Mpc [181, 182] at Planck 2016 cosmology $\Omega_\mathrm{b} = 0.0486$, $\Omega_\mathrm{m} = 0.3089$, $h = 0.6774$, $\sigma_8 = 0.8159$, and $n_s = 0.9667$ [183]. Sub-grid prescriptions are aimed at modeling black hole feedback, thermal and kinetic AGN feedback, stellar evolution, chemical evolution, galactic winds, and magnetic fields [184], among other physical processes. In this work, we use the already existing κTNG convergence maps presented in [170]. These maps cover 5×5 deg^2 of the sky and include 10.000 realizations for 40 source redshifts up to $z_s = 2.6$, obtained from random rotations, translation, and flips of the snapshots. Each map contains 1024^2 pixels, with a pixel size of 0.29 arcmin.

In total, we have 10.000 realizations of convergence maps at source redshifts based on four baryonic feedback scenarios: BAHAMAS low-AGN, fiducial-AGN, high-AGN, and κTNG. All of them possess gravity-only counterparts, dark-matter-only counterpart hereafter, which are based on the same corresponding simulation and initial conditions, but without baryonic physics. Based on these maps, we generate HSC Y1-like mocks, by following the photometric redshift distribution and galaxy number density in Table 5.1. The following methodology is

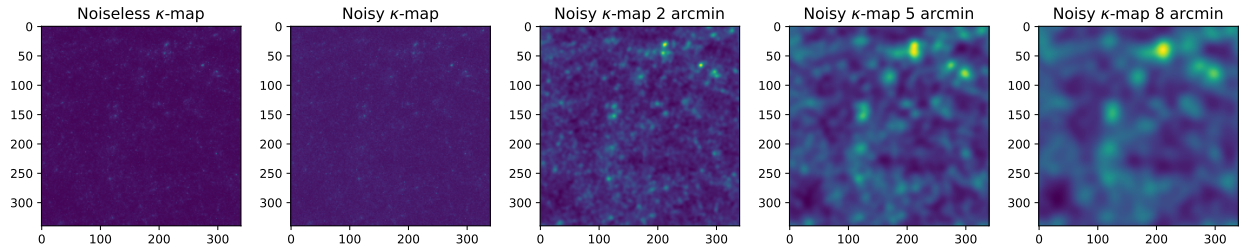


Figure 5.3: The result for the step-by-step mock production. *From left to right:* Original convergence map based on BAHAMAS high-AGN, HSC-Y1-like mock including shape noise, previous HSC-Y1 mock after applying 2 arcmin smoothing scale, the second HSC-Y1 mock after applying 5 arcmin smoothing scale and the same mock after applying 8 arcmin smoothing scale.

equally implemented for the maps based on hydrodynamical simulations, and in parallel to the dark-matter-only counterpart maps.

1. We first split the mocks into the tomographic bins, $0.3 < z_{\rm best} < 0.6$, $0.6 < z_{\rm best} < 0.9$ and $0.9 < z_{\rm best} < 1.2$. We perform a weighted sum to the maps for each tomographic bin, with the weight being based on the HSC Y1 redshift distribution shown in Fig. 5.1. This ensures that our maps are consistent with the observed distribution of galaxies in our data. In this manner, each tomographic bin is associated with the resulting map from the weighted sum.
2. In order to mimick the HSC-Y1 shape noise, we add Gaussian noise to each pixel, whose value is drawn from a Gaussian distribution centered at zero, with variance

$$\sigma^2 = \frac{\sigma_{\rm e}^2}{n_g^{\rm eff} A_{\rm pix}}, \tag{5.1}$$

   where $\sigma_{\rm e} \sim 0.28$ is mean intrinsic ellipticity of galaxies (shape noise), $n_g^{\rm eff}$ is the effective galaxy number density in Table 5.1, and $A_{\rm pix}$ is the solid angle of a pixel in units of arcmin^{-2}. The variance σ^2 depends on tomographic bin and its corresponding $n_g^{\rm eff}$.

3. In the next step, we smooth the maps. Smoothing is a standard procedure to reduce the shape noise per pixel, and to exploit the features encoded in the maps [185]. In our case, this method is extremely useful to investigate the scale at which the baryonic effects can be sufficiently mitigated. Therefore, we apply a Gaussian smoothing filter to the convergence mocks, using the Gaussian kernel $W_{\rm G}$

$$W_{\rm G}(\theta) = \frac{1}{\pi\theta_s^2} \exp\left\{\left(- \frac{\theta^2}{\theta_s^2}\right\}\right), \tag{5.2}$$

   implemented in `LensTools` package [186]. We consider the smoothing scales: $\theta_s = \{1, 2, 5, 8\}$ arcmin.

An example of the effect of our methodology for a map at $z_s = 0.8$ is depicted in Fig. 5.3. The first map corresponds to the original convergence map from BAHAMAS high-AGN, the second illustrates the result after applying noise given in Eq. (5.1). The last three maps are the result of applying a smoothing kernel Eq. (5.2).

5.3.3 HSC-Y1 real data

In order to understand the impacts of different analysis choices on the inferred S_8, particularly in configurations where baryonic effects are relevant, we complement our investigation using convergence maps based on the HSC first data release (Y1) [172, 26]. We utilize the same settings for our analysis of the HSC Y1 real data as those employed in the production of the mock data, including the tomographic redshift bins and smoothing the maps with a Gaussian kernel with the same smoothing scales.

5.4 Summary statistics

The use of non-Gaussian statistics, also known as higher-order statistics (HOS), has gained significant attention in constraining cosmological parameters. Through Fisher forecast analysis, it has been indicated that non-Gaussian statistics can provide improved statistical power in cosmological constraints compared to relying solely on the weak lensing power spectrum [150, 147, 21, 151, 159, 148, 152, 146]. Furthermore, the application of non-Gaussian statistics to real data has demonstrated their effectiveness in constraining cosmological parameters e.g. [148, 187]. These tools capture information beyond the linear regime and have become useful tools for studying the impact of astrophysical systematics and constraining parameters such as neutrino mass [153, 188, 147]. In this work, we focus on the convergence power spectrum, peak counts, minimum counts, and PDF to study the convergence field.

5.4.1 Power spectrum

The lensing signal is measured within a complex sky mask that accounts for the HSC disjoint patches and removes artifacts such as bright stars. Therefore, the estimation of the angular power spectrum can be biased due to the effect of the mask, leading to mode coupling between different scales. For this reason we compute the Pseudo-$C_\ell^{\kappa\kappa}$ power spectrum estimator [189]) on the maps using the public code `NaMaster` [190]. This approach mitigates the masking effect in the angular power spectrum estimation. We obtain the pseudo-power spectrum for each tomographic bin in 14 logarithmical and equally spaced angular multipoles bins spanning the range $80 < \ell < 6500$. To assess the influence of baryonic effects, we consider three different scale cut strategies: $300 < \ell < 900$, $300 < \ell < 1900$, and $900 < \ell < 1900$. We do not consider $\ell < 300$ as [191] found that there are unmodelled systematic errors. For each tomographic bin, we measure the power spectrum in each of the 10.000 realizations. We average over all power spectrum realizations to obtain $\langle C_\kappa^{\rm H}\rangle$ and $\langle C_\kappa^{\rm DM}\rangle$, the average power spectrum for maps containing baryonic physics, and for dark-matter-only maps, respectively (per baryonic feedback scenario and per tomographic bin, with only 1 arcmin smoothing scale considered).

5.4.2 Peaks and Minimum counts

Peak statistic has been studied in the literature extensively and hence it is positioned as one of the most popular non-Gaussian statistics to study the WL convergence field [159, 147]. We employ the definition of peaks which is the counting of pixels in the map whose (convergence κ) value is higher than their surrounding eight pixels. It has been shown that high peaks correspond to massive haloes in the line of sight, whereas low peaks trace the superposition of smaller haloes [192]. Conversely, minimum counts correspond to pixels

with lower values compared to their eight neighbors, tracing underdense regions and probing information complementary to the peak counts alone [21]. We measure the peaks in the linear signal-to-noise ratio $\kappa/\langle\sigma(\kappa)\rangle$ with 19 equally spaced bins from -4 to 4. It is important to note that the average $\langle\sigma(\kappa)\rangle$ refers to the average over the individual $\sigma(\kappa)$ obtained for each map realization. We remove the extreme bins and apply further scale cuts. For the peak counts, we utilize 12 bins in the range $-1.4 < \kappa/\langle\sigma(\kappa)\rangle < 3.6$, whereas for the minimum counts the same number of bins is obtained encompassing the values $-3.6 < \kappa/\langle\sigma(\kappa)\rangle < 1.4$. The measured peaks counts (or minimum counts) based on maps containing baryonic physics are denoted as $N^{\rm H}$, where H stands for 'hydro'. On the other hand, the measured peak counts (minimum counts) based on dark-matter-only maps are denoted as $N^{\rm DM}$. We average the statistics over the 10.000 realizations to obtain $\langle N^{\rm H}\rangle$ and $\langle N^{\rm DM}\rangle$.

5.4.3 Probability distribution function (PDF)

The probability distribution function is a statistical tool that captures a considerable amount of cosmological information. It is sensitive to the NG information contained in the fields and has been shown to contribute to tightening cosmological constraints for a single source redshift or in a tomographic analysis [151, 150]. In this work, the convergence PDF is built upon histogramming pixels of normalized convergence maps. As for peaks and minimum counts, the PDF is calculated in 19 equally spaced signal-to-noise (S/N) bins $\kappa/\langle\sigma(\kappa)\rangle$ from -4 to 4. We cut the first three bins and the 10th bin to prevent of covariance matrix instability, following the analysis methodology implemented in [160]. Each PDF data vector contains 15 bins in total. As for the peak counts and minimum counts, we obtain the average PDF $\langle N^{\rm H}\rangle$ and $\langle N^{\rm DM}\rangle$.

5.5 Parameter inference

We perform a Monte Carlo Markov Chain (MCMC) analysis for Bayesian inference of cosmological parameters Ω_m and S_8, adopting the flat priors $\Omega_m \in [0.15, 0.45]$ and $S_8 \in [0.5, 1]$ and implemented through the `Cobaya` inference framework [193, 194]. We adopt a Gaussian likelihood

$$\mathcal{L}(\mathbf{x}|\mathbf{p}) \propto \exp\left(-\frac{1}{2}[\mathbf{x}-\mathbf{p}]^T C^{-1}[\mathbf{x}-\mathbf{p}]\right), \tag{5.3}$$

where $\mathbf{x}$ is the data vector and $\mathbf{p}$ the theoretical prediction given a set of parameter values. The covariance matrix is denoted as C which is built based on the $N_r = 2668$ realisations of the fiducial model from full-sky N-body simulations with

$$C_{ij} = \frac{1}{N_r - 1}\sum_{n=1}^{N_r}(D_i^n - \hat{D})(D_j^n - \hat{D}), \tag{5.4}$$

where $\hat{D}$ is the mean data vector of the summary statistic. In order to obtain an unbiased inverse of the covariance matrix, we apply the Hartlap factor $(N_r - N_b - 1)/(N_r - 1)$ [195], where N_r is the number of realisations and N_b the number of bins. The theoretical predictions

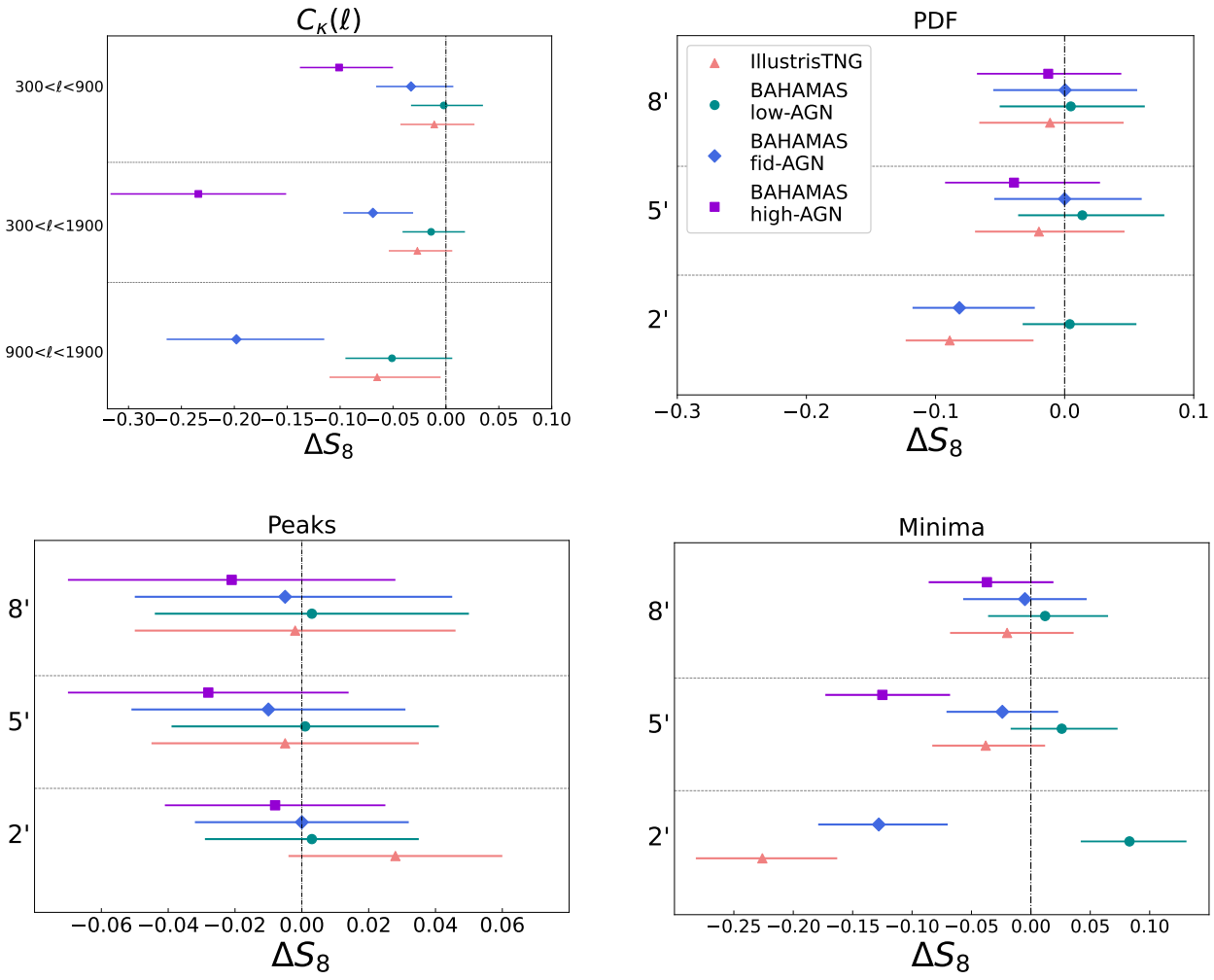


Figure 5.4: The impact of baryonic feedback on S_8 for the power spectrum and non-Gaussian statistics. We show the results for the hydrodynamical simulations BAHAMAS high-AGN (**purple squares**), fiducial-AGN (**blue diamonds**), low-AGN (**teal circles**) and κTNG (**pink triangles**). The vertical axis corresponds to the scales considered and the horizontal axis is the S_8 discrepancies defined as $\Delta S_8 = S_8^{\rm H} - S_8^{\rm DM}$ where H stands for the resulting S_8 from the contaminated synthetic data vector and DM for no contamination.

p are modeled using an emulator, which is trained on the N-body simulations presented in [176]. For the PDF, we utilize the public code presented in [160] to run the likelihood pipeline.

5.5.1 Emulator

In order to model and predict the peak counts, minimum, PDF, and power spectrum for arbitrary cosmologies, we build an emulator based on the Gaussian Process Regression (GPR) method with a Radial-Basis-Function (RBF) kernel implemented in scikit-learn[1]. An individual GPR is trained for each element of the data vectors, for each statistic, tomographic redshift bin and smoothing scale. In total, we train $3 \times 3 \times 12 = 108$ (smoothing scales $\times$ tomographic bin $\times$ S/N bins) emulators for the peak counts, $3 \times 3 \times 12 = 108$ for minimum

[1] `https://scikit-learn.org`

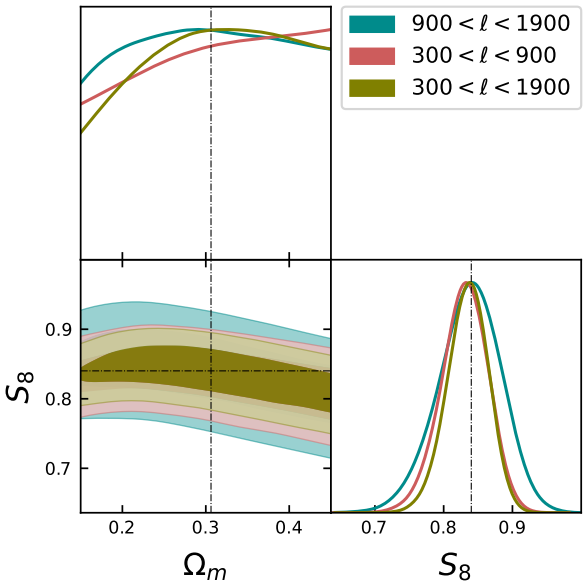


Figure 5.5: Validation of the pipeline, where the data vector is one of the models in the cosmo-varied simulations. The fiducial cosmology is depicted in dashed black lines.

counts and $3 \times 1 \times 14 = 42$ for the power spectrum. The training set consists of the summary statistics outputs from the 100 cosmological models from [176]. We assess the accuracy of the emulators by performing a leave-one-out cross-validation test, at which we compare the truth value with the emulator prediction. Our emulators present on average ~ 3 percent discrepancies between the prediction and truth for all statistics and smoothing scales. In order to further validate the emulator in the inference pipeline, we perform the inference of cosmological parameters in the case when the data vector is based on one of the models in the cosmo-varied mocks. The results for the power spectrum are shown in Fig. 5.5, where the contours are depicted in different colours to differentiate between the different analysis choices employed for the power spectrum (three different scale cuts). From Fig. 5.5, it is seen that we obtain unbiased results, recovering the truth cosmology depicted in dashed-black lines.

5.5.2 Data vector and baryon contamination

The method we employ to account for the presence of baryons proceeds as follows: following the methodology in Sec. 5.4.2 and Sec. 5.4.3, we measure the non-Gaussian statistics and power spectrum on convergence maps that include baryons (for each redshift bin and smoothing scale), and we repeat on the corresponding dark matter-only counterpart. The latter is equally applied to the four feedback scenarios: κTNG, BAHAMAS low-AGN, BAHAMAS fid-AGN and BAHAMAS high-AGN. We divide the statistics vectors $\langle N^H \rangle / \langle N^{DM} \rangle$ (or $\langle C_\kappa^{\rm H} \rangle / \langle C_\kappa^{\rm DM} \rangle$) to obtain the contamination factor which is infused to the data vector as a simple multiplication. Hence, the baryon-contaminated data vector in the likelihood pipeline is

$$\mathbf{x}^H = \frac{\langle N^H \rangle}{\langle N^{DM} \rangle} \mathbf{x}. \tag{5.5}$$

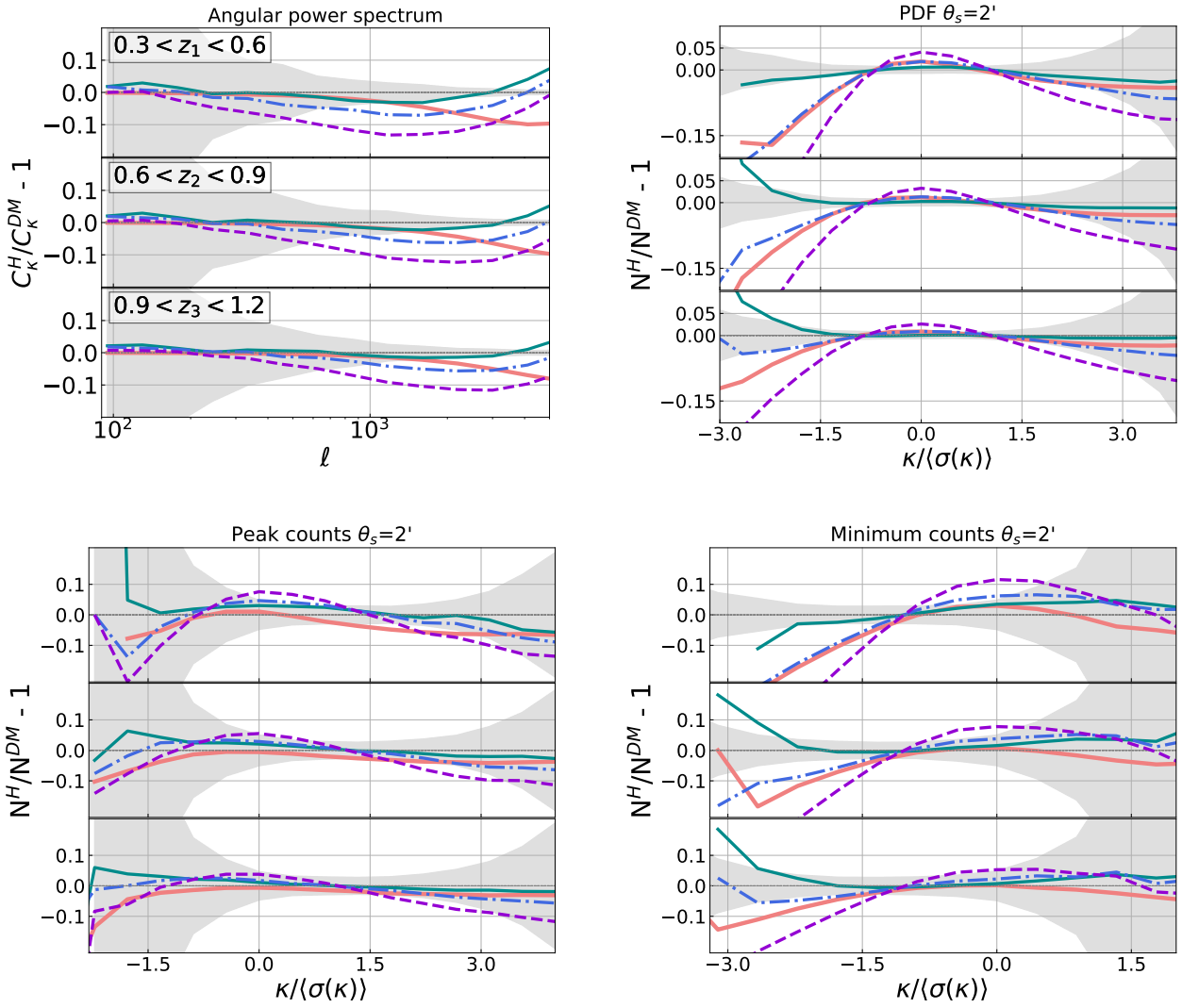


Figure 5.6: The impact of baryonic feedback on the convergence power spectrum (top left panel), PDF (top right panel), peak counts (bottom left panel), and minimum counts (bottom right panel) for κ-maps smoothed with a Gaussian smoothing kernel of $\theta_s = 2$ arcmin. We show the results for the hydrodynamical simulations BAHAMAS high-AGN (**dashed purple line**), fiducial-AGN (**dashed-dot blue line**), low-AGN (**solid teal line**) and κTNG (**solid pink line**). The vertical axis corresponds to the fractional difference where C^H corresponds to the baryon-contaminated angular power spectrum data vector, and C^{DM} the uncontaminated data vector (based on the dark matter-only runs of each simulation). The grey-shaded region corresponds to 1σ HSC-Y1 uncertainty.

We propagate the impact of baryons into the cosmological parameter inference as a two steps process: First, we sample the posterior distributions of the parameters using the likelihood based on dark matter-only data vectors $\mathbf{x}$. Then, we introduce the effect of baryons in the pipeline by replacing $\mathbf{x}$ for $\mathbf{x}^H$. This allows us to quantify the shift on Ω_m and S_8 due to baryonic physics under four feedback scenarios and scale cuts.

5.5.3 Data Compression

For the non-Gaussian statistics, we implement the MOPED (Massively Optimised Parameter Estimation and Data compression) data compression method to the data. This method reduces the size of the data vector to the number of parameters to be constrained (in this case, two parameters: Ω_m and S_8), while preserving the Fisher information (the Fisher matrix of the compressed dataset is identical to the Fisher matrix based on the original dataset) [95, 96]. The advantages are two-fold: it 'Gaussianizes' the likelihood as a consequence of the central limit theorem and makes the sampling of the posteriors more computationally efficient without losing information. The compressed data vector ($D^{\rm compr}$) is computed as follows

$$D^{\rm compr} = \frac{\partial D^{\rm T}}{\partial p_\alpha} C^{-1} D, \tag{5.6}$$

where $\frac{\partial D}{\partial p_\alpha}$ represents the partial derivative of the model data vector with respect to the α-th parameter. The power spectrum data vectors range from 2 to 6 bins, with its distribution known to be nearly Gaussian. Consequently, we apply the compression method only to the non-Gaussian statistics.

5.6 Results

- In Fig. 5.6, we show the impact of baryons on each statistic and tomographic bin for the mocks-based analysis with $\theta_s = 2$ arcmin smoothing scale. The vertical axis stands for the fractional difference $\Delta N = (N^{\rm H} - N^{\rm DM})/N^{\rm DM}$. As seen in Fig. 5.6, the four baryonic feedback prescriptions considered in this work, namely κTNG (solid pink line), low-AGN (solid teal line), fiducial-AGN (dashed-dot blue line) and high-AGN (dashed purple line), suppress the clustering at small scales. These prescriptions exhibit variations in amplitude and scale, consistent with previous studies [7].
  On the top left panel, we show the fractional difference for the **power spectrum**, where the BAHAMAS high-AGN reaches $\sim 10\%$ suppression for all tomographic bins at $\ell > 1000$, with a more pronounced effect observed in the first tomographic bin z_1. For the power spectrum, all the ratios exhibit a turn-around where baryons start to enhance the clustering, possibly attributed to gas cooling and star formation at the smaller scales. This turn-around appears earlier in ℓ when comparing all BAHAMAS cases to κTNG, also consistent with previous findings for other analysis choices [170]. In all redshift bins, the low-AGN is the least pronounced feedback, mainly due to the lower heating produced by the AGN-feedback mechanism. Moreover, low-AGN behaves in the same manner as TNG up to $\ell \sim 1000$.
  On the other hand, the **peak counts** (lower left panel) exhibit a suppression in the counts of peaks of approximately $\sim 10\%$ for high S/N, under BAHAMAS high-AGN. This implies that the presence of baryons dilutes the overdense regions, as observed indeed across all baryonic feedback implementations. However, these deviations from zero are relatively small in comparison to the 1σ uncertainties of the HSCY1 data. In the case of κTNG results, there is an overall suppression of the number of peaks for all S/N bins. In contrast, the BAHAMAS simulations exhibit a slight increase in the number of peaks for z_1 and z_2 at $-1.5 < \nu < 1.5$. One possibility is that the removed material from high S/N is placed in lower S/N regions. This is not seen in TNG as the effect in high S/N is mild. As seen in Fig. A.1 and A.2 in the Appendix **??**, the

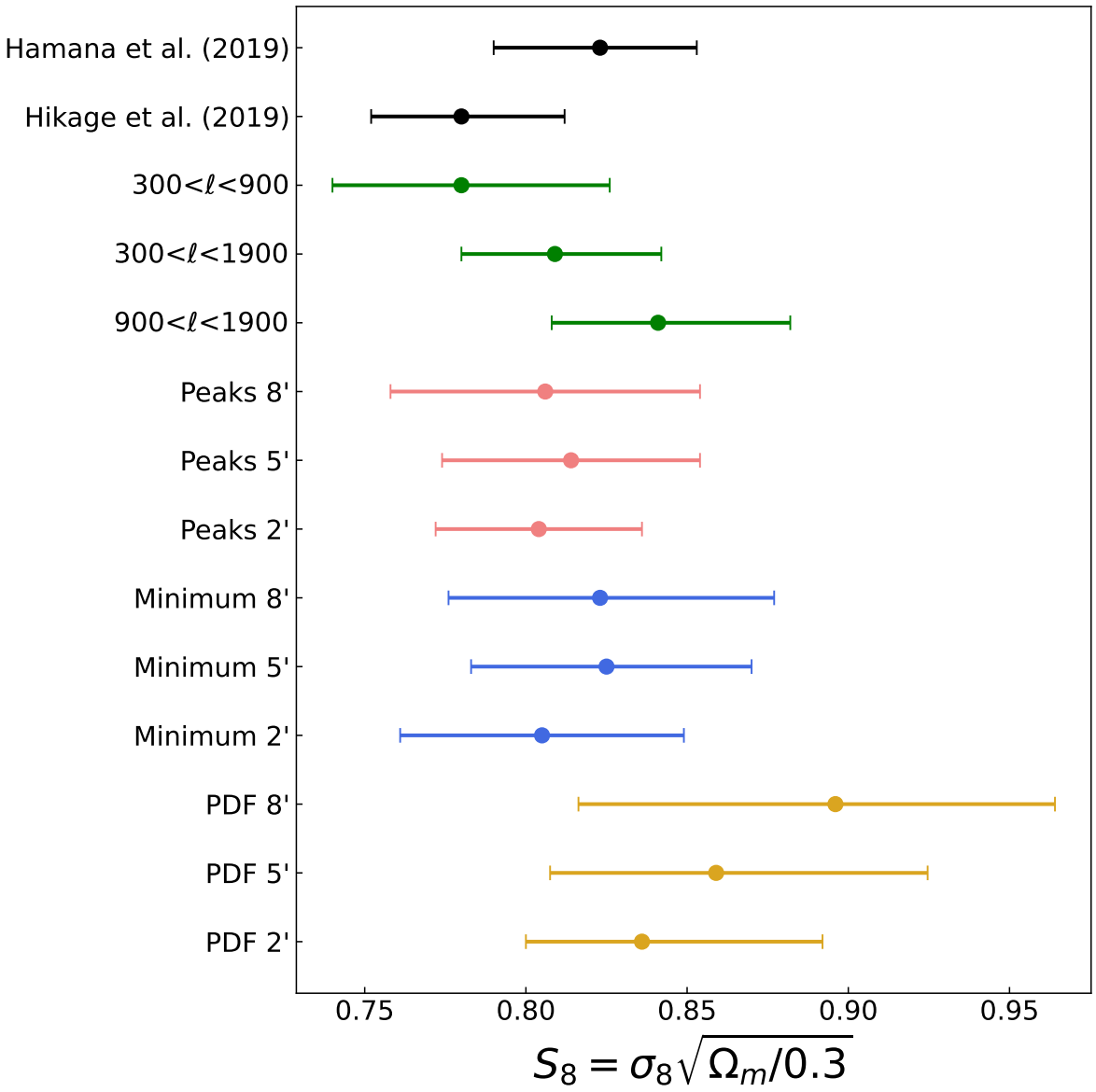


Figure 5.7: Results for HSC-Y1 real data analysis, where error bars show 68% credible regions. The horizontal axis corresponds to $S_8 \equiv \sigma_8\sqrt{\Omega_m/0.3}$ cosmological constraints for the power spectrum (**green**), peak counts (**pink**), minimum counts (**blue**), and PDF (**yellow**). For comparison, we show the results from previous two-point analysis for HSC-Y1 [9, 10] (**black**).

baryonic feedback becomes almost negligible for $\theta_s = \{5, 8\}$ arcmin.

We also found that baryons produce an important impact on the **PDF** and **minimum counts**, resulting in deviations beyond the 1σ error of the HSC-Y1 data. Since both NG statistics capture information from underdense regions, the impact of baryons in Fig. 5.6 shows consistent results between each other. The PDF and minimum counts show the suppression due to baryons in negative S/N, while the PDF also traces the peak counts information accordingly. In that sense, PDF is impacted in both tails, with a $> 15\%$ effect in the left tail, and $\sim 10\%$ in the right tail. Moreover, the κTNG follows the low-AGN trend for the positive S/N region, whereas for negative S/N it closely follows the fiducial-AGN tendency at z_1. In the low-AGN case, we observe an

	Scale cuts		
	$300 < \ell < 900$	$300 < \ell < 1900$	$900 < \ell < 1900$
$C_\ell^{\kappa\kappa}$	$0.780^{+0.042}_{-0.033}$	$0.809^{+0.032}_{-0.028}$	$0.841^{+0.046}_{-0.040}$
	Smoothing scales		
	8'	5'	2'
Peak	0.806 ± 0.048	0.814 ± 0.04	0.804 ± 0.032
Minimum	$0.823^{+0.054}_{-0.047}$	$0.825^{+0.046}_{-0.042}$	0.805 ± 0.044
PDF	$0.896^{+0.0681}_{-0.0797}$	$0.859^{+0.0656}_{-0.0515}$	$0.836^{+0.056}_{-0.036}$

Table 5.2: Cosmological constraints for HSC-Y1 real data. For the power spectrum, we consider three scale cuts: $300 < \ell < 900$, $300 < \ell < 1900$, and $900 < \ell < 1900$. The second row refers to the smoothing scales θ_s considered for computing the non-Gaussian statistics. For the power spectrum, we choose $\theta_s = 1$ arcmin. We report 68% credible level.

interesting trend in the minimum counts, where the inclusion of baryons leads to an increase in the number of minimum counts for the negative signal-to-noise (S/N) at z_2 and z_3, which stands in complete contrast to the results obtained in the high-AGN scenario. The discrepancy can be explained by the characteristics of the low-AGN model, which is characterized by weaker feedback strength. Consequently, underdense regions can form and are not diluted as much as in the high-AGN case, where stronger energy injection occurs.

Effect on S_8: In Fig. 5.4, we observe how the effects described above result in changes in S_8, represented by the difference between the baryon-contaminated case ($S_8^{\rm H}$) and the dark matter-only data vector ($S_8^{\rm DM}$). The error bars of the data points in the figure correspond to the 68% credible region, while the vertical axis displays the various analysis choices under consideration.
When using the power spectrum on small scales ($900 < \ell < 1900$), displayed in the top-left panel, we observe that the impact of baryons leads to a shift of $\sim -2.5\sigma$ for the fiducial-AGN scenario and $\sim -1.1\sigma$ for TNG. However, for the low-AGN model, the result is statistically consistent, with a shift of $< 1\sigma$ compared to the S_8 constraints from dark matter-only simulations. It is important to mention that the high-AGN case demonstrates a significant effect at these scales, leading to constraints that could not be obtained, as it falls outside the prior limits covered by the simulations. As baryonic effects occur mainly at small scales, it is expected to see a negligible bias for large scales ($300 < \ell < 900$), except for the high-AGN model, which exhibits a shift of $\sim -2\sigma$. When considering the whole scales range ($300 < \ell < 1900$), the shifts are -2.8σ and -1.8σ for high-AGN and fiducial-AGN models, respectively. Hence, even for conservative scale cuts, the high-AGN feedback mechanism bias the cosmological constraints from the power spectrum.

The fractional difference results obtained from Fig. 5.6 translate into a direct impact on S_8 for both the minimum and the PDF shown in the top and bottom right panels

of Fig. 5.4, respectively. At large scales, the ratio is close to unity, i.e., no significant baryonic effects. Similarly as observed for the power spectrum on small scales, the PDF and minimum counts under the influence of the high-AGN model introduces substantial shifts in the S_8 constraints for maps smoothed with a Gaussian kernel of $\theta_s = 2$ arcmin. These shifts exceed the prior range covered by our simulation set, which is why we do not have S_8 constraints for these statistics at this particular smoothing scale and baryonic model. The other models, fiducial-AGN, lowAGN and κTNG, also introduced biases, with shifts of -2.2σ, 2σ and -3.6σ, respectively, for minimum counts at 2 arcmin. The low-AGN exhibits a positive shift S_8 value for minimum, mainly due to the enhancement seen in 5.6 in the negative S/N tail, however, this is not captured in the PDF results. For PDF and minimum counts analyses of maps with larger smoothing scales, $\theta_s = 5$ and 8 arcmin, we do not observe significant ($< 1\sigma$) shifts on S_8 caused by the inclusion of baryonic effects, except for the high-AGN model, with -2.2σ bias.
The S_8 constraints from peak counts are less affected by baryonic physics, with no significant shifts ($< 1\sigma$) for all the feedback mechanisms.

As a final remark for the mock-based analysis, it is interesting to notice the direction of the shifts caused by the baryonic mechanisms for the different summary statistics. In the case of the power spectrum, all feedback mechanisms tend to shift S_8 towards lower values. However, the behavior is not necessarily the same for the NG statistics (pdf, peaks, and minimum counts). This variability indicates that baryonic effects might have a more nuanced influence on the NG statistics, and the direction of the shifts in S_8 can depend on the specific feedback scenario and analysis setup.

- We repeat the parameter constraints using summary statistics computed from the HSC Y1 real data rather than a synthetic data vector contaminated by baryonic effects. We aim to investigate whether we observe any significant shifts in the cosmological parameters derived from real data when employing the same analysis choices as for the convergence mocks based on simulations. We show the results for the HSC Y1 real data analysis in Fig. 5.7 and Table 5.2. The real data results show no statistically significant shift in S_8 when comparing large and small scales. In Fig. 5.7, we also show the results from previous works that use the 2-point information of the HSC-Y1 data $S_8 = 0.780^{+0.030}_{-0.033}$, [9], and $S_8 = 0.823^{+0.032}_{-0.028}$, [10]. In all scenarios, we find good agreement with these previous findings, although our analysis differs from these studies in many aspects, including the tomographic bins, scale cuts, the model of the matter power spectrum, and the parameters considered during the inference process. The power spectrum result for C_ℓ $900 < \ell < 1900$ exhibits a 1σ bias towards higher S_8, whereas minimum 2 arcmin shows a 0.3σ bias to lower S_8. The larger S_8 for power spectrum small scales is consistent with other HSC results [196] based on the two-point correlation function.

  Finally, from Table 5.2 we observe that the constraints from peak counts at 2 arcmin outperform all statistics, and hence a combined analysis of two-point information with the peak counts can provide tighter constraints. In our companion paper [26], and in [160], we further explore the constraining power of the NG statistics compared to the power spectrum alone, using smoothing scales that are not expected to be impacted by baryons as given by the hydrodynamical simulations among other systematic effects.

5.7 Discussion

In this study, we investigate the impact of baryons on the convergence power spectrum and non-Gaussian statistics peak counts, minimum counts, and the PDF for lensing convergence maps based on HSC-Year1 mocks. To quantify the effect of baryons on the inferred amplitude of matter clustering S_8, we consider different analysis choices, such as varying smoothing scales, applying scale cuts, and implementing different baryonic feedback models. We employ cosmological hydrodynamical simulations κTNG and BAHAMAS low-AGN, fiducial-AGN, and high-AGN to measure the fractional difference between the statistics obtained from hydrodynamical simulations-based data vectors and dark matter-only data vectors. These comparisons help us understand the impact of baryons on the cosmological constraints and how it varies based on the different choices considered in our analysis.

The impact of baryons goes up to 10% for the power spectrum and the peak counts and reaches $> 15\%$ impact on the negative S/N tails of the PDF and minimum counts. Our results show that the BAHAMAS high-AGN feedback model produces the strongest impact of baryonic effects on the inferred S_8 when considering small scales. This effect is stronger for the power spectrum, PDF and minimum counts. However, most of them show no significant bias for scales larger than 5 arcmin in our mock-based analysis. If nature behaves similarly to the high-AGN model, our analysis indicates that significant cuts are required in all the statistics considered in this study in order to obtain unbiased results. This observation aligns with the expected behavior based on the fractional differences observed in Fig.5.6. However, most of them show no significant bias for scales larger than 5 arcmin in our mock-based analysis. In general, the peak counts is less sensitive to baryons, contrary to what [21] found (although for a different survey and analysis choice).

Interestingly, even though most of the baryonic mechanisms shift S_8 towards lower values due to the suppression of the structures on small scales, we found that the NG statistics can exhibit a different trend depending on the specific setup of the analysis (Fig. 5.4). This highlights the importance of exploring different summary statistics to perform cosmological analysis. By considering various statistical measures, we can gain a more comprehensive understanding of how baryonic physics impacts cosmological parameter estimation. Additionally, non-Gaussian statistics can improve the constraining power on S_8 in a combined analysis, as we show in our companion paper [26].

We run the parameter constraints using HSC-Y1 real data as data vector, under the different smoothing scales and scale cuts we considered for the simulations. For real data, we do not find any indication of significant shifts in S_8, as observed in the simulations when small scales are included Fig. 5.7. Hence, our results suggest that weak baryonic feedback scenarios, such as BAHAMAS low-AGN, are more likely to reproduce HSC-Y1 real data systematics at small scales. Our results are consistent with other Stage-III analyses based on the two-point information, also showing that weak baryonic feedback scenarios are preferred for other Stage-III surveys [197].

It is essential to consider the complexity and limitations of both the simulations and the real data when interpreting these results. In real data, other sources of bias may arise from unaccounted, scale-dependent effects. For instance, the intrinsic alignment of galaxies can introduce correlations between the galaxy shapes that are scale-dependent and may impact cosmic shear measurements [198]. Neutrino masses can also introduce scale-dependent effects on the matter power spectrum, influencing the overall cosmological constraints. Given the multitude of potential scale-dependent effects, it is crucial to carefully account for and model

all relevant astrophysical and cosmological factors when conducting cosmological analyses. Our main results may also indicate that other factors or systematic uncertainties might play a more significant role in shaping the cosmological constraints obtained from the real observational data. We expect future analysis with improved constraining power will be able to discriminate more precisely between baryonic feedback prescriptions and determine if the lower amplitude in S_8 found in some LSS studies can be attributed to baryonic feedback systematics. The NG statistics are particularly valuable for this task, as they provide crucial insights into the impact of different feedback models on cosmological constraints.
Future Stage-IV surveys will provide constraints on cosmological parameters with lower statistical errors, and thus the biases produced by unmodelled baryonic physics will be more statistically significant than those found in this work. For such surveys, scale cuts based on contaminated-data vectors with weak baryonic feedback strength need to be revisited.

Chapter 6

Parametric and non-parametric reconstruction of dark energy evolution

6.1 Overview

Since the discovery in 1998 [13, 14] of cosmic acceleration, the scientific community has incorporated the existence of an exotic component in the Universe, which in its simplest form is known as a cosmological constant of a very specific magnitude. Beyond the cosmological constant hypothesis, other candidates have emerged as an explanation to the observed accelerated expansion: a new component filling the universe, such as a quintessence field, modifications of the theory of gravity, or inhomogeneous models that do not fulfill the cosmological principle [70]. Because of our ignorance on the nature of dark energy (DE), it becomes important to consider and test alternatives to the standard cosmological constant. In this context, many studies have followed a phenomenological approach. For instance, the Chevalier-Polarski-Linder [199] parametrization [69] extends the conventional first order expansion of the dark energy equation of state (EoS) to the parametric redshift-dependent form $w(z) = w_0 + w_a z/(1+z)$. This model free parameters w_0 and w_a has been widely tested and included in the forecast of cosmological parameters of future galaxy surveys [18, 15]. Another approach is to take the opposite direction: using the data to reconstruct the form of cosmological functions, such as $w(z)$. However, as was demonstrated in [200], the $w(z)$ function is behind two integrals from the data we use to reconstruct it, and hence any dynamical information presented in $w(z)$ would be certainly erased after the analysis. For this reason, in [201] a reconstruction was made using $H(z)$ and the growth function, whereas the authors of [202] propose a reconstruction of the scalar field potential assuming that a quintessence field is the nature of DE. In this context, reconstruction schemes using the DE density $\rho_{\rm de}$ were proposed by [203] and [204].

In this work, we use the dark energy density $X(z) = \rho_{\rm de}/\rho^0_{\rm de}$ as a probe to detect departures from the concordance ΛCDM model. With this methodology, we can study whether an evolving dark energy (DE) component, whose nature remains yet to be determined, emerge from cosmological data. Our methodology follows two theory-agnostic approaches: a parametric and non-parametric statistical analysis of late-time cosmological observations.

First, we consider a reconstruction scheme where the dark energy density is assumed under

two separate parametric models: a quadratic and a cubic parametrization. And second, we consider a data-driven approach where no functional form for DE density is assumed. The latter approach has been widely used in cosmology [205, 206, 207] because it makes no assumptions about the underlying cosmological model. Moreover, in order to make the results reflective of any possible influence from the Hubble tension, we employ four priors on H_0. In this chapter, we first present the parametric methods proposed in our work Grandón et al. [28], where the first test of the cubic parametrization against Type-Ia supernova and CC data was provided. Later, we introduce the non-parametric approach, based on our paper [29]. To assess the performance of the methods, we introduce a few metrics and compare our results based on cosmic chronometers data, and Type-Ia supernovae data.

6.2 Parametric method

A general way to test whether a cosmological constant Λ or a dynamical fluid drives the expansion consists in considering an arbitrary function $X(z)$ describing the dark energy fluid. To the best of our knowledge, this was first proposed in [204] assuming a linear interpolation between redshifts, and later tested for a quadratic interpolation in [208, 209]. The results of all these first explorations were that the DE density showed a slight increase with redshift, being consistent with Λ at 2σ. With more recent data, the problem was revisited in [210] where a quadratic interpolation was implemented and tested for data from SN Ia, gas mass fraction in galaxy clusters, baryonic acoustic oscillations, and the cosmic microwave background. Surprisingly, the trend obtained was contrary to the previous one, indicating a DE density that decreases with redshift towards negative values for $z > 1.5$ at 1σ (which is consistent with other studies [211, 212]). The extension of this work in light of more recent data is presented in our work [28], where evidence for DE evolution using a quadratic and also a cubic interpolation was studied using the latest SN Ia and cosmic chronometeres data.
The evolution of the dark energy sector enters through the spatially flat Friedmann equation with $E(z) = H(z)/H_0$

$$E(z)^2 = \Omega_{m0}(1+z)^3 + (1-\Omega_{m0})X(z)\,, \tag{6.1}$$

where $X(z)$ is the normalized DE density. In what follows, we aim to explore two parametrizations techniques presented in [28].

6.2.1 Quadratic parametrization

We use first a quadratic interpolation for $X(z)$ given by

$$X(z) = \frac{x_0(z-z_1)(z-z_2)}{(z_0-z_1)(z_0-z_2)} + \frac{x_1(z-z_0)(z-z_2)}{(z_1-z_0)(z_1-z_2)} + \frac{x_2(z-z_0)(z-z_1)}{(z_2-z_0)(z_2-z_1)}, \tag{6.2}$$

where x_0, x_1 and x_2 are the constant values of $X(z)$ evaluated at z_0, z_1 and z_2 and we assume that $z_2 > z_1 > z_0$. We set $X(z = z_0 = 0) = 1$ and define z_2 as the largest redshift in the data z_m that allows us to properly estimate the parameters, and the intermediate redshift value $z_1 = z_m/2$. With these definitions, the free parameters are: $x_1 = X(z_m/2)$ and $x_2 = X(z_m)$, and (6.2) reduces to

$$X(z) = 1 + \frac{z(4x_1 - x_2 - 3)}{z_m} - \frac{2z^2(2x_1 - x_2 - 1)}{z_m^2}. \tag{6.3}$$

When all the free parameters are $x_i = 1$, the dark energy density gives $X(z) = 1$, recovering the ΛCDM limit.

6.2.2 Cubic parametrization

Knowing the limitations of a quadratic interpolation – the parametric function can only detect a single trend, such as a parabolic increase(decrease) with the evolution – in this section we extend to a cubic parametrization. The cubic parametrization corresponds to

$$\begin{aligned} X(z) = &\frac{x_0(z-z_1)(z-z_2)(z-z_3)}{(z_0-z_1)(z_0-z_2)(z_0-z_3)} + \\ &+\frac{x_1(z-z_0)(z-z_2)(z-z_3)}{(z_1-z_0)(z_1-z_2)(z_1-z_3)} + \\ &+\frac{x_2(z-z_0)(z-z_1)(z-z_3)}{(z_2-z_0)(z_2-z_1)(z_2-z_3)} + \\ &+\frac{x_3(z-z_0)(z-z_1)(z-z_2)}{(z_3-z_0)(z_3-z_1)(z_3-z_2)}. \end{aligned} \tag{6.4}$$

As in the quadratic form, we assume $z_0 < z_1 < z_2 < z_3$ and set $z_0 = 0$ in such a way that $x_0 = X(z = z_0 = 0) = 1$. We set $z_3 = z_m$ as the largest redshift in the data sets, and define $\{z_2 = 2z_m/3, z_1 = z_m/3\}$. In this model, the free parameters are $x_1 = X(z_1)$, $x_2 = X(z_2)$ and $x_3 = X(z_3)$, and Eq. (6.4) reduces to

$$\begin{aligned} X(z) = 1 &+ 9(-1+3x_1-3x_2+x_3)\frac{z^3}{2z_m^3} + \\ &-9(-2+5x_1-4x_2+x_3)\frac{z^2}{2z_m^2} + \\ &+(-11+18x_1-9x_2+2x_3)\frac{z}{2z_m} \end{aligned} \tag{6.5}$$

As before, the ΛCDM limit corresponds to $x_1 = x_2 = x_3 = 1$.

6.3 Non-parametric method

Non-parametric methods offer a physics-agnostic alternative to reconstruct the cosmological functions, providing credibility intervals for the reconstructed mean. In this study, we consider a Gaussian Process (GP) reconstruction method, which is based purely on how data points are statistical correlated based on a kernel function. The GP has emerged as a popular machine learning tool in cosmology, as for the reconstruction of physical functions or for emulating cosmological observables any model [205, 206, 207]. The training process consists in iteratively tuning the kernel hyperparameters (non-physical parameters) by maximizing the likelihood function $\ln\mathcal{L}$.

Consider N observations of $H(z)$ with a covariance matrix C of size $N \times N$. The GP kernel is defined as $K(z^*, \tilde{z}^*)$, which relates the function values at coordinates z^* and $\tilde{z}^*$, with $\tilde{z}^* \neq z^*$. Then, the mean and covariance of the GP reconstruction of $H(z)$ are given by

$$\langle H^* \rangle = K(z^*, Z)\left[K(Z,Z) + C\right]^{-1} H(Z) \ , \tag{6.6}$$

and

$$\mathrm{cov}\left(H^*\right) = K\left(z^*, z^*\right) - K\left(z^*, Z\right)\left[K\left(Z, Z\right) + C\right]^{-1} K\left(Z, z^*\right) \,, \tag{6.7}$$

where Z stands the redshifts in the observation. The kernel hyperparameters θ are then obtained by maximizing

$$\ln \mathcal{L} = -\frac{1}{2} H\left(Z\right)^T \left[K\left(Z, Z\right) + C\right]^{-1} H\left(Z\right) - \frac{1}{2} \ln \left|K\left(Z, Z\right) + C\right| - \frac{N}{2} \ln\left(2\pi\right) \,. \tag{6.8}$$

given the input data $H(Z)$. Eqs. (6.6–6.8) therefore exhibits the simplicity of the GP method and why it is constantly pursued in cosmology.
To assess the performance of each method, we consider three metrics that can be defined for parametric and non-parametric methods. First, we study the χ^2 test defined as

$$\chi^2 = \sum_z \left(\frac{H_{\rm rec}(z) - H_{\rm obs}(z)}{\sigma_{\rm obs}(z)} \right)^2 \,, \tag{6.9}$$

where $H_{\rm rec}(z)$ is the reconstructed Hubble function while $H_{\rm obs}(z)$ and $\sigma_{\rm obs}(z)$ are the mean and 1σ uncertainty of the data. This statistic measures the discrepancy between the reconstructed H(z) and data in units of the uncertainty of the data, and it has the advantage that can be defined for both parametric and non-parametric approaches. The good fit corresponds to χ^2 close to the size of the data N, while overfitting in a parametric sense corresponds to $\chi^2 < N$. We further consider two statistical measures which have been used previously to compare nonparametric reconstruction methods [213]. These are given by

$$\mathcal{D} = \sum_z \left(\frac{H_{\rm rec}(z) - H_{\rm obs}(z)}{\sqrt{\sigma_{\rm rec}(z)^2 + \sigma_{\rm obs}(z)^2}} \right) \,, \tag{6.10}$$

and

$$\gamma^2 = \sum_z \left(\frac{H_{\rm rec}(z) - H_{\rm obs}(z)}{\sqrt{\sigma_{\rm rec}(z)^2 + \sigma_{\rm obs}(z)^2}} \right)^2 \,, \tag{6.11}$$

where $\sigma_{\rm rec}(z)$ is the uncertainty in the reconstruction. The main difference between χ^2 and the statistics $\mathcal{D}$ and γ^2 is how they treat the uncertainties in the reconstruction. The $\mathcal{D}$ and γ^2 metrics consider the uncertainty in a reconstruction on an equal footing with the uncertainty in the data, but $\mathcal{D}$ can be positive or negative (depending on whether the data points lie mostly above/below the best fit) while γ^2 is strictly positive.

6.4 Data

6.4.1 $H(z)$ measurements

We consider measurements of the Hubble function $H(z)$ from cosmic chronometers (CC) and baryon acoustic oscillations (BAO). The CC data set includes 31 data points [214, 215, 216, 217, 218, 219] obtained through the differential aging method described in Chapter 1. The second set of Hubble data consists of 26 points derived from BAO measurements, a characteristic scale in the distribution of galaxies with its origin in the fluctuations of the baryon density in the early Universe. The BAO directly measures the combination $H(z)r_{\rm d}$

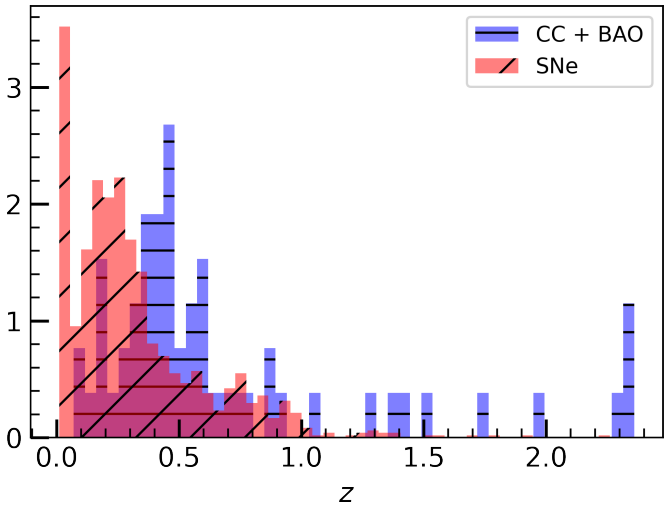


Figure 6.1: Redshift distributions of the base Hubble data coming from CC and BAO (blue bars), and SN-Ia observations (red bars).

where $r_{\rm d}$ is the radius of the sound horizon during baryon drag. This additionally relies on the ΛCDM model which sets the scale of the sound horizon ($r_{\rm d} = 147.74$ Mpc) during baryon drag and supplements more points to the Hubble data [220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231]. The compiled Hubble data from CC and BAO is presented in Table B.1.

6.4.2 Type Ia Supernovae

We make use of 1048 supernovae (SNe) type Ia observations from the Pantheon sample [232]. These measurements span the redshift range $0.01 < z < 2.3$. We utilize the SNe measurements of $E(z)$ of the CANDELS and CLASH Multi-Cycle Treasury data (MCT) [233]. This compression of the Pantheon data in terms of the normalized expansion function is utilized for the GP which together with an H_0 prior can be used to reconstruct $H(z) = H_0 E(z)$ [234, 235]. When considering the SNe observations, we marginalize analytically over the SNe absolute magnitude, as detailed in Ref. [236].
The redshift distribution of the Hubble data and SNIa observations employed in this work is shown in Figure 6.1.

6.5 Priors

In order to sample parameter posterior distribution based on the parametric models, we make use of a Gaussian likelihood in a separate analysis for Hubble data and SNIa data.
To make the results reflective of any possible influence from the Hubble tension, we perform parameter inference in four scenarios, each of them considering a different Gaussian prior on H_0. The priors on H_0 corresponds to

- $H_0^{\rm P18} = 67.4 \pm 0.5$ km s^{-1}Mpc^{-1} [165]
- $H_0^{\rm TRGB} = 69.8 \pm 1.9$ km s^{-1}Mpc^{-1} [237]

- $H_0^{\rm A21} = 71.5 \pm 1.8$ km s^{-1}Mpc^{-1} [238]
- $H_0^{\rm R21} = 73.04 \pm 1.04$ km s^{-1}Mpc^{-1} [239].

These H_0 values represent the current Hubble tension and are considered in this analysis to shed more light on this intriguing puzzle [74]. We emphasize that the usage of parameter priors such as those on H_0 compromise the notion of "model-independence" depending on which assumptions were considered to obtain the priors in the first place. On the other hand, the local H_0 values differ in terms of how the supernovae were calibrated. The derived $H_0^{\rm R21}$ comes from the SNIa calibration using cepheids, while the derived $H_0^{\rm TRGB}$ comes from calibration based on the tip of the red giant branch. $H_0^{\rm A21}$ is a recent reanalysis of the TRGB which lead to a slightly different value of H_0. In summary, taking into account these different measurements of H_0 may provide valuable insight into how the Hubble tension may influence the estimates of cosmological parameters. Finally, we consider the matter fraction prior $\Omega_{m0}h^2 = 0.1430 \pm 0.0011$ [165] which is measured from the cosmic microwave background observations using the angular peaks in the damping tail. Therefore, even though the methods are completely different in nature, they stand on at least a common ground in this analysis, which is the Planck prior on combination of the matter density and the present expansion rate. This is particularly needed for the GP since unlike parametric methods, the GP does not estimate parameters outside of the information it is provided, but rather it reconstructs a particular data set which it is given. The sampling of posteriors is performed using the public code `emcee` [168]. This is a stable, well tested Python implementation of the affine-invariant ensemble sampler for Markov chain Monte Carlo (MCMC) proposed by Goodman & Weare [240]. The output from the chains are visualized using GetDist [169].

6.6 Results

We first show our results for the posterior distribution of the quadratic model parameters based on the $H(z)$ data. These results are shown in Fig. 6.2. From this corner plot, it can be seen that the estimated parameters are influenced by the choice of H_0 prior. Clearly, this is the effect of the Gaussian prior over the likelihood, which directly impacts the parameter posteriors. The most notable feature of Fig. 6.2 is that the posteriors of x_1 and x_2 continue to deviate further from ΛCDM ($x_1 = x_2 = 1$) for increasing values of H_0. The reason for this trend might be due to the Planck $\Omega_{m0}h^2$ prior. However, even for the Planck prior, the inferred quadratic model parameters are $x_1 = 0.8 \pm 0.1$ and $x_2 = -0.6 \pm 0.6$, which disfavours ΛCDM ($x_1 = x_2 = 1$) at 2σ for x_2. Also, Figure 6.2 shows the effect of $\Omega_{m0}h^2$ prior, where the value inferred $\Omega_{m0}h^2$ is fairly constant irrespective of the value of priors on H_0. On the other hand, x_1 and x_2 seem to be correlated giving lower(higher) values with respectively higher(lower) H_0 priors. Furthermore, it can be seen that the tension in H_0 induces a tension in the matter density Ω_{m0} when $\Omega_{m0}h^2$ is practically fixed given the cosmic microwave background data (also, consdering that this parameter is well measured from CMB power spectrum). Figure 6.3 shows the results for the cubic parametrization. Again, we see the impact of the H_0 priors on the model parameters. The increasing trend of deviation from the ΛCDM model can also be seen in this case, i.e., the deviation of the posteriors of x_1, x_2, and x_3 away from $x_1 = x_2 = x_3 = 1$ increases in the order P18, TRGB, A21, and R21, which is more significant in the x_1 parameter. As in the quadratic case, this can be traced to the use of the Planck prior on the combination $\Omega_{m0}h^2$ of the matter fraction and the Hubble

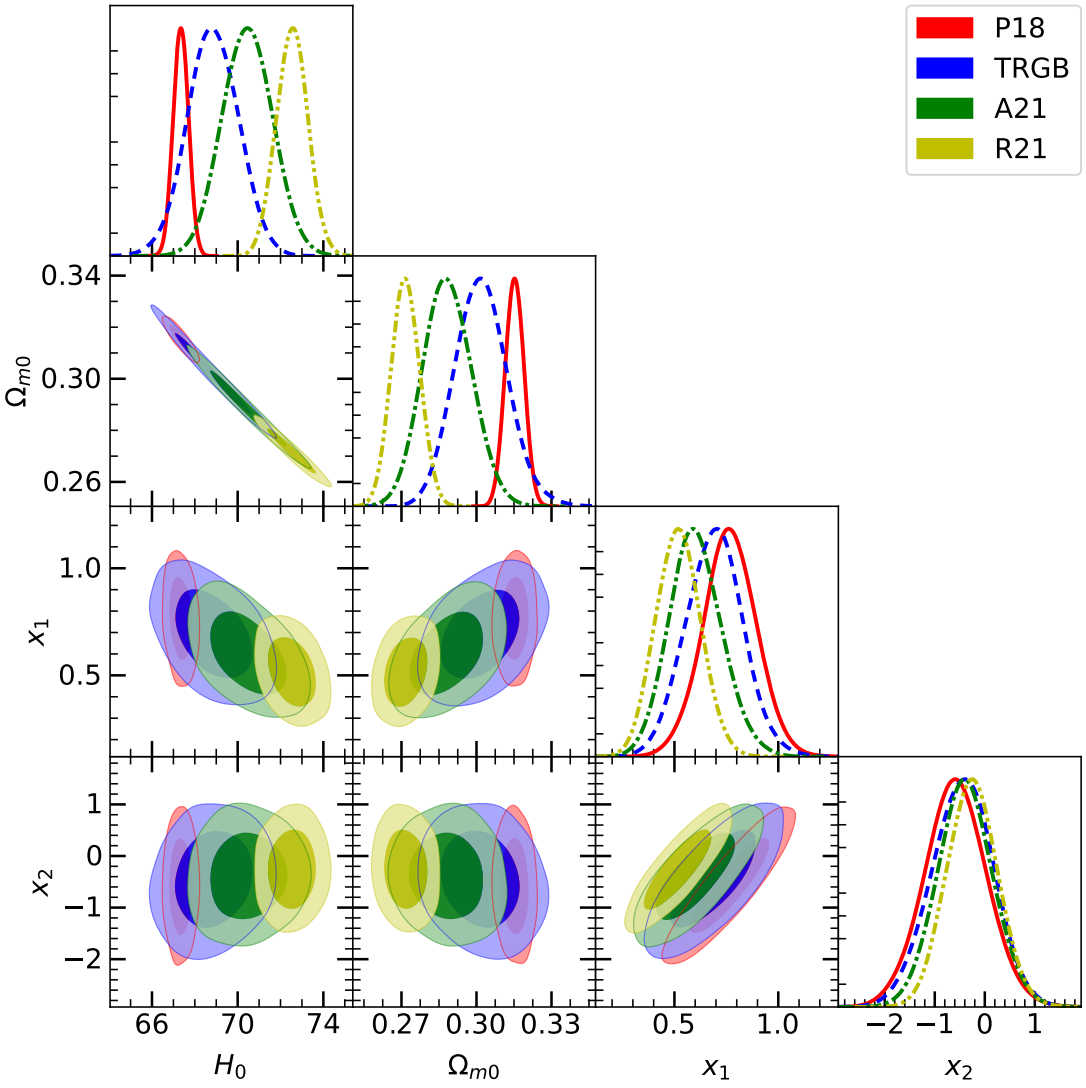


Figure 6.2: The sampled posteriors of the parameters $(H_0, \Omega_{m0}, x_1, x_2)$ in quadratic parametrized DE for each H_0 prior: $H_0^{\rm P18} = 67.4 \pm 0.5$ km s^{-1}Mpc^{-1}, $H_0^{\rm TRGB} = 69.8 \pm 1.9$ km s^{-1}Mpc^{-1}, and $H_0^{\rm A21} = 71.5 \pm 1.8$ km s^{-1}Mpc^{-1}, and $H_0^{\rm R21} = 73.04 \pm 1.04$ km s^{-1}Mpc^{-1}. These were obtained with the base Hubble data (CC + BAO).

constant. Again, a deviation from ΛCDM is also present for the cubic parametrization even for $H_0^{\rm P18}$, particularly with the marginalized posterior of $x_3 = -0.9 \pm 0.5$ excluding $x_3 = 1$ at more than 2σ. It is worth to note that the Hubble tension is practically only influencing the matter density and *not* some of the dark energy parameters, such as x_2 and x_3 in the cubic case. This means that we may draw conclusions on dark energy evolution regardless of the Hubble tension, motivating our use of the various H_0 priors.

Figure 6.4 shows the Hubble function reconstruction based on the GP algorithm using the Matern($\nu = 5/2$) kernel, provided the compiled Hubble data from the $H(z)$ measurements (CC+BAO) and H_0 values. In the inset, we show how the different priors on H_0 impact the reconstructed H_0 values at low redshift. On the other hand, it is important to note that the GP as in this straightforward application also comes with flawbacks. The most obvious being

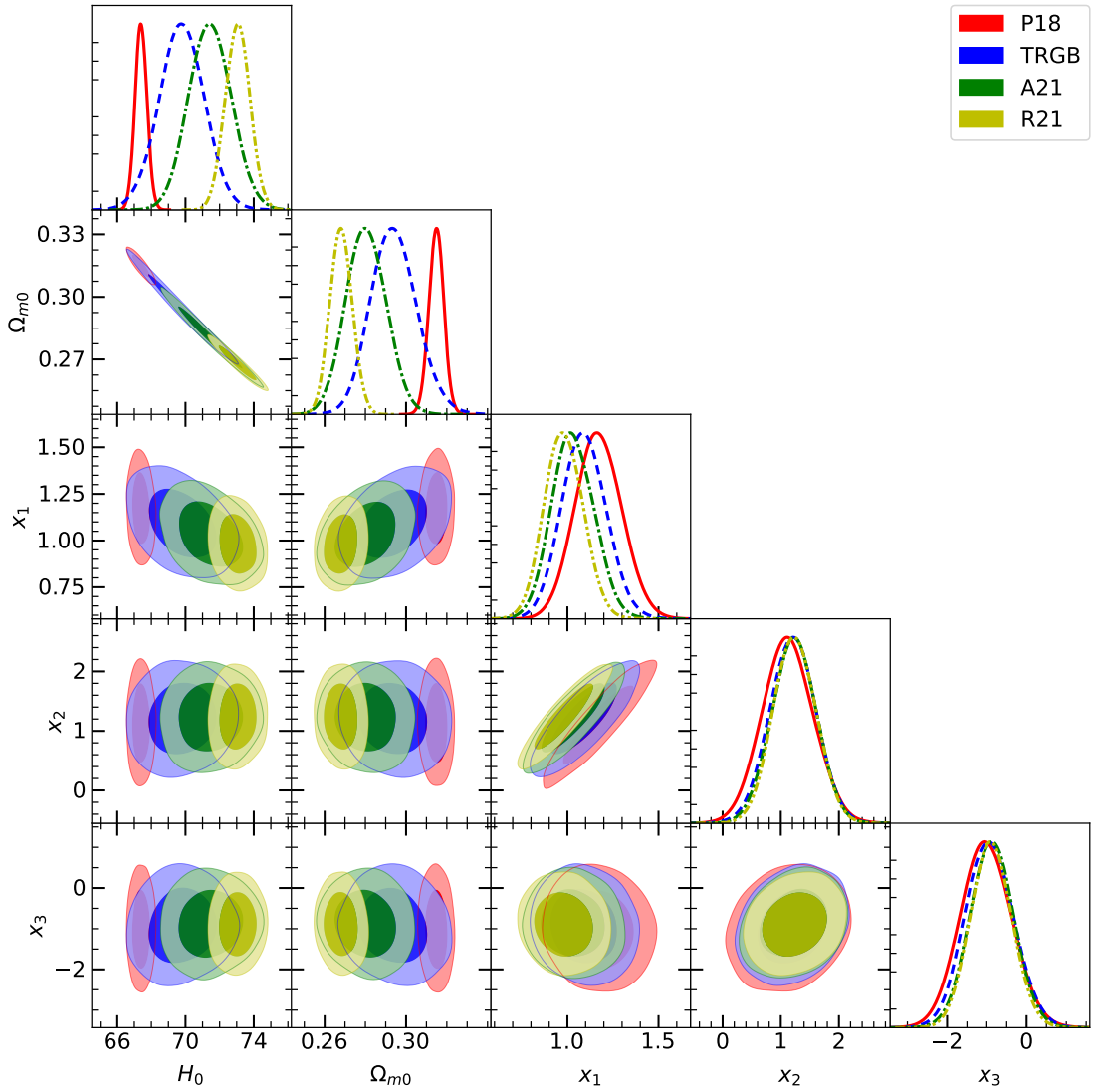


Figure 6.3: The sampled posteriors of the parameters $\{H_0, \Omega_{m0}, x_1, x_2, x_3\}$ in cubic parametrized DE for each H_0 prior: $H_0^{\rm P18} = 67.4 \pm 0.5$ km s^{-1}Mpc^{-1}, $H_0^{\rm TRGB} = 69.8 \pm 1.9$ km s^{-1}Mpc^{-1}, and $H_0^{\rm A21} = 71.5 \pm 1.8$ km s^{-1}Mpc^{-1}, and $H_0^{\rm R21} = 73.04 \pm 1.04$ km s^{-1}Mpc^{-1}. These were obtained with the base Hubble data (CC + BAO).

underestimating uncertainties [241], as seen in 6.4, but also the risk of overfitting [213]. The result of the metrics χ^2, $\mathcal{D}$, and γ^2 from the Hubble data are presented in Table 6.1. We can see that both parametric and non-parametric methods outperforms the ΛCDM model. This holds independent of the choice of an H_0 prior and can be observed for each of the metrics where the ΛCDM values always bring the largest deviation throughout. This may also be viewed as unsurprising, considering the fact that the parametric models fit more parameters than ΛCDM, while the GP tend to overfitting. From Table 6.1, we see the proposed methods predict a $\chi^2 < \chi^2_{\Lambda\rm CDM} < N$ where $\chi^2_{\Lambda\rm CDM}$ is based on the the best fit ΛCDM model. We find that for the quadratic method, $\chi^2_{\rm P18} < \chi^2_{\rm TRGB} < \chi^2_{\rm R21} < N$, while for both the cubic method and the GP, $\chi^2_{\rm R21} < \chi^2_{\rm TRGB} < \chi^2_{\rm P18} < N$. It is also worth noting that the χ^2 for the quadratic method is generally larger than those of the cubic method and the GP which are

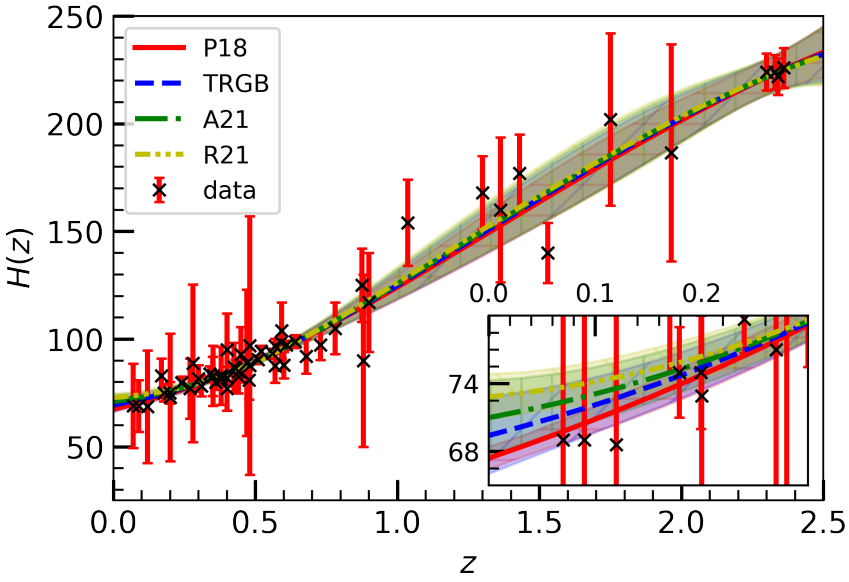


Figure 6.4: The GP reconstructed Hubble function per H_0 prior: $H_0^{\rm P18} = 67.4 \pm 0.5$ km s^{-1}Mpc^{-1}, $H_0^{\rm TRGB} = 69.8 \pm 1.9$ km s^{-1}Mpc^{-1}, $H_0^{\rm A21} = 71.5 \pm 1.8$ km s^{-1}Mpc^{-1}, and $H_0^{\rm R21} = 73.04 \pm 1.04$ km s^{-1}Mpc^{-1}. The colored and hatched parts show the region within 2σ of the GP. Hatches: (P18: "$-$"), (TRGB: "/"), (A21: "|"), (R21: "\"). The inset shows the low redshift region $z \in (0, 0.3)$ of the GP reconstructed Hubble function.

coincidentally of comparable sizes to within a few percent. It should be further noted that the results from the cubic parametrization and the GP are less sensitive to the choice of the H_0 prior. For instance, within the ΛCDM model, we find $|\chi^2_{\rm P18} - \chi^2_{\rm R21}| = 22.4$, while the corresponding values are $\Delta\chi^2 = 1.70$ and 1.00 for the cubic method and the GP, respectively. For the statistic $\mathcal{D}$, the general result is that $\mathcal{D} < 0$, implying that most of the data points can be found above the best fit line, albeit the $H(z)$ reconstruction is within the data error bars. It is also interesting to see the results for γ^2, suggesting that the quadratic method performs better than the cubic method and the GP. A more consistent trend for each method can be observed for γ^2, that $\gamma^2_{\rm quad} < \gamma^2_{\rm cubic} < \gamma^2_{\rm GP} < \gamma^2_{\Lambda\rm CDM}$, where we remind that γ^2 considers the uncertainty of the reconstruction and the data on an equal footing.

6.6.1 Reconstructed dark energy from Hubble data

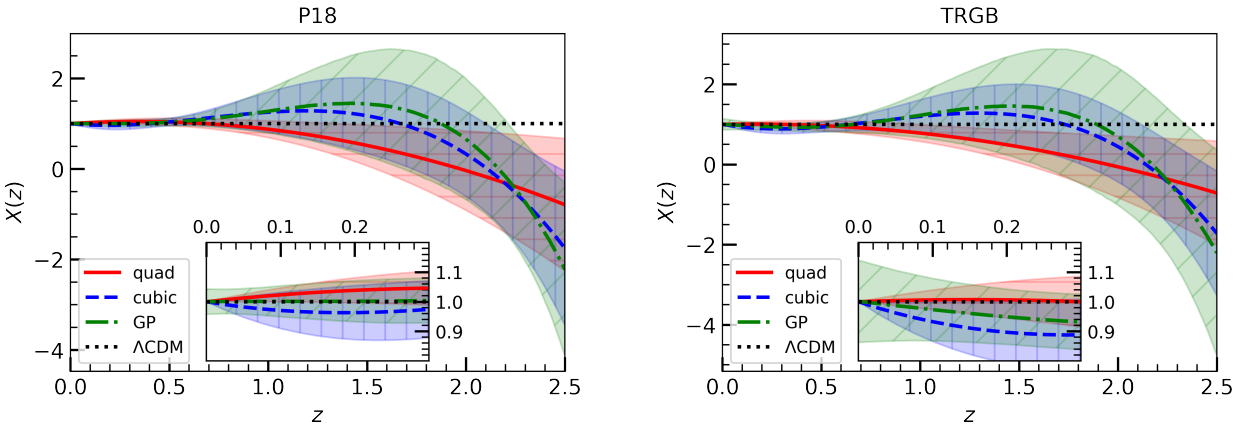
In this section, we show the reconstructed DE density and evaluate the performance of each reconstruction approach based on the three metrics presented in Section 6.3.
Figure 6.5 shows the (normalized) DE density $X(z)$ obtained by means of all the methods for each H_0 prior. The ΛCDM curves $X = 1$ appear as horizontal black dotted line.

From Fig. 6.5, it is seen that regardless of the H_0 prior, the reconstructions are statistically consistent with the ΛCDM model at low redshifts, albeit slight deviations emerge for the $H_0^{\rm R21}$ prior. Moreover, deviations from $X = 1$ emerge at high redshifts for all values of H_0, for all the methods. In general, larger H_0 lead to slightly larger deviations, with the quadratic method disfavouring ΛCDM at 2σ for $z \gtrsim 1.5$. The latter can also be seen for the cubic method and the GP method, albeit deviating from $z \sim 2.3$ where the earliest observational data can be found. The influence of the H_0 priors also come into play at low redshifts

Table 6.1: An assessment of the performance of each method using the statistics given by Eqs. (6.9), (6.10), and (6.11). The Hubble data corresponds to $N = 57$ points coming from CC (31 points) and BAO (26 points). The Planck prior for the matter fraction is $\Omega_{m0}h^2 = 0.1430 \pm 0.0011$ [2].

H_0 prior	Method/Model	χ^2	$\mathcal{D}$	γ^2
P18	ΛCDM	37.6	-3.11	30.7
	Parametric (quadratic)	30.6	-3.06	7.01
	Parametric (cubic)	27.0	-3.08	14.3
	Nonparametric (GP)	26.6	-2.19	25.2
TRGB	ΛCDM	38.6	1.27	24.6
	Parametric (quadratic)	32.1	-2.28	6.15
	Parametric (cubic)	26.1	-1.69	10.8
	Nonparametric (GP)	26.1	-1.03	24.4
A21	ΛCDM	42.8	4.29	25.2
	Parametric (quadratic)	33.8	-1.84	6.99
	Parametric (cubic)	25.7	-1.13	11.3
	Nonparametric (GP)	25.8	-0.25	24.0
R21	ΛCDM	60.0	15.5	34.8
	Parametric (quadratic)	36.4	-0.97	8.16
	Parametric (cubic)	25.3	-0.72	13.3
	Nonparametric (GP)	25.6	0.53	23.6

$z \lesssim 0.3$, as seen in the insets of Figure 6.5. Nevertheless, the $H_0^{\rm P18}$ prior shows consistency with ΛCDM model in this low redshift region. Overall, the three methods, despite their inherent differences, appear to suggest dynamical dark energy, or more precisely, an evolving $X(z)$ as shown in Figure 6.5. Figure 6.5 shows an interesting trend, which is that all reconstruction methods exhibit $X(z) < 0$ at the high redshift regime. Our findings are consistent with other recent reports using different methodologies and data sets [242, 212, 243]. We refer the reader to [242] for the implications of the sign-change in $X(z)$ (but also on the theoretical implications of $w(z) = -1$ crossing) in the context of modified gravity scenarios. We also highlight the difference between the low redshift reconstruction of the GP and the parametrized approaches. In the parametrized reconstructions, the uncertainties shrinks to an infinitely narrow size closer to $z = 0$. This can be traced from the fact that these



(a) $H_0^{\mathrm{P18}} = 67.4 \pm 0.5$ km s^{-1}Mpc^{-1}

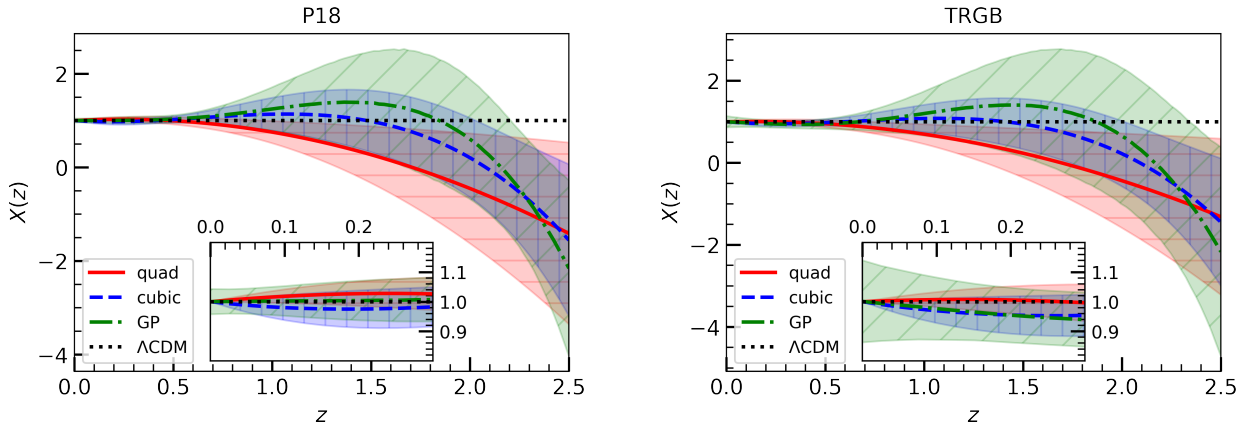
(b) $H_0^{\mathrm{TRGB}} = 69.8 \pm 1.9$ km s^{-1}Mpc^{-1}

(c) $H_0^{\mathrm{A21}} = 71.5 \pm 1.8$ km s^{-1}Mpc^{-1}

(d) $H_0^{\mathrm{R21}} = 73.04 \pm 1.04$ km s^{-1}Mpc^{-1}

Figure 6.5: The reconstructed normalized DE per method derived from the base Hubble data (CC + BAO) for each H_0 prior: (a) P18, (b) TRGB, (c) A21, and (d) R21. *Legends:* "quad" and "cubic" stands for the quadratic and cubic parametrized DE, respectively; "GP" for the GPs. The colored and hatched regions show the 2σ confidence interval of the reconstructions. *Hatches:* (quad: "−"), (cubic: "|"), (GP: "/"). The inset zooms in on the low redshift region $z \in (0, 0.3)$.

parametrized methods are an expansion about the redshift, i.e., $X(z) \sim 1 + az + bz^2 + O(z^3)$ for constants $a, b, \cdots$. The GP, on the other hand, does not share this feature, and continues to be able to interpolate for low redshifts even down to $z = 0$ since $X(z = 0) = 1$ by construction. In summary, from the $X(z)$ reconstructions, our result show that both parametric and non-parametric approaches somehow agree in their macrophysical implications, despite their inherent differences. The overall trend in the redshift evolution of $X(z)$ is shared by all the methods, with only the cubic case and GP differing from the quadratic evolution at intermediate redshifts due to their intrinscally larger freedom to fit the data. Even though each method is unequivocally challenging in its own way and likely leaving traces of non-physical artefacts in their reconstruction, all agree on a conclusion despite this difference.



(a) $H_0^{\rm P18} = 67.4 \pm 0.5$ km s^{-1}Mpc^{-1}

(b) $H_0^{\rm TRGB} = 69.8 \pm 1.9$ km s^{-1}Mpc^{-1}

(c) $H_0^{\rm A21} = 71.5 \pm 1.8$ km s^{-1}Mpc^{-1}

(d) $H_0^{\rm R21} = 73.04 \pm 1.04$ km s^{-1}Mpc^{-1}

Figure 6.6: The reconstructed normalized DE density per method derived from the base Hubble data (CC + BAO) and supernovae observations (Pantheon/MCT) for each H_0 prior: (a) P18, (b) TRGB, (c) A21, and (d) R21. *Legends:* "quad" and "cubic" stands for the quadratic and cubic parametrized DE, respectively; "GP" for the Gaussian processes. The colored-hatched regions show the 2σ-region of the reconstructions while the insets reveal a magnified view of the low redshift region $z \in (0, 0.3)$. *Hatches:* (quad: "−"), (cubic: "|"), (GP: "/").

6.6.2 Reconstructed dark energy from Supernovae

In this section, we study the reconstructed $X(z)$ by including supernovae observations in the analysis. This is done using the 1048 SNe observations from the Pantheon sample [232] for the parametric methods and through the $E(z)$ measurements from the CANDELS and CLASH Multi-Cycle Treasury (MCT) data [233] for the GP. For the GP, we use the H_0 priors to convert this into $H(z) = H_0 E(z)$ measurements, which are then subsequently used together with the base Hubble data [234, 235]. The reconstructed normalized DE is shown in 6.6, with the ΛCDM curve being depicted by horizontal dotted lines.

We found an overall trend similar to the reconstructions obtained with the base Hubble data only. Therefore, the same observations made for the $X(z)$ reconstructions with the base

Hubble data (CC + BAO) therefore holds in this case including information from supernovae (Pantheon/MCT). That is, there is a more than 2σ deviation from the ΛCDM model at high redshifts $z \gtrsim 2.3$ for any method and H_0 prior, while this deviation occurs earlier in redshift for the quadratic model.
The overall picture from this is that the deductions that were made in the previous sections about the dynamical nature of DE still hold, or even strengthened, with the inclusion of supernovae observations.

6.7 Discussion

Despite its discovery many decades ago, a theoretical understanding of dark energy (DE) remains elusive. To explore and shed light on the properties of this component, we adopted a phenomenological approach to study the evolution of dark energy. In our analysis, we incorporate four H_0 priors that account for the Hubble tension, along with the Planck prior on $\Omega_{m0}h^2$ (therefore, compromising the notion of a model-independent approach). Interestingly, even with the Planck priors on the Hubble constant and matter density, which are consistent with the ΛCDM model, both the parametric and non-parametric approaches employed in this work support a deviation from standard cosmology in the form of an evolving DE component. The most robust evidence for dark energy evolution is found at high redshifts ($z \sim 2.3$), particularly near the Lyman alpha BAO measurements. This finding aligns with previous reports of BAO measurements [223, 220, 224, 231, 244], which also indicate tension with the ΛCDM model. We have demonstrated that despite the inherent difference in how the different methods analyze the data, there is an overall trend for the evolution of dark energy. This evolving dark energy revealed by all the methods, data sets and H_0 priors is the highlight of this work. If this feature had emerged using only parametric or nonparametric approaches, it could have been easily doubted as a mere artifact of the methods.
However, it is important to note that the methods utilized in our study come with their own limitations. For instance, parametric methods often face criticism for the arbitrary nature of the parametrization of phenomenological functions [245], while the GP method may suffer of the underestimation of uncertainties [213, 241, 246]. Indeed, the GP reconstruction of the Hubble function $H(z)$ presented in Fig. 6.4 indicates a clear underestimation of uncertainties. By appropriately propagating errors, the observed deviations in the reconstruction shown in Fig. 6.6 may become less significant. Given these considerations, a natural extension of this work is to incorporate various other approaches [247, 248, 242, 243] and study whether the trends we found in this paper for the dark energy evolution can still be observed.
On the other hand, it should be emphasized as well that dynamical DE does not necessarily imply modified gravity as alluded in [249]. To make robust conclusions about modified gravity, for example, one can instead extend this work to accommodate observations of redshift space distorsions. Relevant progress in this direction are [250, 251].
Finally, we find it interesting to investigate how the methods employed in this work perform when considering observables at the perturbative level. We will leave this for future work.

Conclusions

In cosmology, our ultimate goal is to unveil the physical model that describes the properties and evolution of our Universe. This endeavor has been pursued for years, leading to significant advances in our understanding of the Universe. Observations of the cosmic microwave background radiation, Supernovae Ia, and the large-scale structure have established the ΛCDM model as the concordance model, albeit requiring the inclusion of two exotic and as-yet-unknown components: dark matter and dark energy. Numerous efforts have been devoted to designing cosmological experiments capable of yielding valuable data, sensitive to the properties and nature of these components. However, without careful study of the methods we employ to analyze such data, we may not be capable of drawing accurate conclusions about the underlying cosmological model. The work presented in this thesis contributes to the latter, employing statistical tools and machine learning techniques to analyze cosmological data and provide unbiased estimates of the model parameters.

After an overview of machine learning tools and statistical methods in Chapter 2, we present the formalism behind weak gravitational lensing in Chapter 3. The study of weak gravitational lensing has become increasingly considered as it is one of the main cosmological probes to be targeted by the next generation of cosmology experiments. In Chapter 4, we provide for the first time a Bayesian framework to perform parameter inference when the theoretical predictions are assisted by neural networks. The latter is a method firmly considered for the analysis of large-scale structure probes, such as weak gravitational lensing, as the matter perturbations rely on the computation of accurate but slow Boltzmann codes. Our formalism consist in a new posterior distribution for parameter inference, whose shape accounts for the error introduced by the neural net approximation of the target function. In this manner, this adapted-posterior propagates the neural net uncertainties and debias parameter inference. We show that our method can be easily implemented in the case of Gaussian distributions. We train two different neural networks to learn the mapping from cosmological parameters to the matter power spectrum at $z = 1$. In both cases, the parameter inference assisted by these neural networks can recover the true cosmology when our methodology is applied. Therefore, we consider it highly beneficial to incorporate this method in future galaxy survey pipelines, either to directly debias inference or to examine the accuracy needed in the neural network predictions to obtain unbiased and precise cosmological constraints. For future work, we plan to extend this analysis to directly emulate the weak lensing power spectrum in a tomographic analysis, making it easier to implement in the galaxy surveys pipelines. Along the same lines, we intend to build a public code that can facilitate the incorporation of our methodology.

In Chapter 5 we investigate the impact of baryonic feedback on the weak lensing power

spectrum and non-Gaussian statistics. We utilize HSC-Y1 mocks based on hydrodynamical and N-body simulations. Our methodology consists in simulation-based analysis where a baryon contamination factor is injected in the likelihood data vector. The results show that baryons can introduce severe biases in the inferred S_8 parameter. The BAHAMAS high-AGN simulation introduces the more drastic effects, especially impacting the power spectrum, minimum counts and PDF measurements. On the other hand, we found that the peak counts is the statistics with better constraining power and, at the same time, less sensitive to the influence of baryons (as accounted in this paper). This is an important result, as the peak counts can be easily implemented in future weak lensing pipelines. Finally, we conducted an analysis on HSC-Y1 real data, finding that the results obtained from both large and small scales are statistically consistent. This leads to the conclusion that, for current Stage-III surveys such as HSC, baryonic physics does not seem to play a major role in the bias of cosmological parameters. This is an interesting finding as the baryonic feedback is considered one of the main potential systematic that can induce a lower amplitude of structures S_8 [20] compared to the value derived by Planck. The results obtained in this work requires further validation for the upcoming surveys. For this reason, ongoing work is being conducted by applying this same methodology to LSST mock data. However, the precision of cosmological parameters estimates in LSST forecasts demands a careful analysis of the emulators and likelihood. As future research, we plan to delve more deeply into the accuracy needed for non-Gaussian statistics emulators, following a similar approach as presented in Chapter 4. Additionally, other strategies to account for baryons will be explored, so that we do not have to rely solely on simulations at different cosmologies and resolutions.

In Chapter 6, we present our results regarding the study of dark energy evolution using both parametric and non-parametric methods. The non-parametric method employs the Gaussian Process algorithm, which is also applied in Chapter 5. We found that all methods, prior distributions, and data sets agree on a dynamical dark energy scenario, characterized by a decreasing value of $X(z)$ with redshift. Moreover, the cubic parametrization and the GP method exhibit strong agreement across all reconstruction schemes. However, the statistical significance of the non-parametric method needs to be revisited, as the method clearly underestimates its uncertainties. For future work, we are exploring the inclusion of a regularization parameter in the GP, which could lead to better posterior estimates. The results presented in this work are obtained from a completely phenomenological approach, nonetheless it is still important to account for the physical theories that can explain a dynamical dark energy. For instance, models with interaction between dark components [252, 28] can potentially lead to an effective $X(z)$ that decreases with redshift.

Bibliography

[1] Catherine Heymans, Ludovic Van Waerbeke, Lance Miller, Thomas Erben, Hendrik Hildebrandt, Henk Hoekstra, Thomas D Kitching, Yannick Mellier, Patrick Simon, Christopher Bonnett, et al. Cfhtlens: the canada–france–hawaii telescope lensing survey. *Monthly Notices of the Royal Astronomical Society*, 427(1):146–166, 2012.

[2] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020. [Erratum: Astron.Astrophys. 652, C4 (2021)].

[3] N. Aghanim et al. Planck 2018 results. I. Overview and the cosmological legacy of Planck. *Astron. Astrophys.*, 641:A1, 2020.

[4] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and The Supernova Cosmology Project. Measurements of Ω and Λ from 42 High-Redshift Supernovae. ApJ, 517(2):565–586, June 1999.

[5] Michele Moresco et al. Unveiling the Universe with emerging cosmological probes. *Living Rev. Rel.*, 25(1):6, 2022.

[6] Md. Shoaibur Rahman. Computations, optimization and tuning of deep feedforward neural networks. *bioRxiv*, 2019.

[7] Nora Elisa Chisari et al. Modelling baryonic feedback for survey cosmology. *Open J. Astrophys.*, 2(1):4, 2019.

[8] Masayuki Tanaka, Jean Coupon, Bau-Ching Hsieh, Sogo Mineo, Atsushi J Nishizawa, Joshua Speagle, Hisanori Furusawa, Satoshi Miyazaki, and Hitoshi Murayama. Photometric redshifts for hyper suprime-cam subaru strategic program data release 1. *Publications of the Astronomical Society of Japan*, 70(SP1):S9, 2018.

[9] Chiaki Hikage, Masamune Oguri, Takashi Hamana, Surhud More, Rachel Mandelbaum, Masahiro Takada, Fabian Köhlinger, Hironao Miyatake, Atsushi J Nishizawa, Hiroaki Aihara, et al. Cosmology from cosmic shear power spectra with subaru hyper suprime-cam first-year data. *Publications of the Astronomical Society of Japan*, 71(2):43, 2019.

[10] Takashi Hamana, Masato Shirasaki, Satoshi Miyazaki, Chiaki Hikage, Masamune Oguri, Surhud More, Robert Armstrong, Alexie Leauthaud, Rachel Mandelbaum, Hironao Miyatake, Atsushi J. Nishizawa, Melanie Simet, Masahiro Takada, Hiroaki Aihara, James Bosch, Yutaka Komiyama, Robert Lupton, Hitoshi Murayama, Michael A. Strauss, and Masayuki Tanaka. Cosmological constraints from cosmic shear two-point correlation functions with HSC survey first-year data. PASJ, 72(1):16, February 2020.

[11] A. Einstein. Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik*, 354(7):769–822, January 1916.

[12] B. P. Abbott et al. LIGO: The Laser interferometer gravitational-wave observatory. *Rept. Prog. Phys.*, 72:076901, 2009.

[13] Adam G. Riess et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.*, 116:1009–1038, 1998.

[14] S. Perlmutter et al. Measurements of the cosmological parameters Omega and Lambda from the first 7 supernovae at z>=0.35. *Astrophys. J.*, 483:565, 1997.

[15] LSST Science Collaboration, Paul A. Abell, Julius Allison, Scott F. Anderson, John R. Andrew, J. Roger P. Angel, Lee Armus, David Arnett, S. J. Asztalos, Tim S. Axelrod, Stephen Bailey, D. R. Ballantyne, Justin R. Bankert, Wayne A. Barkhouse, Jeffrey D. Barr, L. Felipe Barrientos, Aaron J. Barth, James G. Bartlett, Andrew C. Becker, Jacek Becla, Timothy C. Beers, Joseph P. Bernstein, Rahul Biswas, Michael R. Blanton, Joshua S. Bloom, John J. Bochanski, Pat Boeshaar, Kirk D. Borne, Marusa Bradac, W. N. Brandt, Carrie R. Bridge, Michael E. Brown, Robert J. Brunner, James S. Bullock, Adam J. Burgasser, James H. Burge, David L. Burke, Phillip A. Cargile, Srinivasan Chandrasekharan, George Chartas, Steven R. Chesley, You-Hua Chu, David Cinabro, Mark W. Claire, Charles F. Claver, Douglas Clowe, A. J. Connolly, Kem H. Cook, Jeff Cooke, Asantha Cooray, Kevin R. Covey, Christopher S. Culliton, Roelof de Jong, Willem H. de Vries, Victor P. Debattista, Francisco Delgado, Ian P. Dell'Antonio, Saurav Dhital, Rosanne Di Stefano, Mark Dickinson, Benjamin Dilday, S. G. Djorgovski, Gregory Dobler, Ciro Donalek, Gregory Dubois-Felsmann, Josef Durech, Ardis Eliasdottir, Michael Eracleous, Laurent Eyer, Emilio E. Falco, Xiaohui Fan, Christopher D. Fassnacht, Harry C. Ferguson, Yanga R. Fernandez, Brian D. Fields, Douglas Finkbeiner, Eduardo E. Figueroa, Derek B. Fox, Harold Francke, James S. Frank, Josh Frieman, Sebastien Fromenteau, Muhammad Furqan, Gaspar Galaz, A. Gal-Yam, Peter Garnavich, Eric Gawiser, John Geary, Perry Gee, Robert R. Gibson, Kirk Gilmore, Emily A. Grace, Richard F. Green, William J. Gressler, Carl J. Grillmair, Salman Habib, J. S. Haggerty, Mario Hamuy, Alan W. Harris, Suzanne L. Hawley, Alan F. Heavens, Leslie Hebb, Todd J. Henry, Edward Hileman, Eric J. Hilton, Keri Hoadley, J. B. Holberg, Matt J. Holman, Steve B. Howell, Leopoldo Infante, Zeljko Ivezic, Suzanne H. Jacoby, Bhuvnesh Jain, R, Jedicke, M. James Jee, J. Garrett Jernigan, Saurabh W. Jha, Kathryn V. Johnston, R. Lynne Jones, Mario Juric, Mikko Kaasalainen, Styliani, Kafka, Steven M. Kahn, Nathan A. Kaib, Jason Kalirai, Jeff Kantor, Mansi M. Kasliwal, Charles R. Keeton, Richard Kessler, Zoran Knezevic, Adam Kowalski, Victor L. Krabbendam, K. Simon Krughoff, Shrinivas Kulkarni, Stephen Kuhlman, Mark Lacy, Sebastien Lepine, Ming Liang, Amy Lien, Paulina Lira,

Knox S. Long, Suzanne Lorenz, Jennifer M. Lotz, R. H. Lupton, Julie Lutz, Lucas M. Macri, Ashish A. Mahabal, Rachel Mandelbaum, Phil Marshall, Morgan May, Peregrine M. McGehee, Brian T. Meadows, Alan Meert, Andrea Milani, Christopher J. Miller, Michelle Miller, David Mills, Dante Minniti, David Monet, Anjum S. Mukadam, Ehud Nakar, Douglas R. Neill, Jeffrey A. Newman, Sergei Nikolaev, Martin Nordby, Paul O'Connor, Masamune Oguri, John Oliver, Scot S. Olivier, Julia K. Olsen, Knut Olsen, Edward W. Olszewski, Hakeem Oluseyi, Nelson D. Padilla, Alex Parker, Joshua Pepper, John R. Peterson, Catherine Petry, Philip A. Pinto, James L. Pizagno, Bogdan Popescu, Andrej Prsa, Veljko Radcka, M. Jordan Raddick, Andrew Rasmussen, Arne Rau, Jeonghee Rho, James E. Rhoads, Gordon T. Richards, Stephen T. Ridgway, Brant E. Robertson, Rok Roskar, Abhijit Saha, Ata Sarajedini, Evan Scannapieco, Terry Schalk, Rafe Schindler, Samuel Schmidt, Sarah Schmidt, Donald P. Schneider, German Schumacher, Ryan Scranton, Jacques Sebag, Lynn G. Seppala, Ohad Shemmer, Joshua D. Simon, M. Sivertz, Howard A. Smith, J. Allyn Smith, Nathan Smith, Anna H. Spitz, Adam Stanford, Keivan G. Stassun, Jay Strader, Michael A. Strauss, Christopher W. Stubbs, Donald W. Sweeney, Alex Szalay, Paula Szkody, Masahiro Takada, Paul Thorman, David E. Trilling, Virginia Trimble, Anthony Tyson, Richard Van Berg, Daniel Vanden Berk, Jake VanderPlas, Licia Verde, Bojan Vrsnak, Lucianne M. Walkowicz, Benjamin D. Wandelt, Sheng Wang, Yun Wang, Michael Warner, Risa H. Wechsler, Andrew A. West, Oliver Wiecha, Benjamin F. Williams, Beth Willman, David Wittman, Sidney C. Wolff, W. Michael Wood-Vasey, Przemek Wozniak, Patrick Young, Andrew Zentner, and Hu Zhan. LSST Science Book, Version 2.0. *arXiv e-prints*, page arXiv:0912.0201, December 2009.

[16] R. Laureijs, J. Amiaux, S. Arduini, J. L. Augueres, J. Brinchmann, R. Cole, et al. Euclid definition study report, 2011.

[17] Alessio Spurio Mancini, Davide Piras, Justin Alsing, Benjamin Joachimi, and Michael P. Hobson. COSMOPOWER: emulating cosmological power spectra for accelerated Bayesian inference from next-generation surveys. MNRAS, 511(2):1771–1788, April 2022.

[18] Andrea Manrique-Yus and Elena Sellentin. Euclid-era cosmology for everyone: neural net assisted MCMC sampling for the joint 3×2 likelihood. MNRAS, 491(2):2655–2663, Jan 2020.

[19] Supranta S. Boruah, Tim Eifler, Vivian Miranda, and Sai Krishanth P. M. Accelerating cosmological inference with Gaussian processes and neural networks – an application to LSST Y1 weak lensing and galaxy clustering. *arXiv e-prints*, page arXiv:2203.06124, March 2022.

[20] Alexandra Amon and George Efstathiou. A non-linear solution to the S8 tension? *Monthly Notices of the Royal Astronomical Society*, 516(4):5355–5366, 09 2022.

[21] William R. Coulton, Jia Liu, Ian G. McCarthy, and Ken Osato. Weak Lensing Minima and Peaks: Cosmological Constraints and the Impact of Baryons. *Mon. Not. Roy. Astron. Soc.*, 495(3):2531–2542, 2020.

[22] Roohi Dalal et al. Hyper Suprime-Cam Year 3 Results: Cosmology from Cosmic Shear Power Spectra. 4 2023.

[23] Tianhuan Lu, Zoltán Haiman, and Xiangchong Li. Cosmological constraints from HSC survey first-year data using deep learning. *Mon. Not. Roy. Astron. Soc.*, 521(2):2050–2066, 2023.

[24] Daniela Grandón and Elena Sellentin. Bayesian error propagation for neural-net based parameter inference. *Open J. Astrophys.*, 5, 2022.

[25] Daniela Grandón, Gabriela A. Marques, Leander Thiele, Jia Liu, Masato Shirasaki, and Ken Osato. The impact of baryonic feedback on HSC Y1 weak lensing non-gaussian statistics. *in prep*, 2023.

[26] Gabriela A. Marques, Jia Liu, Masato Shirasaki, Leander Thiele, Daniela Grandón, Kevin M. Huffenberger, Sihao Cheng, Joachim Harnois-Déraps, and Ken Osato. Cosmology from weak lensing peaks and minima with Subaru Hyper Suprime-Cam survey first-year data. *arXiv preprint 2308.10866*, 2023.

[27] Hiroaki Aihara, Nobuo Arimoto, Robert Armstrong, Stéphane Arnouts, Neta A Bahcall, et al. The Hyper Suprime-Cam SSP Survey: Overview and survey design. *Publications of the Astronomical Society of Japan*, 70(SP1), 09 2017. S4.

[28] Daniela Grandón and Victor H Cardenas. Studies on dark energy evolution. *Class. Quantum Gravity*, 38(14):145008, jun 2021.

[29] Reginald Christian Bernardo, Daniela Grandón, Jackson Levi Said, and Víctor H. Cárdenas. Parametric and nonparametric methods hint dark energy evolution. *Physics of the Dark Universe*, 36:101017, 2022.

[30] Scott Dodelson. *Modern Cosmology*. Academic Press, Amsterdam, 2003.

[31] Dragan Huterer. *A Course in Cosmology: From Theory to Practice*. Cambridge University Press, 2023.

[32] P. J. E. Peebles. *The large-scale structure of the universe*. 1980.

[33] N. Kaiser. *Elements of Astrophysics*. CreateSpace Independent Publishing Platform, 2014.

[34] Edwin Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the National Academy of Sciences*, 15(3):168–173, 1929.

[35] A. Friedmann. Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes. *Zeitschrift fur Physik*, 21(1):326–332, December 1924.

[36] Abbé G. Lemaître. A Homogeneous Universe of Constant Mass and Increasing Radius accounting for the Radial Velocity of Extra-galactic Nebulæ. *Monthly Notices of the Royal Astronomical Society*, 91(5):483–490, 03 1931.

[37] H. P. Robertson. Kinematics and World-Structure. ApJ, 82:284, November 1935.

[38] A. G. Walker. On Milne's Theory of World-Structure*. *Proceedings of the London Mathematical Society*, s2-42(1):90–127, 01 1937.

[39] James M. Bardeen, J. R. Bond, Nick Kaiser, and A. S. Szalay. The Statistics of Peaks of Gaussian Random Fields. *Astrophys. J.*, 304:15–61, 1986.

[40] Antony Lewis, Anthony Challinor, and Anthony Lasenby. Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. ApJ, 538(2):473–476, August 2000.

[41] Antony Lewis and Anthony Challinor. CAMB: Code for Anisotropies in the Microwave Background. Astrophysics Source Code Library, record ascl:1102.026, February 2011.

[42] Julien Lesgourgues. The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview. *arXiv preprint arXiv:1104.2932*, 2011.

[43] Raul E. Angulo and Oliver Hahn. Large-scale dark matter simulations. *Living Rev Comput Astrophys 8, 1*, 2022.

[44] George F. Smoot et al. First results of the COBE satellite measurement of the anisotropy of the cosmic microwave background radiation. In *2nd Rencontre de Blois on Physical Cosmology: 25th Anniversary of the Cosmic Background Radiation Discovery*, pages 105–120, 5 1991.

[45] G. F. R. Ellis and J. E. Baldwin. On the expected anisotropy of radio source counts. *Monthly Notices of the Royal Astronomical Society*, 206(2):377–381, 01 1984.

[46] M. Hazumi et al. LiteBIRD: A Satellite for the Studies of B-Mode Polarization and Inflation from Cosmic Background Radiation Detection. *J. Low Temp. Phys.*, 194(5-6):443–452, 2019.

[47] T. Matsumura et al. Mission design of LiteBIRD. *J. Low Temp. Phys.*, 176:733, 2014.

[48] Peter Ade et al. The Simons Observatory: Science goals and forecasts. *JCAP*, 02:056, 2019.

[49] Frank J Qu, Blake D Sherwin, Mathew S Madhavacheril, Dongwon Han, Kevin T Crowley, Irene Abril-Cabezas, Peter AR Ade, Simone Aiola, Tommy Alford, Mandana Amiri, et al. The Atacama Cosmology Telescope: A Measurement of the DR6 CMB Lensing Power Spectrum and its Implications for Structure Growth. *arXiv preprint arXiv:2304.05202*, 2023.

[50] GA Marques, MS Madhavacheril, O Darwish, S Shaikh, M Aguena, O Alves, S Avila, D Bacon, EJ Baxter, K Bechtol, et al. Cosmological constraints from the tomography of des-y3 galaxies with cmb lensing from act dr4. *arXiv preprint arXiv:2306.17268*, 2023.

[51] M. M. Phillips. The absolute magnitudes of Type IA supernovae. *Astrophys. J. Lett.*, 413:L105–L108, 1993.

[52] Mario Hamuy. The acceleration of the Universe in the light of supernovae – The key role of the Cerro Tololo Inter-American Observatory. 2013.

[53] S. Perlmutter, S. Gabi, G. Goldhaber, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, R. Pain, C. R. Pennypacker, I. A. Small, R. S. Ellis, R. G. McMahon, B. J. Boyle, P. S. Bunclark, D. Carter, M. J. Irwin, K. Glazebrook, H. J. M. Newberg, A. V. Filippenko, T. Matheson, M. Dopita, and W. J. Couch. Measurements of the Cosmological Parameters Ω and Λ from the First Seven Supernovae at z >= 0.35. ApJ, 483(2):565–581, July 1997.

[54] Mario Hamuy, M. M. Phillips, Robert A. Schommer, Nicholas B. Suntzeff, Jose Maza, and R. Aviles. The Absolute luminosities of the Calan/Tololo type IA supernovae. *Astron. J.*, 112:2391, 1996.

[55] M. M. Phillips, P. Lira, N. B. a Suntzeff, R. A. Schommer, M. Hamuy, and J. Maza. The reddening-free decline rate versus luminosity relationship for type ia supernovae. *Astron. J.*, 118:1766, 1999.

[56] Dragan Huterer and Michael S. Turner. Prospects for probing the dark energy via supernova distance measurements. *Phys. Rev. D*, 60:081301, 1999.

[57] Raul Jimenez and Abraham Loeb. Constraining cosmological parameters based on relative galaxy ages. *Astrophys. J.*, 573:37–42, 2002.

[58] M. Moresco et al. Improved constraints on the expansion rate of the Universe up to z˜1.1 from the spectroscopic evolution of cosmic chronometers. *JCAP*, 08:006, 2012.

[59] Ana Achúcarro et al. Inflation: Theory and Observations. *arXiv preprint arXiv:2203.08128*, 2022.

[60] Daniel Baumann. Inflation. In *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics of the Large and the Small*, pages 523–686, 2011.

[61] D. Kazanas. Dynamics of the universe and spontaneous symmetry breaking. ApJ, 241:L59–L63, October 1980.

[62] Alan H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D*, 23:347–356, Jan 1981.

[63] A.A. Starobinsky. A new type of isotropic cosmological models without singularity. *Physics Letters B*, 91(1):99–102, 1980.

[64] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, January 1933.

[65] Vera C. Rubin. The rotation of spiral galaxies. *Science*, 220(4604):1339–1344, 1983.

[66] Leszek Roszkowski, Enrico Maria Sessolo, and Sebastian Trojanowski. WIMP dark matter candidates and searches—current status and future prospects. *Rept. Prog. Phys.*, 81(6):066201, 2018.

[67] R. Alves Batista et al. EuCAPT White Paper: Opportunities and Challenges for Theoretical Astroparticle Physics in the Next Decade. 10 2021.

[68] Albert Einstein. Cosmological Considerations in the General Theory of Relativity. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1917:142–152, 1917.

[69] Eric V. Linder. Exploring the expansion history of the universe. *Phys. Rev. Lett.*, 90:091301, 2003.

[70] Luca Amendola and Shinji Tsujikawa. *Dark Energy: Theory and Observations*. Cambridge University Press, 2010.

[71] Ariel G. Sánchez, C. M. Baugh, W. J. Percival, J. A. Peacock, N. D. Padilla, S. Cole, C. S. Frenk, and P. Norberg. Cosmological parameters from cosmic microwave background measurements and the final 2dF Galaxy Redshift Survey power spectrum. MNRAS, 366(1):189–207, February 2006.

[72] Eleonora Di Valentino. Challenges of the Standard Cosmological Model. *Universe*, 8(8):399, 2022.

[73] Dillon Brout et al. The Pantheon+ Analysis: Cosmological Constraints. *Astrophys. J.*, 938(2):110, 2022.

[74] Eleonora Di Valentino et al. Snowmass2021 - Letter of interest cosmology intertwined II: The hubble constant tension. *Astropart. Phys.*, 131:102605, 2021.

[75] Leandros Perivolaropoulos and Foteini Skara. Challenges for ΛCDM: An update. *New Astron. Rev.*, 95:101659, 2022.

[76] Yang Liu, Hongwei Yu, and Puxun Wu. Cosmological-model-independent Determination of Hubble Constant from Fast Radio Bursts and Hubble Parameter Measurements. *Astrophys. J. Lett.*, 946(2):L49, 2023.

[77] Marc Kamionkowski and Adam G. Riess. The Hubble Tension and Early Dark Energy. *arXiv preprint arXiv:2211.04492*, 11 2022.

[78] Yan-Hong Yao and Xin-He Meng. Can interacting dark energy with dynamical coupling resolve the Hubble tension. *arXiv preprint arXiv:2207.05955*, 2022.

[79] George Efstathiou. H0 Revisited. *Mon. Not. Roy. Astron. Soc.*, 440(2):1138–1152, 2014.

[80] Eleonora Di Valentino et al. Snowmass2021 - Letter of interest cosmology intertwined I: Perspectives for the next decade. *Astropart. Phys.*, 131:102606, 2021.

[81] T. Abbott et al. The Dark Energy Survey: more than dark energy – an overview. *Mon. Not. Roy. Astron. Soc.*, 460(2):1270–1299, 2016.

[82] A. Amon et al. Dark Energy Survey Year 3 results: Cosmology from cosmic shear and robustness to data calibration. *Phys. Rev. D*, 105(2):023514, 2022.

[83] Konrad Kuijken et al. Gravitational Lensing Analysis of the Kilo Degree Survey. *Mon. Not. Roy. Astron. Soc.*, 454(4):3500–3532, 2015.

[84] Eleonora Di Valentino, Céline Bøehm, Eric Hivon, and François R. Bouchet. Reducing the H_0 and σ_8 tensions with Dark Matter-neutrino interactions. *Phys. Rev. D*, 97(4):043513, 2018.

[85] Eleonora Di Valentino and Sarah Bridle. Exploring the Tension between Current Cosmic Microwave Background and Cosmic Shear Data. *Symmetry*, 10(11):585, 2018.

[86] Calvin Preston, Alexandra Amon, and George Efstathiou. A non-linear solution to the S_8 tension II: Analysis of DES Year 3 cosmic shear. *arXiv preprint arXiv:2305.09827*, 2023.

[87] T. M. C. Abbott et al. DES Y3 + KiDS-1000: Consistent cosmology combining cosmic shear surveys. *arXiv preprint arXiv:2305.17173*, 2023.

[88] A.N. Kolmogorov. *Foundations of the Theory of Probability*. Martino Fine Books, 2013.

[89] James Joyce. Bayes' Theorem. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2019 edition, 2019.

[90] Karl Pearson. "Das Fehlergesetz Und Seine Verallgemeiner-Ungen Durch Fechner Und Pearson*." A Rejoinder. *Biometrika*, 4(1-2):169–212, 06 1905.

[91] Phil Gregory. *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support*. Cambridge University Press, 2005.

[92] Henrique S. Xavier, Filipe B. Abdalla, and Benjamin Joachimi. Improving lognormal models for cosmological fields. *Mon. Not. Roy. Astron. Soc.*, 459(4):3693–3710, 2016.

[93] Robert L. Schuhmann, Benjamin Joachimi, and Hiranya V. Peiris. Gaussianization for fast and accurate inference from cosmological data. *Mon. Not. Roy. Astron. Soc.*, 459(2):1916–1928, 2016.

[94] B. Joachimi and A. N. Taylor. Forecasts of non-Gaussian parameter spaces using Box–Cox transformations. *Monthly Notices of the Royal Astronomical Society*, 416(2):1010–1022, 09 2011.

[95] Alan F. Heavens, Raul Jimenez, and Ofer Lahav. Massive lossless data compression and multiple parameter estimation from galaxy spectra. MNRAS, 317(4):965–972, October 2000.

[96] Alan F. Heavens, Elena Sellentin, Damien de Mijolla, and Alvise Vianello. Massive data compression for parameter-dependent covariance matrices. MNRAS, 472(4):4244–4250, December 2017.

[97] Karl Pearson F.R.S. X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157–175, 1900.

[98] S. D. Poisson. English translation of poisson's "recherches sur la probabilité des jugements en matière criminelle et en matière civile" / "researches into the probabilities of judgements in criminal and civil cases". *arXiv preprint arXiv:1902.02782*, 2019.

[99] Elena Sellentin, Catherine Heymans, and Joachim Harnois-Déraps. The skewed weak lensing likelihood: why biases arise, despite data and theory being sound. *Mon. Not. Roy. Astron. Soc.*, 477(4):4879–4895, 2018.

[100] James A. D. Diacoumis and Yvonne Y. Y. Wong. On the prior dependence of cosmological constraints on some dark matter interactions. *JCAP*, 05:025, 2019.

[101] Andrew Gelman and Jennifer Hill. *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press, 2006.

[102] Curtis A Parvin. An Introduction to Multivariate Statistical Analysis, 3rd ed. T.W. Anderson. Hoboken, NJ: John Wiley. *Clinical Chemistry*, 50(5):981–982, 05 2004.

[103] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 12 2004.

[104] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 04 1970.

[105] Amir Hajian. Efficient Cosmological Parameter Estimation with Hamiltonian Monte Carlo. *Phys. Rev. D*, 75:083525, 2007.

[106] John Skilling. Nested sampling for general Bayesian computation. *Bayesian Analysis*, 1(4):833 – 859, 2006.

[107] Ofer Lahav. Deep Machine Learning in Cosmology: Evolution or Revolution? *arXiv preprint arXiv:2302.04324*, 2023.

[108] Adrian A. Collister and Ofer Lahav. ANNz: Estimating photometric redshifts using artificial neural networks. *Publ. Astron. Soc. Pac.*, 116:345–351, 2004.

[109] Yogesh Wadadekar. Estimating photometric redshifts using support vector machines. *Publ. Astron. Soc. Pac.*, 117:79, 2005.

[110] Arrykrishna Mootoovaloo, Alan F. Heavens, Andrew H. Jaffe, and Florent Leclercq.

Parameter inference for weak lensing using Gaussian Processes and MOPED. MNRAS, 497(2):2213–2226, September 2020.

[111] Arrykrishna Mootoovaloo, Andrew H. Jaffe, Alan F. Heavens, and Florent Leclercq. Kernel-based emulator for the 3D matter power spectrum from CLASS. *Astron. Comput.*, 38:100508, 2022.

[112] Joseph DeRose, Shi-Fan Chen, Martin White, and Nickolas Kokron. Neural network acceleration of large-scale structure theory calculations. *JCAP*, 04(04):056, 2022.

[113] Sven Günther. Uncertainty-aware and Data-efficient Cosmological Emulation using Gaussian Processes and PCA. *arXiv preprint arXiv:2307.01138*, 2023.

[114] Rodrigo Carrasco-Davis et al. Alert Classification for the ALeRCE Broker System: The Real-time Stamp Classifier. *Astron. J.*, 162(6):231, 2021.

[115] F. Förster et al. The Automatic Learning for the Rapid Classification of Events (ALeRCE) Alert Broker. *Astron. J.*, 161(5):242, 2021.

[116] F. Förster et al. DELIGHT: Deep Learning Identification of Galaxy Hosts of Transients using Multiresolution Images. *Astron. J.*, 164(5):195, 2022.

[117] Aurlien Gron. *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, Inc., 1st edition, 2017.

[118] Xue Ying. An overview of overfitting and its solutions. *Journal of Physics: Conference Series*, 1168(2):022022, feb 2019.

[119] Warren S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. Bulletin of mathematical biophysics. *The Journal of Symbolic Logic*, 9:115–133, 1943.

[120] (1958) f. rosenblatt, the perceptron: a probabilistic model for information storage and organization in the brain, psychological review,65:386-408. In *Neurocomputing, Volume 1: Foundations of Research*. The MIT Press, 04 1988.

[121] C. A. L. Bailer-Jones, R. Gupta, and H. P. Singh. An introduction to artificial neural networks. *arXiv preprint astro-ph/0102224*, 2001.

[122] Thomas Beckers. An introduction to gaussian process models, 2021.

[123] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.

[124] Rachel Mandelbaum. Weak lensing for precision cosmology. *Ann. Rev. Astron. Astrophys.*, 56:393–433, 2018.

[125] M. Meneghetti. *Introduction to Gravitational Lensing: With Python Examples*. Lecture

Notes in Physics. Springer International Publishing, 2021.

[126] Martin Kilbinger. Cosmology with cosmic shear observations: a review. *Reports on Progress in Physics*, 78(8):086901, July 2015.

[127] Andreas Albrecht, Gary Bernstein, Robert Cahn, Wendy L Freedman, Jacqueline Hewitt, Wayne Hu, John Huth, Marc Kamionkowski, Edward W Kolb, Lloyd Knox, et al. Report of the dark energy task force. *arXiv preprint astro-ph/0609591*, 2006.

[128] Matthias Bartelmann. Gravitational Lensing. *Class. Quant. Grav.*, 27:233001, 2010.

[129] Volker Perlick. *Ray Optics, Fermat's Principle, and Applications to General Relativity*, volume 61. 2000.

[130] D. Munshi, P. Valageas, Ludovic Van Waerbeke, and A. Heavens. Cosmology with Weak Lensing Surveys. *Phys. Rept.*, 462:67–121, 2008.

[131] Henk Hoekstra. The effect of imperfect models of point spread function anisotropy on cosmic shear measurements. MNRAS, 347(4):1337–1344, February 2004.

[132] James Bosch, Robert Armstrong, Steven Bickerton, Hisanori Furusawa, Hiroyuki Ikeda, Michitaro Koike, Robert Lupton, Sogo Mineo, Paul Price, Tadafumi Takata, Masayuki Tanaka, Naoki Yasuda, Yusra AlSayyad, Andrew C. Becker, William Coulton, Jean Coupon, Jose Garmilla, Song Huang, K. Simon Krughoff, Dustin Lang, Alexie Leauthaud, Kian-Tat Lim, Nate B. Lust, Lauren A. MacArthur, Rachel Mandelbaum, Hironao Miyatake, Satoshi Miyazaki, Ryoma Murata, Surhud More, Yuki Okura, Russell Owen, John D. Swinbank, Michael A. Strauss, Yoshihiko Yamada, and Hitomi Yamanoi. The Hyper Suprime-Cam software pipeline. PASJ, 70:S5, January 2018.

[133] Daniel M. Jones and Alan F. Heavens. Bayesian photometric redshifts of blended sources. *Mon. Not. Roy. Astron. Soc.*, 483(2):2487–2505, 2019.

[134] Nora Elisa Chisari, Mark L. A. Richardson, Julien Devriendt, Yohan Dubois, Aurel Schneider, Amandine Le Brun, M. C., Ricarda S. Beckmann, Sebastien Peirani, Adrianne Slyz, and Christophe Pichon. The impact of baryons on the matter power spectrum from the Horizon-AGN cosmological hydrodynamical simulation. *Mon. Not. Roy. Astron. Soc.*, 480(3):3962–3977, 2018.

[135] James Sunseri, Zack Li, and Jia Liu. Effects of baryonic feedback on the cosmic web. *Phys. Rev. D*, 107(2):023514, 2023.

[136] T. M. C. Abbott et al. Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing. *Phys. Rev. D*, 98(4):043526, 2018.

[137] T. M. C. Abbott et al. Dark Energy Survey Year 3 results: Cosmological constraints from galaxy clustering and weak lensing. *Phys. Rev. D*, 105(2):023520, 2022.

[138] Giovanni Aricò, Raul E. Angulo, Carlos Hernández-Monteagudo, Sergio Contreras, and Matteo Zennaro. Simultaneous modelling of matter power spectrum and bispectrum

in the presence of baryons. MNRAS, 503(3):3596–3609, May 2021.

[139] Aurel Schneider and Romain Teyssier. A new method to quantify the effects of baryons on the matter power spectrum. *JCAP*, 12:049, 2015.

[140] Aurel Schneider, Romain Teyssier, Joachim Stadel, Nora Elisa Chisari, Amandine M. C. Le Brun, Adam Amara, and Alexandre Refregier. Quantifying baryon effects on the matter power spectrum and the weak lensing shear correlation. *JCAP*, 03:020, 2019.

[141] Max E. Lee, Tianhuan Lu, Zoltán Haiman, Jia Liu, and Ken Osato. Comparing weak lensing peak counts in baryonic correction models to hydrodynamical simulations. *Mon. Not. Roy. Astron. Soc.*, 519(1):573–584, 2023.

[142] R. E. Smith, J. A. Peacock, A. Jenkins, S. D. M. White, C. S. Frenk, F. R. Pearce, P. A. Thomas, G. Efstathiou, and H. M. P. Couchmann. Stable clustering, the halo model and nonlinear cosmological power spectra. *Mon. Not. Roy. Astron. Soc.*, 341:1311, 2003.

[143] Katrin Heitmann, David Higdon, Martin White, Salman Habib, Brian J. Williams, Earl Lawrence, and Christian Wagner. The Coyote Universe. II. Cosmological Models and Precision Emulation of the Nonlinear Matter Power Spectrum. ApJ, 705(1):156–174, November 2009.

[144] Ryuichi Takahashi, Masanori Sato, Takahiro Nishimichi, Atsushi Taruya, and Masamune Oguri. Revising the Halofit Model for the Nonlinear Matter Power Spectrum. *Astrophys. J.*, 761:152, 2012.

[145] Joachim Harnois-Déraps, Ludovic van Waerbeke, Massimo Viola, and Catherine Heymans. Baryons, Neutrinos, Feedback and Weak Gravitational Lensing. *Mon. Not. Roy. Astron. Soc.*, 450(2):1212–1223, 2015.

[146] V. Ajani et al. Euclid Preparation. XXVIII. Forecasts for ten different higher-order weak lensing statistics. *Astron. Astrophys.*, 675:A120, 2023.

[147] William R. Coulton, Jia Liu, Mathew S. Madhavacheril, Vanessa Böhm, and David N. Spergel. Constraining Neutrino Mass with the Tomographic Weak Lensing Bispectrum. *JCAP*, 05:043, 2019.

[148] Jia Liu, Andrea Petri, Zoltán Haiman, Lam Hui, Jan M. Kratochvil, and Morgan May. Cosmology constraints from the weak lensing peak counts and the power spectrum in CFHTLenS data. *Phys. Rev. D*, 91(6):063507, 2015.

[149] Zack Li, Jia Liu, José Manuel Zorrilla Matilla, and William R. Coulton. Constraining neutrino mass with tomographic weak lensing peak counts. *Phys. Rev. D*, 99(6):063527, 2019.

[150] Aoife Boyle, Cora Uhlemann, Oliver Friedrich, Alexandre Barthelemy, Sandrine Codis, Francis Bernardeau, Carlo Giocoli, and Marco Baldi. Nuw CDM cosmology from the weak-lensing convergence PDF. *Mon. Not. Roy. Astron. Soc.*, 505(2):2886–2902, 2021.

[151] Jia Liu and Mathew S. Madhavacheril. Constraining neutrino mass with the tomographic weak lensing one-point probability distribution function and power spectrum. *Phys. Rev. D*, 99(8):083508, 2019.

[152] Leander Thiele, J. Colin Hill, and Kendrick M. Smith. Accurate analytic model for the weak lensing convergence one-point probability distribution function and its autocovariance. *Phys. Rev. D*, 102(12):123545, 2020.

[153] Gabriela A. Marques, Jia Liu, José Manuel Zorrilla Matilla, Zoltán Haiman, Armando Bernui, and Camila P. Novaes. Constraining neutrino mass with weak lensing Minkowski Functionals. *JCAP*, 06:019, 2019.

[154] P. Schneider and M. Lombardi. The three-point correlation function of cosmic shear. I. The natural components. A&A, 397:809–818, January 2003.

[155] Masahiro Takada and Bhuvnesh Jain. Three-point correlations in weak lensing surveys: Model predictions and applications. *Mon. Not. Roy. Astron. Soc.*, 344:857, 2003.

[156] Masahiro Takada and Bhuvnesh Jain. Cosmological parameters from lensing power spectrum and bispectrum tomography. *Mon. Not. Roy. Astron. Soc.*, 348:897, 2004.

[157] Joachim Harnois-Déraps, Nicolas Martinet, Tiago Castro, Klaus Dolag, Benjamin Giblin, Catherine Heymans, Hendrik Hildebrandt, and Qianli Xia. Cosmic shear cosmology beyond two-point statistics: a combined peak count and correlation function analysis of DES-Y1. *Monthly Notices of the Royal Astronomical Society*, 506(2):1623–1650, 06 2021.

[158] D. Zürcher et al. Dark energy survey year 3 results: Cosmology with peaks using an emulator approach. *Mon. Not. Roy. Astron. Soc.*, 511(2):2075–2104, 2022.

[159] Jia Liu, Zoltán Haiman, Lam Hui, Jan M. Kratochvil, and Morgan May. The Impact of Magnification and Size Bias on Weak Lensing Power Spectrum and Peak Statistics. *Phys. Rev. D*, 89(2):023515, 2014.

[160] Leander Thiele, Gabriela A Marques, Jia Liu, and Masato Shirasaki. Cosmological constraints from HSC Y1 lensing convergence PDF. *arXiv preprint arXiv:2304.05928*, 2023.

[161] Jens Schmalzing, Martin Kerscher, and Thomas Buchert. Minkowski functionals in cosmology. *Proc. Int. Sch. Phys. Fermi*, 132:281–291, 1996.

[162] Jan M. Kratochvil, Eugene A. Lim, Sheng Wang, Zoltán Haiman, Morgan May, and Kevin Huffenberger. Probing cosmology with weak lensing Minkowski functionals. Phys. Rev. D, 85(10):103513, May 2012.

[163] Euclid Collaboration, A. Blanchard, S. Camera, C. Carbone, V. F. Cardone, S. Casas, S. Ilić, M. Kilbinger, T. Kitching, M. Kunz, F. Lacasa, E. Linder, E. Majerotto, K. Markovič, M. Martinelli, V. Pettorino, et al. Euclid preparation: VII. Forecast validation for Euclid cosmological probes. *arXiv e-prints*, page arXiv:1910.09273, Oct

2019.

[164] Garrelt Mellema, Léon V. E. Koopmans, Filipe A. Abdalla, Gianni Bernardi, Benedetta Ciardi, Soobash Daiboo, A. G. de Bruyn, Kanan K. Datta, Heino Falcke, Andrea Ferrara, Ilian T. Iliev, Fabio Iocco, Vibor Jelić, Hannes Jensen, Ronniy Joseph, Panos Labroupoulos, Avery Meiksin, Andrei Mesinger, André R. Offringa, V. N. Pandey, Jonathan R. Pritchard, Mario G. Santos, Dominik J. Schwarz, Benoit Semelin, Harish Vedantham, Sarod Yatawatta, and Saleem Zaroubi. Reionization and the Cosmic Dawn with the Square Kilometre Array. *Experimental Astronomy*, 36(1-2):235–318, August 2013.

[165] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020.

[166] Jasper Albers, Christian Fidler, Julien Lesgourgues, Nils Schöneberg, and Jesus Torrado. CosmicNet. Part I. Physics-driven implementation of neural networks within Einstein-Boltzmann Solvers. *JCAP*, 09:028, 2019.

[167] J. Bayron Orjuela-Quintana, Savvas Nesseris, and Domenico Sapone. Machine learning unveils the linear matter power spectrum of modified gravity. *arXiv preprint arXiv:2307.03643*, 2023.

[168] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. *Publications of the Astronomical Society of the Pacific*, 125(925):306–312, Mar 2013.

[169] Antony Lewis. GetDist: a Python package for analysing Monte Carlo samples. *arXiv preprint arXiv:1910.13970*, 2019.

[170] Ken Osato, Jia Liu, and Zoltán Haiman. κTNG: effect of baryonic processes on weak lensing with IllustrisTNG simulations. *Mon. Not. Roy. Astron. Soc.*, 502(4):5593–5602, 2021.

[171] Hiroaki Aihara, Nobuo Arimoto, Robert Armstrong, Stéphane Arnouts, Neta A. Bahcall, Steven Bickerton, James Bosch, Kevin Bundy, Peter L. Capak, James H. H. Chan, Masashi Chiba, Jean Coupon, Eiichi Egami, Motohiro Enoki, Francois Finet, Hiroki Fujimori, Seiji Fujimoto, Hisanori Furusawa, Junko Furusawa, Tomotsugu Goto, Andy Goulding, Johnny P. Greco, Jenny E. Greene, James E. Gunn, Takashi Hamana, Yuichi Harikane, Yasuhiro Hashimoto, Takashi Hattori, Masao Hayashi, Yusuke Hayashi, Krzysztof G. Hełminiak, Ryo Higuchi, Chiaki Hikage, Paul T. P. Ho, Bau-Ching Hsieh, Kuiyun Huang, Song Huang, Hiroyuki Ikeda, Masatoshi Imanishi, Akio K. Inoue, Kazushi Iwasawa, Ikuru Iwata, Anton T. Jaelani, Hung-Yu Jian, Yukiko Kamata, Hiroshi Karoji, Nobunari Kashikawa, Nobuhiko Katayama, Satoshi Kawanomoto, Issha Kayo, Jin Koda, Michitaro Koike, Takashi Kojima, Yutaka Komiyama, Akira Konno, Shintaro Koshida, Yusei Koyama, Haruka Kusakabe, Alexie Leauthaud, Chien-Hsiu Lee, Lihwai Lin, Yen-Ting Lin, Robert H. Lupton, Rachel Mandelbaum, Yoshiki Matsuoka, Elinor Medezinski, Sogo Mineo, Shoken Miyama, Hironao Miyatake, Satoshi Miyazaki, Rieko Momose, Anupreeta More, Surhud More, Yuki Moritani, Takashi J. Moriya, Tomoki Morokuma, Shiro Mukae, Ryoma Murata, Hitoshi Murayama, Tohru

Nagao, Fumiaki Nakata, Mana Niida, Hiroko Niikura, Atsushi J. Nishizawa, Yoshiyuki Obuchi, Masamune Oguri, Yukie Oishi, Nobuhiro Okabe, Sakurako Okamoto, Yuki Okura, Yoshiaki Ono, Masato Onodera, Masafusa Onoue, Ken Osato, Masami Ouchi, Paul A. Price, Tae-Soo Pyo, Masao Sako, Marcin Sawicki, Takatoshi Shibuya, Kazuhiro Shimasaku, Atsushi Shimono, Masato Shirasaki, John D. Silverman, Melanie Simet, Joshua Speagle, David N. Spergel, Michael A. Strauss, Yuma Sugahara, Naoshi Sugiyama, Yasushi Suto, Sherry H. Suyu, Nao Suzuki, Philip J. Tait, Masahiro Takada, Tadafumi Takata, Naoyuki Tamura, Manobu M. Tanaka, Masaomi Tanaka, Masayuki Tanaka, Yoko Tanaka, Tsuyoshi Terai, Yuichi Terashima, Yoshiki Toba, Nozomu Tominaga, Jun Toshikawa, Edwin L. Turner, Tomohisa Uchida, Hisakazu Uchiyama, Keiichi Umetsu, Fumihiro Uraguchi, Yuji Urata, Tomonori Usuda, Yousuke Utsumi, Shiang-Yu Wang, Wei-Hao Wang, Kenneth C. Wong, Kiyoto Yabe, Yoshihiko Yamada, Hitomi Yamanoi, Naoki Yasuda, Sherry Yeh, Atsunori Yonehara, and Suraphong Yuma. The Hyper Suprime-Cam SSP Survey: Overview and survey design. PASJ, 70:S4, January 2018.

[172] Rachel Mandelbaum, Hironao Miyatake, Takashi Hamana, Masamune Oguri, Melanie Simet, Robert Armstrong, James Bosch, Ryoma Murata, François Lanusse, Alexie Leauthaud, Jean Coupon, Surhud More, Masahiro Takada, Satoshi Miyazaki, Joshua S. Speagle, Masato Shirasaki, Cristóbal Sifón, Song Huang, Atsushi J. Nishizawa, Elinor Medezinski, Yuki Okura, Nobuhiro Okabe, Nicole Czakon, Ryuichi Takahashi, William R. Coulton, Chiaki Hikage, Yutaka Komiyama, Robert H. Lupton, Michael A. Strauss, Masayuki Tanaka, and Yousuke Utsumi. The first-year shear catalog of the Subaru Hyper Suprime-Cam Subaru Strategic Program Survey. PASJ, 70:S25, January 2018.

[173] Satoshi Kawanomoto, Fumihiro Uraguchi, Yutaka Komiyama, Satoshi Miyazaki, Hisanori Furusawa, François Finet, Takashi Hattori, Shiang-Yu Wang, Naoki Yasuda, and Naotaka Suzuki. Hyper Suprime-Cam: Filters. *Publications of the Astronomical Society of Japan*, 70(4):66, 06 2018.

[174] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. ApJS, 208(2):19, oct 2013.

[175] Ryuichi Takahashi, Takashi Hamana, Masato Shirasaki, Toshiya Namikawa, Takahiro Nishimichi, Ken Osato, and Kosei Shiroyama. Full-sky Gravitational Lensing Simulation for Large-area Galaxy Surveys and Cosmic Microwave Background Experiments. *Astrophys. J.*, 850(1):24, 2017.

[176] Masato Shirasaki, Kana Moriwaki, Taira Oogi, Naoki Yoshida, Shiro Ikeda, and Takahiro Nishimichi. Noise reduction for weak lensing mass mapping: an application of generative adversarial networks to Subaru Hyper Suprime-Cam first-year data. *Mon. Not. Roy. Astron. Soc.*, 504(2):1825–1839, 2021.

[177] Masato Shirasaki, Kana Moriwaki, Taira Oogi, Naoki Yoshida, Shiro Ikeda, and

Takahiro Nishimichi. Noise reduction for weak lensing mass mapping: an application of generative adversarial networks to subaru hyper suprime-cam first-year data. *Monthly Notices of the Royal Astronomical Society*, 504(2):1825–1839, 2021.

[178] Ian G. McCarthy, Joop Schaye, Simeon Bird, and Amandine M. C. Le Brun. The BAHAMAS project: Calibrated hydrodynamical simulations for large-scale structure cosmology. *Mon. Not. Roy. Astron. Soc.*, 465(3):2936–2965, 2017.

[179] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. ApJS, 208(2):19, October 2013.

[180] Ian G. Mccarthy, Simeon Bird, Joop Schaye, Joachim Harnois-Deraps, Andreea S. Font, and Ludovic Van Waerbeke. The BAHAMAS project: the CMB–large-scale structure tension and the roles of massive neutrinos and galaxy formation. *Mon. Not. Roy. Astron. Soc.*, 476(3):2999–3030, 2018.

[181] Mark Vogelsberger, Shy Genel, Volker Springel, Paul Torrey, Debora Sijacki, Dandan Xu, Gregory F. Snyder, Dylan Nelson, and Lars Hernquist. Introducing the Illustris Project: Simulating the coevolution of dark and visible matter in the Universe. *Mon. Not. Roy. Astron. Soc.*, 444(2):1518–1547, 2014.

[182] Dylan Nelson et al. The IllustrisTNG Simulations: Public Data Release. *arXiv preprint arXiv:1812.05609*, 2018.

[183] Planck Collaboration. Planck 2015 results. XIII. Cosmological parameters. A&A, 594:A13, September 2016.

[184] Annalisa Pillepich et al. First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies. *Mon. Not. Roy. Astron. Soc.*, 475(1):648–675, 2018.

[185] Andreas J. Weiss, Aurel Schneider, Raphael Sgier, Tomasz Kacprzak, Adam Amara, and Alexandre Refregier. Effects of baryons on weak lensing peak statistics. *JCAP*, 10:011, 2019.

[186] A. Petri. Mocking the weak lensing universe: The lenstools python computing package. *Astronomy and Computing*, 17:73–79, 2016.

[187] HuanYuan Shan et al. KiDS-450: cosmological constraints from weak lensing peak statistics – I. Inference from analytical prediction of high signal-to-noise ratio convergence peaks. *Mon. Not. Roy. Astron. Soc.*, 474(1):1116–1134, 2018.

[188] Virginia Ajani, Austin Peel, Valeria Pettorino, Jean-Luc Starck, Zack Li, and Jia Liu. Constraining neutrino masses with weak-lensing multiscale peak counts. *Phys. Rev. D*, 102(10):103531, 2020.

[189] Eric Hivon, Krzysztof M Górski, C Barth Netterfield, Brendan P Crill, Simon Prunet, and Frode Hansen. Master of the cosmic microwave background anisotropy power spectrum: a fast method for statistical analysis of large and complex cosmic microwave background data sets. *The Astrophysical Journal*, 567(1):2, 2002.

[190] David Alonso, Javier Sanchez, and Anže Slosar. A unified pseudo-C_ℓ framework. *Mon. Not. Roy. Astron. Soc.*, 484(3):4127–4151, 2019.

[191] Masamune Oguri, Satoshi Miyazaki, Chiaki Hikage, Rachel Mandelbaum, Yousuke Utsumi, Hironao Miyatake, Masahiro Takada, Robert Armstrong, James Bosch, Yutaka Komiyama, et al. Two-and three-dimensional wide-field weak lensing mass maps from the hyper suprime-cam subaru strategic program s16a data. *Publications of the Astronomical Society of Japan*, 70(SP1):S26, 2018.

[192] Jia Liu and Zoltan Haiman. Origin of weak lensing convergence peaks. *Phys. Rev. D*, 94(4):043533, 2016.

[193] Jesús Torrado and Antony Lewis. Cobaya: code for bayesian analysis of hierarchical physical models. *Journal of Cosmology and Astroparticle Physics*, 2021(05):057, may 2021.

[194] Jesus Torrado and Antony Lewis. Cobaya: Code for Bayesian Analysis of hierarchical physical models. *JCAP*, 05:057, 2021.

[195] J. Hartlap, Patrick Simon, and P. Schneider. Why your model parameter confidences might be too optimistic: Unbiased estimation of the inverse covariance matrix. *Astron. Astrophys.*, 464:399, 2007.

[196] Xiangchong Li et al. Hyper Suprime-Cam Year 3 Results: Cosmology from Cosmic Shear Two-point Correlation Functions. *arXiv preprint arXiv:2304.00702*, 2023.

[197] Hung-Jin Huang, Tim Eifler, Rachel Mandelbaum, Gary M Bernstein, Anqi Chen, Ami Choi, Juan García-Bellido, Dragan Huterer, Elisabeth Krause, Eduardo Rozo, et al. Dark energy survey year 1 results: Constraining baryonic physics in the universe. *Monthly Notices of the Royal Astronomical Society*, 502(4):6010–6031, 2021.

[198] Joachim Harnois-Déraps, Nicolas Martinet, and Robert Reischke. Cosmic shear beyond 2-point statistics: Accounting for galaxy intrinsic alignment with projected tidal fields. *Monthly Notices of the Royal Astronomical Society*, 509(3):3868–3888, 2022.

[199] Michel Chevallier and David Polarski. *Int. J. Mod. Phys. D*, 10:213–224, 2001.

[200] Irit Maor, Ram Brustein, and Paul J. Steinhardt. Limitations in using luminosity distance to determine the equation of state of the universe. *Physical Review Letters*, 86(1):6–9, Jan 2001.

[201] Max Tegmark. Measuring the metric: A parametrized post-friedmannian approach to the cosmic dark energy problem. *Physical Review D*, 66(10), Nov 2002.

[202] Dragan Huterer and Glenn Starkman. *Phys. Rev. Lett.*, 90:031301, 2003.

[203] Yun Wang and Geoffrey Lovelace. Unbiased estimate of dark energy density from type IA supernova data. *Astrophys. J. Lett.*, 562:L115–L120, 2001.

[204] Yun Wang and Peter M. Garnavich. Measuring time dependence of dark energy density from type Ia supernova data. *Astrophys. J.*, 552:445, 2001.

[205] Marina Seikel, Chris Clarkson, and Mathew Smith. Reconstruction of dark energy and expansion dynamics using Gaussian processes. *JCAP*, 2012(6):036, 2012.

[206] Marina Seikel and Chris Clarkson. Optimising Gaussian processes for reconstructing dark energy dynamics from supernovae. *arXiv preprint arXiv:1311.6678*, 2013.

[207] Arman Shafieloo, Alex G. Kim, and Eric V. Linder. Gaussian Process Cosmography. *Phys. Rev. D*, 85:123530, 2012.

[208] Yun Wang and Pia Mukherjee. Model-independent constraints on dark energy density from flux-averaging analysis of type ia supernova data. *The Astrophysical Journal*, 606(2):654–663, May 2004.

[209] Yun Wang and Katherine Freese. Probing dark energy using its density instead of its equation of state. *Phys. Lett. B*, 632:449–452, 2006.

[210] Victor H. Cardenas. Exploring hints for dark energy density evolution in light of recent data. *Phys. Lett. B*, 750:128–134, 2015.

[211] Özgür Akarsu, Suresh Kumar, Emre Özülker, and J. Alberto Vazquez. Relaxing cosmological tensions with a sign switching cosmological constant. *Phys. Rev. D*, 104(12):123512, 2021.

[212] Özgür Akarsu, John D. Barrow, Luis A. Escamilla, and J. Alberto Vazquez. Graduated dark energy: Observational hints of a spontaneous sign switch in the cosmological constant. *Phys. Rev. D*, 101(6):063528, 2020.

[213] Celia Escamilla-Rivera, Jackson Levi Said, and Jurgen Mifsud. Performance of non-parametric reconstruction techniques in the late-time universe. *JCAP*, 10:016, 2021.

[214] Michele Moresco, Lucia Pozzetti, Andrea Cimatti, Raul Jimenez, Claudia Maraston, Licia Verde, Daniel Thomas, Annalisa Citro, Rita Tojeiro, and David Wilkinson. A 6% measurement of the Hubble parameter at $z \sim 0.45$: direct evidence of the epoch of cosmic re-acceleration. *JCAP*, 05:014, 2016.

[215] Michele Moresco. Raising the bar: new constraints on the Hubble parameter with cosmic chronometers at z $\sim$ 2. *Mon. Not. Roy. Astron. Soc.*, 450(1):L16–L20, 2015.

[216] Cong Zhang, Han Zhang, Shuo Yuan, Siqi Liu, Tong-Jie Zhang, and Yan-Chun Sun. Four new observational H(z) data from luminous red galaxies in the Sloan Digital Sky Survey data release seven. *Research in Astronomy and Astrophysics*, 14(10):1221–1233,

2014.

[217] Daniel Stern, Raul Jimenez, Licia Verde, Marc Kamionkowski, and S. Adam Stanford. Cosmic chronometers: constraining the equation of state of dark energy. I: H(z) measurements. *JCAP*, 2010(2):008, 2010.

[218] M. Moresco et al. Improved constraints on the expansion rate of the Universe up to z ~1.1 from the spectroscopic evolution of cosmic chronometers. *JCAP*, 2012(8):006, 2012.

[219] A. L. Ratsimbazafy, S. I. Loubser, S. M. Crawford, C. M. Cress, B. A. Bassett, R. C. Nichol, and P. Väisänen. Age-dating Luminous Red Galaxies observed with the Southern African Large Telescope. *Mon. Not. Roy. Astron. Soc.*, 467(3):3239–3254, 2017.

[220] Timothée Delubac et al. Baryon acoustic oscillations in the Lyα forest of BOSS DR11 quasars. *Astron. Astrophys.*, 574:A59, 2015.

[221] Chris Blake, Sarah Brough, Matthew Colless, Carlos Contreras, Warrick Couch, Scott Croom, Darren Croton, Tamara M. Davis, Michael J. Drinkwater, Karl Forster, David Gilbank, Mike Gladders, Karl Glazebrook, Ben Jelliffe, Russell J. Jurek, I. hui Li, Barry Madore, D. Christopher Martin, Kevin Pimbblet, Gregory B. Poole, Michael Pracy, Rob Sharp, Emily Wisnioski, David Woods, Ted K. Wyder, and H. K. C. Yee. The WiggleZ Dark Energy Survey: joint measurements of the expansion and growth history at z < 1. *Mon. Not. Roy. Astron. Soc.*, 425(1):405–414, September 2012.

[222] Chia-Hsun Chuang et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: single-probe measurements and the strong power of normalized growth rate on constraining dark energy. *Mon. Not. Roy. Astron. Soc.*, 433:3559, 2013.

[223] Andreu Font-Ribera et al. Quasar-Lyman α Forest Cross-Correlation from BOSS DR11 : Baryon Acoustic Oscillations. *JCAP*, 05:027, 2014.

[224] Julian E. Bautista et al. Measurement of baryon acoustic oscillation correlations at $z = 2.3$ with SDSS DR12 Lyα-Forests. *Astron. Astrophys.*, 603:A12, 2017.

[225] Enrique Gaztanaga, Anna Cabre, and Lam Hui. Clustering of Luminous Red Galaxies IV: Baryon Acoustic Peak in the Line-of-Sight Direction and a Direct Measurement of H(z). *Mon. Not. Roy. Astron. Soc.*, 399:1663–1680, 2009.

[226] Akira Oka, Shun Saito, Takahiro Nishimichi, Atsushi Taruya, and Kazuhiro Yamamoto. Simultaneous constraints on the growth of structure and cosmic expansion from the multipole power spectra of the SDSS DR7 LRG sample. *Mon. Not. Roy. Astron. Soc.*, 439:2515–2530, 2014.

[227] Yuting Wang et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: tomographic BAO analysis of DR12 combined sample in configuration space. *Mon. Not. Roy. Astron. Soc.*, 469(3):3762–3774, 2017.

[228] Chia-Hsun Chuang and Yun Wang. Modeling the Anisotropic Two-Point Galaxy Cor-

relation Function on Small Scales and Improved Measurements of $H(z)$, $D_A(z)$, and $\beta(z)$ from the Sloan Digital Sky Survey DR7 Luminous Red Galaxies. *Mon. Not. Roy. Astron. Soc.*, 435:255–262, 2013.

[229] Shadab Alam et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. *Mon. Not. Roy. Astron. Soc.*, 470(3):2617–2652, 2017.

[230] Lauren Anderson et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples. *Mon. Not. Roy. Astron. Soc.*, 441(1):24–62, 2014.

[231] N. G. Busca et al. Baryon acoustic oscillations in the Lyα forest of BOSS quasars. *A&A*, 552:A96, April 2013.

[232] D. M. Scolnic et al. The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample. *Astrophys. J.*, 859(2):101, 2018.

[233] Adam G. Riess et al. Type Ia Supernova Distances at Redshift > 1.5 from the Hubble Space Telescope Multi-cycle Treasury Programs: The Early Expansion Rate. *Astrophys. J.*, 853(2):126, 2018.

[234] Adrià Gómez-Valent and Luca Amendola. H_0 from cosmic chronometers and Type Ia supernovae, with Gaussian Processes and the novel Weighted Polynomial Regression method. *JCAP*, 04:051, 2018.

[235] Rebecca Briffa, Salvatore Capozziello, Jackson Levi Said, Jurgen Mifsud, and Emmanuel N. Saridakis. Constraining teleparallel gravity through Gaussian processes. *Class. Quant. Grav.*, 38(5):055007, 2020.

[236] A. Conley et al. Supernova Constraints and Systematic Uncertainties from the First 3 Years of the Supernova Legacy Survey. *Astrophys. J. Suppl.*, 192:1, 2011.

[237] Wendy L. Freedman et al. The Carnegie-Chicago Hubble Program. VIII. An Independent Determination of the Hubble Constant Based on the Tip of the Red Giant Branch. *Astrophys. J.*, 882(1):34, 2019.

[238] Gagandeep S. Anand, R. Brent Tully, Luca Rizzi, Adam G. Riess, and Wenlong Yuan. Comparing Tip of the Red Giant Branch Distance Scales: An Independent Reduction of the Carnegie-Chicago Hubble Program and the Value of the Hubble Constant. *Astrophys. J.*, 932(1):15, 2022.

[239] Adam G. Riess et al. A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km s^1 Mpc^1 Uncertainty from the Hubble Space Telescope and the SH0ES Team. *Astrophys. J. Lett.*, 934(1):L7, 2022.

[240] Jonathan Goodman and Jonathan Weare. *Communications in Applied Mathematics and Computational Science*, 5(1):65–80, January 2010.

[241] Eoin Ó Colgáin and M. M. Sheikh-Jabbari. Elucidating cosmological model dependence with H_0. *Eur. Phys. J. C*, 81(10):892, 2021.

[242] Yuting Wang, Levon Pogosian, Gong-Bo Zhao, and Alex Zucca. Evolution of dark energy reconstructed from the latest observations. *Astrophys. J. Lett.*, 869:L8, 2018.

[243] Luis A. Escamilla and J. Alberto Vazquez. Model selection applied to reconstructions of the Dark Energy. *Eur. Phys. J. C*, 83(3):251, 2023.

[244] Carlos Bengaly. A null test of the Cosmological Principle with BAO measurements. *Phys. Dark Univ.*, 35:100966, 2022.

[245] Eoin Ó. Colgáin, M. M. Sheikh-Jabbari, and Lu Yin. Can dark energy be dynamical? *Phys. Rev. D*, 104(2):023510, 2021.

[246] Reginald Christian Bernardo and Jackson Levi Said. Towards a model-independent reconstruction approach for late-time Hubble data. *JCAP*, 08:027, 2021.

[247] J. Alberto Vazquez, M. Bridges, M. P. Hobson, and A. N. Lasenby. Reconstruction of the Dark Energy equation of state. *JCAP*, 09:020, 2012.

[248] Gong-Bo Zhao et al. Dynamical dark energy in light of the latest observations. *Nature Astron.*, 1(9):627–632, 2017.

[249] Yuewei Wen, Eva Nesbit, Dragan Huterer, and Scott Watson. Misinterpreting modified gravity as dark energy: a quantitative study. *JCAP*, 03(03):042, 2022.

[250] Levon Pogosian, Marco Raveri, Kazuya Koyama, Matteo Martinelli, Alessandra Silvestri, Gong-Bo Zhao, Jian Li, Simone Peirone, and Alex Zucca. Imprints of cosmological tensions in reconstructed gravity. *Nature Astron.*, 6(12):1484–1490, 2022.

[251] Marco Raveri, Levon Pogosian, Matteo Martinelli, Kazuya Koyama, Alessandra Silvestri, Gong-Bo Zhao, Jian Li, Simone Peirone, and Alex Zucca. Principal reconstructed modes of dark energy and gravity. *JCAP*, 02:061, 2023.

[252] Daniela Grandon and Victor H. Cardenas. Exploring evidence of interaction between dark energy and dark matter. *Gen Relativ Gravit*, 51, 2019.

Annex A

The impact of baryons for larger smoothing scales

Similar to Fig. 5.6, we show the impact of baryonic feedback on the fractional difference for the smoothing scales of 5 arcmin in Fig. A.1 and 8 arcmin in Fig. A.2.

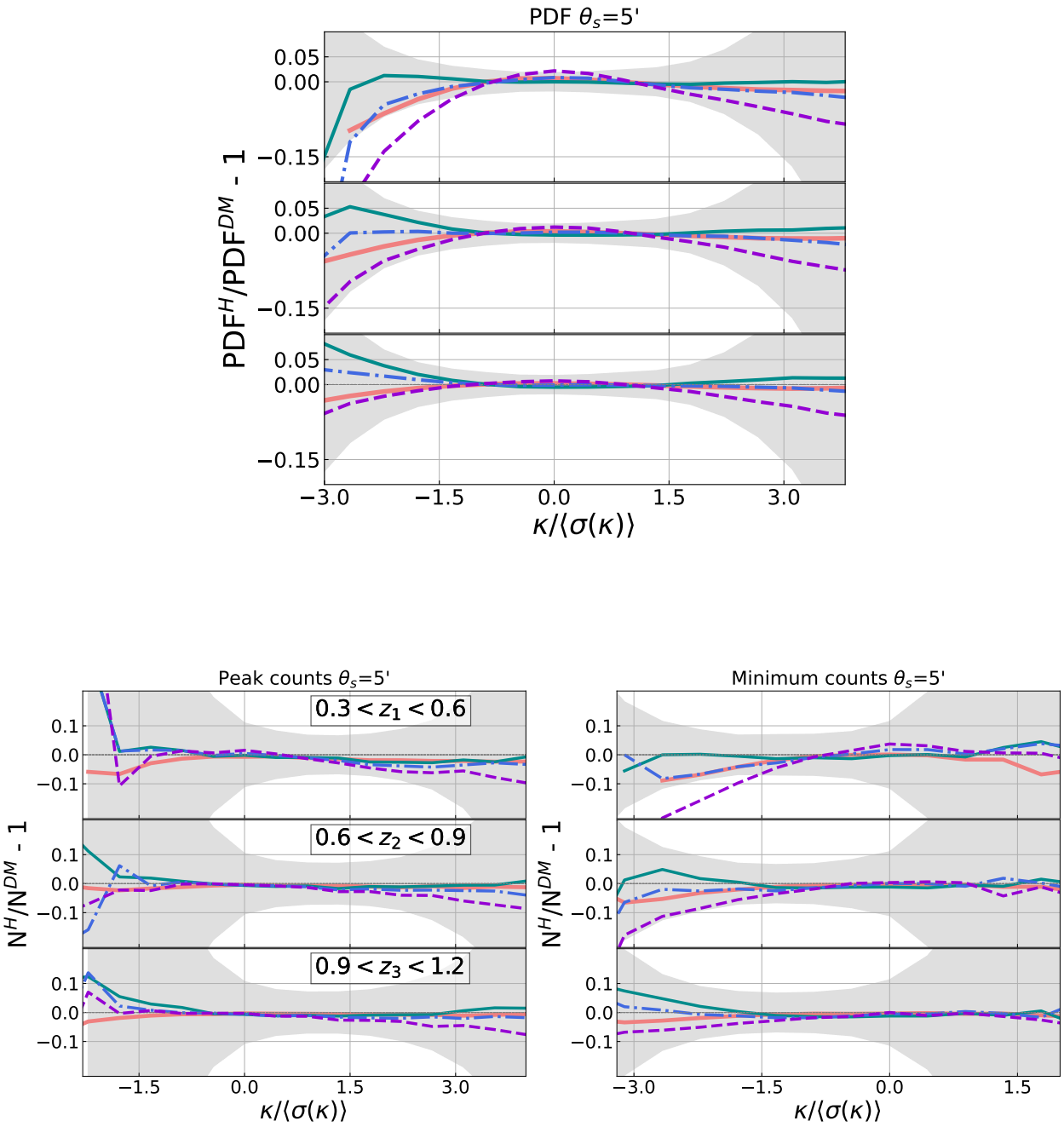


Figure A.1: The impact of baryonic feedback on the power spectrum, PDF, peak counts, and minimum counts for convergence maps smoothed with Gaussian smoothing kernel of $\theta_s = 5$ arcmin. We show the results for the hydrodynamic simulations BAHAMAS high-AGN (**dashed purple line**), fiducial-AGN (**dashed-dot blue line**), low-AGN (**solid teal line**) and κTNG (**solid pink line**). Vertical axis corresponds to fractional difference, where C_κ^H is the baryon-contaminated angular power spectrum and C_{DM} the angular power spectrum measured in the dark matter only runs. The grey-shaded region corresponds to HSC-Y1 uncertainty.

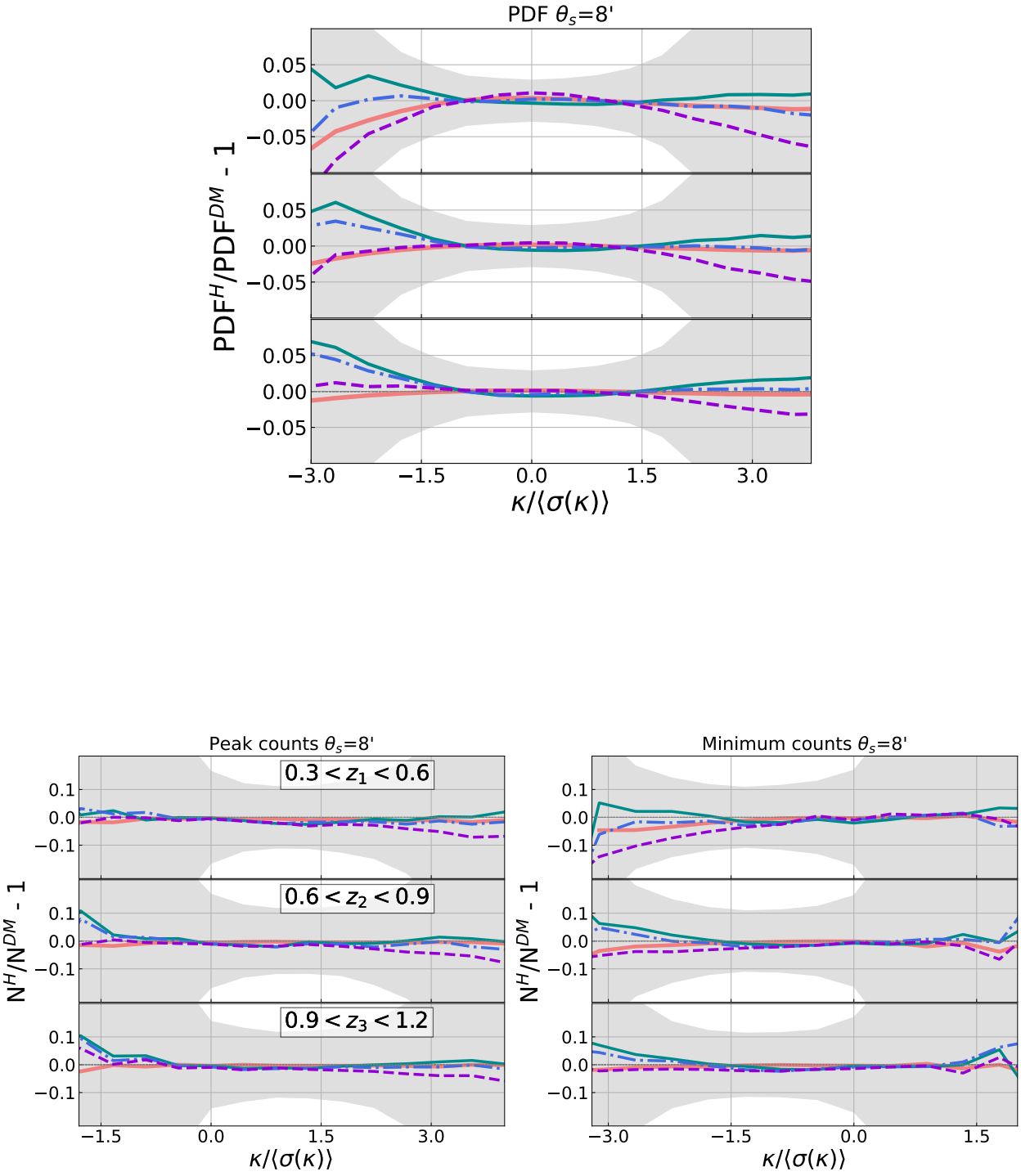


Figure A.2: The same as Fig. A.1, but for maps smoothed with Gaussian smoothing kernel of $\theta_s = 8$ arcmin.

Annex B

Table of H(z) data

We present the Hubble expansion data used in this paper in Table B.1.

Table B.1: Expansion data from cosmic chronometers (31 points, left column) and baryon acoustic oscillations (26 points, right column).

z	H [km s^{-1} Mpc^{-1}]	Ref.	z	H [km s^{-1} Mpc^{-1}]	Ref.
0.07	69 ± 19.6	[216]	0.24	79.69 ± 2.99	[225]
0.09	69 ± 12	[217]	0.3	81.7 ± 6.22	[226]
0.12	68.6 ± 26.2	[216]	0.31	78.18 ± 4.74	[227]
0.17	83 ± 8	[217]	0.34	83.8 ± 3.66	[225]
0.1791	75 ± 4	[218]	0.35	82.7 ± 9.1	[228]
0.1993	75 ± 5	[218]	0.36	79.94 ± 3.38	[227]
0.2	72.9 ± 29.6	[216]	0.38	81.5 ± 1.9	[229]
0.27	77 ± 14	[217]	0.4	82.04 ± 2.03	[227]
0.28	88.8 ± 36.6	[216]	0.43	86.45 ± 3.97	[225]
0.3519	83 ± 14	[218]	0.44	84.81 ± 1.83	[227]
0.3802	83 ± 13.5	[214]	0.44	82.6 ± 7.8	[221]
0.4	95 ± 17	[217]	0.48	87.79 ± 2.03	[227]
0.4004	77 ± 10.2	[214]	0.51	90.4 ± 1.9	[229]
0.4247	87.1 ± 11.2	[214]	0.52	94.35 ± 2.64	[227]
0.4497	92.8 ± 12.9	[214]	0.56	93.34 ± 2.3	[227]
0.47	89 ± 34	[219]	0.57	87.6 ± 7.8	[222]
0.4783	80.9 ± 9	[214]	0.57	96.8 ± 3.4	[230]
0.48	97 ± 62	[217]	0.59	98.48 ± 3.18	[227]
0.5929	104 ± 13	[218]	0.6	87.9 ± 6.1	[221]
0.6797	92 ± 8	[218]	0.61	97.3 ± 2.1	[229]
0.7812	105 ± 12	[218]	0.64	98.82 ± 2.98	[227]
0.8754	125 ± 17	[218]	0.73	97.3 ± 7.0	[221]
0.88	90 ± 40	[217]	2.3	224 ± 8.6	[231]
0.9	117 ± 23	[217]	2.33	224 ± 8	[224]
1.037	154 ± 20	[218]	2.34	222 ± 8.5	[220]
1.3	168 ± 17	[217]	2.36	226 ± 9.3	[223]
1.363	160 ± 33.6	[215]			
1.43	177 ± 18	[217]			
1.53	140 ± 14	[217]			
1.75	202 ± 40	[217]			
1.965	186.5 ± 50.4	[215]			