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

DATA MINING AND MACHINE LEARNING ALGORITHMS IN COSMOLOGY TO OPTIMIZE THE CONSTRAINING POWER OF PARAMETERS BY BAYESIAN INFERENCE

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

JAVIER IGNACIO SILVA LAFAURIE

PROFESOR GUÍA:
DOMENICO SAPONE

MIEMBROS DE LA COMISIÓN:
GINEVRA FAVOLE
FRANCISCO FÖRSTER
GONZALO PALMA QUILODRÁN

Este trabajo ha sido parcialmente financiado por FONDECYT Iniciación N. 11140496 y FONDECYT Regular N. 1200171

SANTIAGO DE CHILE
2020

RESUMEN DE LA TESIS PARA OPTAR
AL GRADO DE MAGÍSTER EN CIENCIAS, MENCIÓN FÍSICA
POR: JAVIER IGNACIO SILVA LAFAURIE
FECHA: 2020
PROF. GUÍA: DOMENICO SAPONE

DATA MINING AND MACHINE LEARNING ALGORITHMS IN COSMOLOGY TO OPTIMIZE THE CONSTRAINING POWER OF PARAMETERS BY BAYESIAN INFERENCE

El objetivo de esta tesis es explorar algunos algoritmos de minería de datos y aprendizaje de máquinas aplicados a datos cosmológicos, especialmente a los observables obtenidos del agrupamiento de galaxias en la estructura a gran escala de nuestro Universo. Revisamos como obtener esos observables para el catálogo "New York University Value-Added Galaxy Catalog", del cual reducimos aproximadamente 600000^3 coordenadas de galaxias a tan solo 45^3 números. Luego, procedemos a constreñirlos, para adquirir información cosmológica, usando el modelo ΛCDM e inferencia Bayesiana. Además, obtenemos los resultados usando dos algoritmos: Metropolis-Hastings y DELFI; que están basados en cadenas de Marcov y redes neuronales, respectivamente. Donde el último puede ser más eficiente que el primero bajo una arquitectura apropiada. Concluimos que los resultados obtenidos de ambos métodos son consistentes, pero los errores en los parámetros de *Shift* y *Redshift Space Distortions* son subestimados en comparación con otros trabajos que usaron datos similares.

También usamos las medidas de $f\sigma_8(z)$ de diferentes catálogos junto a medidas de $H(z)$, obtenidas de los cronómetros cósmicos, para comparar diferentes modelos de Energía Oscura. Utilizamos diferentes criterios estadísticos, como la evidencia Bayesiana, el criterio de información Bayesiano, el criterio de información de Akaike y la figura de mérito. Finalmente, contrastamos esos criterios para todos los modelos y concluimos que el modelo ΛCDM es siempre el favorecido con estos datos.

Además, exploramos los observables obtenidos del agrupamiento de galaxias del catálogo "SDSS DR7 main galaxy sample", usando diferentes tipos de galaxias como trazadores. El uso de múltiples trazadores puede inducir un sesgo en la función de correlación, que se refleja en su amplitud y en el pico de las oscilaciones acústicas de bariones. Sin embargo, como estas galaxias trazan la misma distribución de Materia Oscura, esperamos que el sesgo sea pequeño y que el uso conjunto de los múltiples trazadores nos ayude a reducir la varianza cósmica. Finalmente, concluimos que el sesgo es bajo y que los resultados son estadísticamente consistentes entre los trazadores. Asimismo, al combinar la covarianza de las galaxias, obtenemos resultados concordantes con otros trabajos que usaron los mismos datos.

Finalmente, estudiamos como el tamaño y el número de celdas, para obtener la covarianza de los multipolos de la función de correlación usando la técnica de *jackknife*, afecta en su precisión para el catálogo "SDSS-III BOSS CMASS sample". Además, los comparamos con covarianzas obtenidas a través de catálogos de galaxias simulados con la estimación *log-normal*, bajo la misma área del catálogo original. Queremos saber cuál es el error propagado a la escala característica de las oscilaciones acústicas de bariones, y para aquello utilizamos la corrección de *tapering* en ambas covarianzas. En conclusión, encontramos que aquel error no depende de la escala asociada del método de *jackknife*.

Abstract

This thesis's main objective is to explore some of the data mining and machine learning algorithms used in cosmological data, especially to the observables from Galaxy-Clustering in the Large Scale Structure of our Universe. We review how to construct them using the New York University Value-Added Galaxy Catalog. We reduce approximately 600000^3 galaxies coordinates to just 45^3. Then, we obtain constraints of the cosmological information encoded in the observables using the ΛCDM scenario and Bayesian inference. Furthermore, we contrast the results from the Metropolis-Hastings algorithm and density-estimation likelihood-free inference. The last is a deep learning algorithm that could be efficient than the first under the correct architecture. We obtain concordant results between both techniques but with underestimated errors in the shift and Redshift Space Distortions parameters, which encode the cosmological information, in comparison with works with similar data.

Besides, we combine the Redshift Space Distortions measurements $f\sigma_8(z)$ from several catalogs with the Hubble data $H(z)$ from cosmic chronometers to compare several Dark Energy models beyond ΛCDM. The performing of different statistical criteria bases our analysis. We use the Evidence comparison, Bayesian Information Criterion, Akaike Information Criteria, and the Figure of Merit. Therefore, we compare them to different frameworks. The flat ΛCDM model is favored under this data-set for all the criteria.

We also explore the observables from Galaxy-Clustering using different galaxy targets or tracers from the SDSS DR7 main galaxy sample. This insight could induce a bias in the two-point correlation function reflected in its amplitude and the Baryon Acoustic Oscillation peak. However, since these targets trace the same underlying dark matter field, we expect a small bias and constraints that help us beat the cosmic variance. We effectively conclude that the bias is small for this data-set, and the results are statistically consistent between tracers. We also find accurate constraints on the shift parameter by combining the tracers' covariances.

Finally, we study how the jackknife size and the number of resamplings impact the precision of the covariance estimate on the correlation function multipoles and the error on the inferred baryon acoustic scale, using the SDSS-III BOSS CMASS sample. We compare the measurements with log-normal mock galaxy catalogs with the same survey geometry. We also apply the tapering scheme to estimate the precision matrix for both paradigms. The results from CMASS and mock catalogs show that the error estimate of the baryon acoustic scale does not depend on the jackknife scale.

"Aequam memento rebus in arduis servare mentem"

Quintus Horatius Flaccus, Odae (II, 3).

Agradecimientos

Durante este camino de aprendizaje, me he topado con muchas personas que han depositado un poco de su conocimiento en mi y estoy profundamente agradecido por aquello. Es por este motivo que agradezco a todos mis profesores de mi enseñanza escolar, que de una u otra forma, hicieron crecer un interés en mi por las ciencias. Particularmente agradezco a la academia de astronomía del colegio encabezada por la profesora María Angélica, que ayudaron a prosperar aun más mi curiosidad por esa disciplina.

También a los profesores de la Universidad de Chile, quienes desde el primer día me hicieron ver la naturaleza con un exquisito detallismo basado en las matemáticas, y además un fuerte entusiasmo en ir más allá. A ellos les agradezco igualmente la visión de multidisciplinariedad que dejaron impregnado en mi como clave para el futuro. Asimismo a todos quienes conforman la Universidad y cada día hacen un gran esfuerzo para que se lleve a cabo su proyecto educativo.

Al mismo tiempo quiero agradecer profundamente a Domenico, mi profesor guía, que me ha motivado en varios momentos durante el desarrollo de esta tesis, y sin quien no hubiese sido posible haber concretado este proyecto.

Agradezco de igual forma a todas las personas que de alguna u otra forma se cruzaron en mi vida y terminaron siendo mis amigos, los aprecio con todo mi corazón. A aquellos y aquellas que conocí por el colegio, en el pregrado y en el postgrado, por una fiesta, o por mera casualidad, estoy muy grato con ustedes por todos los momentos de discusión y ocio que hemos compartido, y que me han hecho crecer como persona.

De igual forma agradezco infinitamente a toda mi familia, que me han apoyado durante mi carrera haciendo que esto sea posible, este logro también en suyo, muchas gracias.

Finalmente agradezco a todas y todos los científicos e investigadores que han plasmado su conocimiento, y del cual he tomado pinceladas para trazar este trabajo. Especialmente a los y las que se han dado el trabajo de pensar y diseñar los experimentos donde se han recolectado los datos que he usado.

Contents

List of Tables

List of Figures

Introduction

How our Universe starts if it does? What is its size? How old is it? Those have been some of the recurrent questions for ourselves as humans during all our existing. Which have been boarded by philosophy, theology, and science, whose replies have not been wholly satisfactory and concordant.

Nonetheless, since the Renaissance in the 15th century, science would start to be structurally ordered by a system known as the scientific method, which we use until nowadays. This change is essential to try to explain the questions mentioned above with a perspective based on the mathematical logic, where the first step was putting the Sun at the center and not the Earth. It also opened the field to the modern Cosmology. Over time, we discover new theories and create technology, which increases our knowledge and capabilities to understand our Universe. Nowadays, we have some explanations, based on the scientific method, for the original questions.

The modern Cosmology is able to partially explain the beginning, evolution, size, and age of our Universe consciously, even what components make it up. Unfortunately, we need to invoke to mystery explanations not entirely cleared at all yet, which are Inflation, the dark matter, and the dark energy. Nonetheless, these pillars are not fictitious because the actual data sustain them, then it is necessary to explain their nature. For that, in the last decades, more resources have been destined to observatories, and experiments focused on reducing our uncertainties to compare new results with several models proposed by theorists.

Even when we do not thoroughly understand the nature of Inflation and the dark sector, we have a favorite model subtended by actual data, and its name is ΛCDM. For it, the dark energy is driven by a cosmological constant in the Einstein equations, known as Λ, which encodes the vacuum energy in a possible quantum gravity theory. On the other hand, the CDM means Cold Dark Matter, which is a type of matter that does not interact (or it interacts very weakly) with photons, and it is the second principal component after dark energy. The particles that constitute stars, planets, and ourselves are known as baryonic matter. We also know pretty well forms of energy like photons, neutrinos, between others. These components, together with inflation, form the ΛCDM model. Nonetheless, other models beyond ΛCDM can change the nature of the dark energy, even some of them modify the Einstein gravity. Meantime, different approaches vary the candidates to dark matter with known and theoretical sources, which seems not to be affected by electromagnetic phenomena.

Nowadays, we have a window open to the observable Universe in all the frequencies of the electromagnetic spectrum. Even new ones are arising due to gravitational waves. This information received requires a lot of resources to catch it, save it, and analyze it, because we can not do it manually. Therefore, we need to develop machines that can perform those tasks, but since the resources are limited, we also need to ensure that they are efficient. These must obtain the most useful information in the least time possible and optimize the resources like memory and storage, for example.

We have developed new techniques to resolve some of the problems listed above, especially in data collection and analysis. For example, the art of reducing millions and millions of data to just a few enriched numbers is known as Data Mining, which uses many math models, available techniques, and others in the state-of-art for its purpose. Also, there are methods in which the machine is trained statistically with known data to apply then that knowledge with a new one. These techniques are what we know as machine learning. And it has been a great success recently since its capability to make good predictions without completely understanding the algorithms behind, especially in fields like Economy and Biology.

In Cosmology, we seek some data to construct what we call observables. These contain a lot of cosmological information that helps to better understand the Universe in what we live. Some of these observables are grouped according to the nature of the data. For example, we have the Supernovae (SN), the Cosmic Microwave Background (CMB), the Galaxy Clustering, and the Weak Lensing, between others. The SN, which are explosions of stars that release a lot of energy, can be observed from far away, allowing us to compute distances in the late Universe. Their magnitude can be standardized and they help us to determine the expansion of our Universe. On the other hand, the photons from the CMB radiation were the first released from the primordial plasma, which contains a lot of information of the early Universe. This radiation has very low energy nowadays, but with significant technological development, we can measure it well. Then, we have the Galaxy Clustering, which encodes how the galaxies are distributed in the late Universe. This geometrical disposition helps us to understand how was the evolution of our Universe from micro to macro cosmos due to the expansion. Finally, the Weak Lensing accounts for the deformations and magnifications of the shape of galaxies, which should be random. Still, like the matter content of our Universe deflects the light, it takes a non-zero signal that effectively measures the matter.

The CMB, Galaxy Clustering, and Weak Lensing signals are in general too small, hence we need to ensure an excellent signal to noise ratio to obtain useful cosmological information. We can do it by seeking a large patch in the sky because we use statistical techniques to construct those signals. For the CMB, we have complete sky maps, but we also manage small patches with a fair resolution that helps us understand the small scales. For the Galaxy Clustering and Weak Lensing, useful statistics requires a lot of galaxies distributed at some scale of interest, and in redshift if we want to consider the evolution. Therefore, we need to construct sensitive receivers that are increasingly crossing technological barriers.

Nonetheless, with the actual data, we can also try to improve the signal with some of the new techniques mentioned, especially those based in Signal Processing. We indeed have a system based on performing experiments every time more sensitives with a useful statistics framework to extract information. Still, maybe we will not be able to do that forever due to the physical limitations of our Universe. Thus, it is vital to research new signals or observables that can help us in our task and spend time on how we can extract the maximal information from the existing ones.

The previous comments motive us to optimally process data for Cosmology, especially the Large Scale Structure of our Universe. This process requires several steps from the data collection until the extract of cosmological information. Therefore, we will review how we can do it and the techniques, traditional and new ones, involved in these processes.

The outline of this thesis is the following: In chapter 1, we underline the Cosmological theory based on the ΛCDM model at the background level, which says to us how our homogeneous and isotropic Universe evolves. Besides, we seek some of the principal distances and how they are related to the redshift, which is one of the attributes best determined. Finally, we mention some of the principal epochs, at a background level, that allowed our Universe.

In chapter 2, we go beyond the Cosmological background theory, and we study our Universe at first order in perturbations in energy density using General Relativity as our framework. We obtain the cosmological equations in real and Fourier space to apply them to the principal components of our Universe: photons, baryons, dark matter, and neutrinos. We excluded dark energy since it has not perturbations by definition of the ΛCDM model.

Once constructed the background and the linear order of perturbations for the ΛCDM model, we can trace the history of our Universe, which we do in chapter 3. For that, we also require the mechanism of inflation, which we explain briefly. Then, we can track the formation of the principal components, their interactions and evolution, and, more importantly, the fingerprints that they left in the actual observables.

In chapter 4, we construct the principal observable for this thesis, which is based on the two-point statistics for the matter density contrast. Besides, we study how to obtain this signal from a galaxy survey and relate it with the theoretical one, considering the distortions induced by the proper movements of galaxies and their distances.

We also need to review the technicalities required to construct our observable. For that reason, in chapter 5, we make a brief review of the design of the principal algorithms to build the observable, and how to extract cosmological information from it. We also need to find methods to make them faster, especially in data mining, since we will analyze a lot of information. For that task, we emphasize some of the more efficient algorithms.

In chapter 6, we take the New York University Value-Added Galaxy Catalog to construct our observable using the techniques studied in the last chapters. We will build several estimators, which we will compare. Finally, we also estimate the uncertainty on the observable.

In chapter 7, we choose one of the observables constructed in the last chapter to apply statistical techniques and algorithms to obtain the cosmological information. We use two ways of inference, but we also compare several models in the literature. Finally, we discuss the results, and we make a brief analysis of how they can improve in the future.

We also present three works related with observables from Galaxy Clustering in chapters 8 to 10, which were developed during the master project. We compare the cosmic chronometers versus the growth rate of structure in chapter 8. The first is an expansion measurement, and the second is a geometrical one since we can obtain it from the analysis of Galaxy Clustering. Then, we use actual data in both measurements to constrain several dark energy models against the ΛCDM.

In chapter 9, we construct the two-point correlation function for three galaxy tracers from the same survey. Then, we constrain that data to get cosmological information and compare how it differs between the targets.

Finally, in chapter 10, we compute the covariance of the two-point correlation function using the jackknife technique and mock catalogs. Then, we compare the final cosmological information obtained from the different methods. The importance of this lies in the lack of a strict consensus on how we can get the uncertainties in Galaxy Clustering. Even when the mock catalogs are the most used, they require some assumptions or additional information.

Chapter 1

Basics of Cosmology: The Background

The main hypothesis of modern Cosmology is the cosmological principle [150], which says that the distribution of matter and energy in our Universe is homogeneous and isotropic on large enough scales. We can see this as a natural generalization of the Copernican principle in which we are not in a privileged point in the Universe, but also that no point is. This statement is not a big deal for a static universe, since it has always been and has stayed in the same way. Nonetheless, the discovery made by Hubble a century ago [129] suggests that our Universe is not static; in fact, the physical distance between two points is increasing on time. Therefore, the cosmological principle and the Universe expansion are going to lay the foundations in the description of the smooth and expanding Universe and its evolution, which we will study in this chapter basing most of our discussions in [23, 31].

As we study our Universe on lower cosmological scales, it seems that the cosmology principle is broken, which happens in almost all physical theories because chaotic processes appear. For that, our Universe on large scales, which is homogeneous and isotropic, will be referred to as the background universe or just background. To describe it, we are going to use the four-dimensional coordinates $X^\mu \equiv (x^0 = t, x^1, x^2, x^3)$ since we are going to work in a General Relativity framework, where t is the cosmic time, and x^i are the three-dimensional coordinates. We also set the fundamental constants to one $c = \hbar = k_B = 1$. Besides we always are going to use dots to indicate derivation with respect to cosmic time $\dot{\vartheta} = \mathrm{d}\vartheta/\mathrm{d}t$ and primes to indicate derivation with respect to conformal time $\vartheta' = \mathrm{d}\vartheta/\mathrm{d}\tau$, which we will be define in the next section. Finally we will use the Minkowski metric under the convention $(-1, 1, 1, 1)$.

1.1 The Expanding Universe

After Hubble discovered that the galaxies are moving away from each other, it was necessary to reconsider the real distance between two points in the space. We know that the physical distance, given a reference system, is $\Delta\vec{r} = \vec{r_2} - \vec{r_1}$ in which $\vec{r_1}$ and $\vec{r_2}$ are the position vectors of the objects under the reference system. Since the coordinates themself are changing, the election of a coordinate system is no longer trivial; therefore, we need to consider an ideal system that contains our Universe, but it is not expanding, this is also known as a system that does not follow the Hubble flow. For that system, we preserve the notion of distance that we had, and then it is just $\Delta\vec{r}$. However, as observers, we only are available to measure distances in a system that follows the Hubble flow because we are in the Universe, not outside it. This defines the comoving frame, in which the distance $\Delta\vec{x} = \vec{x_2} - \vec{x_1}$ is the comoving distance, where $\vec{x_1}$ and $\vec{x_2}$ are the position vectors in comoving coordinates.

The cosmological principle states that our Universe is homogeneous and isotropic, alongside with its expansion. Then, if we observe that two points moving away of $100l.y.$ during a period Δt, then other two points in a different region will also move away of $100l.y.$. We can express that sentence defining a function that relates the physical coordinates with the comoving ones, which is kwon as the scalar factor a, and imposing that it must depend only on time: $a(t)$, because if not, the expansion would stop being homogeneous. Hence we can relate both coordinates systems using $\vec{r} = a(t)\vec{x}$, and therefore, the relation between distances is $\Delta\vec{r} = a(t)\Delta\vec{x}$. We can derive it using the chain rule and obtain

$$\begin{aligned}\frac{\mathrm{d}\Delta\vec{r}}{\mathrm{d}t} &= \dot{a}(t)\Delta\vec{x} + a(t)\dot{\Delta\vec{x}} \\ &= \frac{\dot{a}(t)}{a(t)}\Delta\vec{r} + a(t)\vec{v} \\ &= H(t)\Delta\vec{r} + \vec{v}_p\,,\end{aligned} \tag{1.1}$$

where we defined the peculiar or proper velocity of the object with respect to the Hubble flow $\vec{v}_p \equiv a(t)\vec{v}$ and the Hubble parameter as

$$H(t) \equiv \frac{\dot{a}}{a}\,. \tag{1.2}$$

In the absence of proper velocity, the eq. (1.1) tells us that the velocity between two points is proportional to the distance, this is known as the Hubble law. In our reference system (the Earth) and at present time, the Hubble parameter is known as the Hubble constant $H_0 \equiv H(t = t_{today}) = 100h$ km s^{-1} Mpc^{-1}, where we also defined the dimensionless reduced Hubble parameter h, that describes the uncertainty in the value of H_0. We also are going to introduce the conformal time τ as

$$\tau \equiv \int \frac{\mathrm{d}t}{a(t)}\,, \tag{1.3}$$

which says that for a free particle with velocity $c = 1$ as photons, the comoving distance traveled during a conformal time interval $\Delta\tau$ is just $\Delta\tau$. In the same way that eq. (1.2), we can define a conformal Hubble parameter as

$$\begin{aligned}\mathcal{H} \equiv \frac{1}{a}\frac{\mathrm{d}a}{\mathrm{d}\tau} &= \frac{a'}{a} \\ &= aH\,.\end{aligned} \tag{1.4}$$

1.2 The Friedmann Equations

The Friedmann equations describe the dynamics of the background universe. To obtain them we need the framework of General Relativity theory introduced by Einstein in 1916, a metric that describes the spacetime $g_{\mu\nu}(t, \vec{x})$ and characterizes the line-element $\mathrm{d}s^2 \equiv g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$, and also the energy-momentum tensor $T_{\mu\nu}$ that contains all the possible matter and energy species to consider. Those enter in the Einstein equations as

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} \,, \tag{1.5}$$

where we define the Einstein tensor $G_{\mu\nu}$, G is the Newton gravitational constant, $R_{\mu\nu}$ is the Ricci tensor, and R is the Ricci scalar or curvature scalar, given by

$$\begin{aligned} R &\equiv tr\left(R_{\mu\nu}\right) \\ &= g^{\mu\nu} R_{\mu\nu} \,. \end{aligned} \tag{1.6}$$

In the equation above $g^{\mu\nu}$ is the inverse of the metric that satisfies $g^{\mu\lambda} g_{\lambda\nu} = \delta^{\mu}_{\nu}$, where δ^{μ}_{ν} is the Kronecker delta. The Ricci tensor is defined through

$$R_{\mu\nu} \equiv \partial_\lambda \Gamma^{\lambda}{}_{\mu\nu} - \partial_\nu \Gamma^{\lambda}{}_{\mu\lambda} + \Gamma^{\lambda}{}_{\lambda\rho} \Gamma^{\rho}{}_{\mu\nu} - \Gamma^{\rho}{}_{\mu\lambda} \Gamma^{\lambda}{}_{\nu\rho} \,, \tag{1.7}$$

where ∂_α indicates a derivation with respect to X^α and $\Gamma^{\mu}{}_{\alpha\beta}$ are the Christoffel symbols, which we define using the derivatives of the metric as

$$\Gamma^{\mu}{}_{\alpha\beta} \equiv \frac{1}{2} g^{\mu\lambda} \left(\partial_\alpha g_{\beta\lambda} + \partial_\beta g_{\alpha\lambda} - \partial_\lambda g_{\alpha\beta} \right) \,. \tag{1.8}$$

The eq. (1.5) describes the geometry of the spacetime that controls the dynamics over the objects that it contains. It was applied to study our Solar system, but also to get solutions to static universes. After Hubble observations, the scientists Friedmann, Lemaitre, Roberson, and Walker propose solutions of eq. (1.5) for non-stationary Universes. They used the Friedman-Lemaitre-Roberson-Walker (FLRW) metric to describe a homogeneous, isotropic and expanding Universe, where its line-element is characterized by

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t) \left(\frac{\mathrm{d}r^2}{1 - K r^2} + r^2 \left(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2 \right) \right) \,, \tag{1.9}$$

here we used polar coordinates and where the constant $K = -1, 0, 1$ indicates if we are talking about a closed, open or flat geometry, respectively. Note that the spacetime in eq. (1.9) has a rescaling symmetry under

$$a \to \lambda a \,, \;\; r \to r/\lambda \,, \;\; K \to \lambda^2 K \,, \tag{1.10}$$

which tells us that the geometry of the Universe stays the same if we rescale a, r, K simultaneously by some λ. Thus we can use this freedom to set the scale factor today to one: $a_0 \equiv a(t = t_{today}) = 1$, and then a becomes dimensionless, meanwhile r and $K^{-1/2}$ inherit the dimension of length.

The energy-momentum tensor for one single perfect fluid with energy density $\rho(t)$ and pressure $P(t)$, is:

$$T^{\mu}{}_{\nu} = \left(\rho(t) + P(t) \right) u^\mu u_\nu + P \delta^{\mu}_{\nu} \,, \tag{1.11}$$

where u^μ is the four-velocity of the fluid in comoving coordinates, which obeys $g_{\mu\nu} u^\mu u^\nu = -1$.

To obtain the Einstein equations, we need to compute the Christoffel symbols through eq. (1.8) and the metric in eq. (1.9):

$$\Gamma^{0}{}_{ij} = a^2 H \gamma_{ij} \,, \tag{1.12}$$

$$\Gamma^{i}{}_{0j} = \Gamma^{i}{}_{j0} = H\delta^{i}_{j}\,, \tag{1.13}$$

$$\Gamma^{1}{}_{11} = \frac{Kr}{1-Kr^{2}}\,, \tag{1.14}$$

$$\Gamma^{1}{}_{22} = -r\left(1-Kr^{2}\right)\,, \tag{1.15}$$

$$\Gamma^{1}{}_{33} = -r\left(1-Kr^{2}\right)\sin^{2}\theta\,, \tag{1.16}$$

$$\Gamma^{2}{}_{33} = \sin\theta\cos\theta\,, \tag{1.17}$$

$$\Gamma^{2}{}_{12} = \Gamma^{2}{}_{21} = \Gamma^{3}{}_{13} = \Gamma^{3}{}_{31} = \frac{1}{r}\,, \tag{1.18}$$

$$\Gamma^{3}{}_{23} = \Gamma^{3}{}_{32} = \cot\theta\,, \tag{1.19}$$

whereas, the other Christoffel symbols, which indexes do not appear in the equations above, vanishes. In eq. (1.12), γ_{ij} is defined as

$$\gamma_{ij} \equiv \delta_{ij} + K\frac{x_{i}x_{j}}{1-K\left(x_{k}x^{k}\right)}\,, \tag{1.20}$$

since we choose polar coordinates, γ_{ij} becomes diagonal, and its values are

$$\gamma_{11} = \frac{1}{1-Kr^{2}}\,, \tag{1.21}$$

$$\gamma_{22} = r^{2}\,, \tag{1.22}$$

$$\gamma_{33} = r^{2}\sin^{2}\theta\,. \tag{1.23}$$

Using eqs. (1.12) to (1.19) we proceed to compute the Ricci tensor through eq. (1.7):

$$R_{00} = -3\left(H^{2}+\dot{H}\right)\,, \tag{1.24}$$

$$R_{0i} = R_{i0} = 0\,, \tag{1.25}$$

$$R_{ij} = a^{2}\left(3H^{2}+\dot{H}+\frac{2K}{a^{2}}\right)\gamma_{ij}\,, \tag{1.26}$$

then we can obtain the Ricci tensor using eq. (1.6) and the FLRW metric (eq. (1.9))

$$R = 6\left(2H^{2}+\dot{H}+\frac{K}{a^{2}}\right)\,. \tag{1.27}$$

Finally, the components of the Einstein tensor, using $G^{\mu}{}_{\nu} = g^{\mu\alpha}G_{\alpha_{\nu}}$, are:

$$G^{0}{}_{0} = -3\left(H^{2}+\frac{K}{a^{2}}\right)\,, \tag{1.28}$$

$$G^{0}{}_{i} = G^{i}{}_{0} = 0\,, \tag{1.29}$$

$$G^{i}{}_{j} = -\left(3H^{2}+2\dot{H}+\frac{K}{a^{2}}\right)\delta^{i}_{j}\,, \tag{1.30}$$

then we can connect them directly with the energy momentum ${T^{\mu}}_{\nu}$ in eq. (1.11) through ${G^{\mu}}_{\nu} = 8\pi G {T^{\mu}}_{\nu}$. For the (00) component we have that $T^{0}_{\ 0} = -\rho$, which gives us the first Einstein equation

$$H^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2}\,. \tag{1.31}$$

For the (ij) component, we have that $T^{i}_{\ j} = P\delta^{i}_{j}$, which give us the second Einstein equation

$$3H^2 + 2\dot{H} = -8\pi G P - \frac{K}{a^2}\,. \tag{1.32}$$

If we combine eqs. (1.31) and (1.32) to eliminate the K/a^2 term, we obtain the Friedmann equation

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + 3P\right)\,. \tag{1.33}$$

It is convenient to rewrite eq. (1.31) as

$$\Omega + \Omega_K = 1\,, \tag{1.34}$$

where we had defined the normalized density parameter Ω:

$$\Omega \equiv \frac{8\pi G\rho}{3H^2}\,, \tag{1.35}$$

and the normalized density parameter of curvature Ω_K:

$$\Omega_K \equiv -\frac{K}{\left(aH\right)^2}\,. \tag{1.36}$$

The Einstein tensor satisfies the Bianchi identities $\nabla_{\mu}{G^{\mu}}_{\nu} = 0$, where

$$\nabla_{\mu}{G^{\mu}}_{\nu} \equiv \partial_{\mu}{G^{\mu}}_{\nu} + {\Gamma^{\mu}}_{\mu\alpha}{G^{\alpha}}_{\nu} - {\Gamma^{\alpha}}_{\nu\mu}{G^{\mu}}_{\alpha}\,. \tag{1.37}$$

Then by eq. (1.5), the energy-momentum tensor also satisfies the Bianchi identities which imply that the energy-momentum tensor is conserved

$$\nabla_{\mu}{T^{\mu}}_{\nu} = 0\,, \tag{1.38}$$

therefore we can obtain another equation using the $\nu = 0$ component of eq. (1.38), together with eqs. (1.11) to (1.19) and (1.37):

$$\dot{\rho} + 3H\left(\rho + P\right) = 0\,. \tag{1.39}$$

For adiabatic perfect fluids, the pressure is a function of the energy density $P(t) = w\rho(t)$, where w defines the equation of state parameter. We omitted some time-space dependency or in another variable for the function w, since at the beginning, we are going to study fluids where it is constant *i.e.* just a number. In those cases the eq. (1.39) has a solution given by

$$\rho(t) = \rho_0 \exp\left(-3\left(1+w\right)\int_{t_0}^{t} H(t')\mathrm{d}t'\right)\,. \tag{1.40}$$

1.3 Cosmic Distances

We have shown that in an expanding Universe, our inherent concept of physical distance changes a bit. Thus, we have to find a way to measure distances in our Universe, which we call observable distances, but it is not trivial because of the expansion and the geometry present in the FLRW spacetime. Besides, the observable distances can be affected by astrophysical processes, which varies depending in the source that we are seeking. To start, we are going to write the eq. (1.9) recovering the c units

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a^2(t)\left(\mathrm{d}\chi^2 + f_K^2\left(\chi\right)\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2\right)\right) , \tag{1.41}$$

where we have defined the radial coordinate χ through the change of variables

$$\mathrm{d}\chi \equiv \frac{\mathrm{d}r^2}{\sqrt{1-Kr^2}} , \tag{1.42}$$

therefore under the right integral limits, χ is

$$\chi = \begin{cases} R_0 \sinh^{-1}\left(\frac{r}{R_0}\right) & K=-1 \\ r & K=0 \\ R_0 \sin^{-1}\left(\frac{r}{R_0}\right) & K=1 \end{cases} . \tag{1.43}$$

In eq. (1.41) we also defined the function f_K, that appears after the change of variable, and its values are

$$f_K\left(\chi\right) \equiv \begin{cases} R_0 \sinh\left(\frac{\chi}{R_0}\right) & K=-1 \\ \chi & K=0 \\ R_0 \sin\left(\frac{\chi}{R_0}\right) & K=1 \end{cases} , \tag{1.44}$$

which is also possible to write in a compact way as

$$f_K\left(\chi\right) = \frac{R_0}{\sqrt{-K}}\sinh\left(\sqrt{-K}\frac{\chi}{R_0}\right) , \tag{1.45}$$

where we can recover the value for flat geometries taking the limit $K \to -0$. Note that in eqs. (1.43) to (1.45) we introduced the length scale R_0, which appears by completeness because we choose to make the scale factor dimensionless in the last section.

Because of the expansion, the light that gives us the information from distant objects suffers changes in its wavelength. That process is known as Redshift since the wavelength of the light gets stretched. Suppose we observe a far galaxy at a fixed comoving distance from us that emits a signal whose duration is $\Delta\tau$ in conformal time. In that case, its value still will be $\Delta\tau$ when we receive that signal since, in the conformal frame, the time intervals are the same for free particles. Nonetheless, the physical time intervals are different, those that we can measure at the emission and detection points. In the point of emission the time interval is $\Delta t_1 = a(\tau_1)\Delta\tau$, and in the point of detection the time interval is $\Delta t_0 = a(\tau_0)\Delta\tau$. Using the relation $\lambda = c\Delta t$ we find that

$$\frac{\lambda_0}{\lambda_1} = \frac{a(\tau_0)}{a(\tau_1)} , \tag{1.46}$$

then we check that in absence of expansion *i.e.* $a(\tau_0) = a(\tau_1)$, the wavelength is unchanged. To take account in the change of wavelengths, we are going to define the Redshift parameter z as

$$z \equiv \frac{\lambda_0 - \lambda_1}{\lambda_1}\,, \tag{1.47}$$

then by eq. (1.46), we have that

$$\begin{aligned} 1 + z &= \frac{a(t_0)}{a(t_1)} \\ &= \frac{1}{a(t_1)}\,, \end{aligned} \tag{1.48}$$

where in the last line, we used the freedom given by the rescaling symmetry (eq. (1.10)).

1.3.1 Proper Distance

The proper distance is the distance between two objects measured at some specific time t simultaneously, then $\mathrm{d}t = 0$ in eq. (1.41), and assuming that the two points are in the same direction *i.e.* $\mathrm{d}\Omega^2 \equiv \mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2 = 0$. Therefore, the metric in eq. (1.41) becomes

$$\mathrm{d}s^2 = a^2(t)\mathrm{d}\chi^2\,, \tag{1.49}$$

and the proper distance d_p is the integral over $\mathrm{d}s$

$$\begin{aligned} \mathrm{d}_p &\equiv \int \mathrm{d}s \\ &= a(t)\int_{\chi_1}^{\chi_2} \mathrm{d}\chi = a(t)\Delta\chi\,. \end{aligned} \tag{1.50}$$

Note that we have found a similar expression in section 1.1, where we called $\Delta\chi$ as the comoving distance, which will be introduced in the next section formally. Nonetheless, we must emphasize that the proper distance is not an observable because it violates causality.

1.3.2 Comoving Distance

We already argued that the comoving distance emerges in an expanding universe because, in a static one, we are only concerned about the physical distance. We also saw that the proper distance is not physical since it violates causality, hence we need to introduce new measurable distances. Imagine that we want to measure the distance between two points, where we start measuring at time t_1 from the first and then we travel until the second point at time t_2. If this measure is made at the light speed, we have $\mathrm{d}s = 0$ since light satisfies null geodesic equations. Besides, we are going to assume that the two points are in the same direction ($\mathrm{d}\Omega^2 = 0$). Then, eq. (1.41) takes the form

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a(t)^2\mathrm{d}\chi^2 = 0\,, \tag{1.51}$$

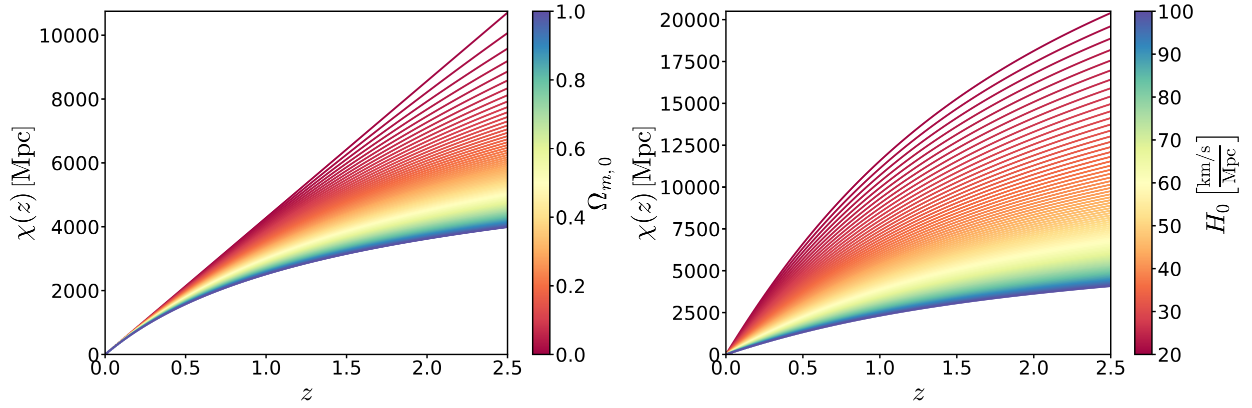


Figure 1.1: The comoving distance as function of the redshift for the flat ΛCDM model, where we also show its variation as function of the $\Omega_{m,0}$ parameter (*Left*), which is equal to $1 - \Omega_{\Lambda,0}$ under this model, and as function of the H_0 parameter (*Right*).

and we define the comoving distance d_c as the integral over $\mathrm{d}\chi$

$$\begin{aligned} \mathrm{d}_c &\equiv \int_{\chi_1}^{\chi_2} \mathrm{d}\chi \\ &= \int_{t_1}^{t_2} \frac{c}{a(t)} \mathrm{d}t \quad = c \int_{a_1}^{a_2} \frac{\mathrm{d}a}{a^2 H(a)} \\ &\qquad\qquad\qquad = -c \int_{z_1}^{z_2} \frac{\mathrm{d}z}{H(z)} \,, \end{aligned} \tag{1.52}$$

where in the second line we used the definition in eq. (1.2) to obtain $\mathrm{d}t = \mathrm{d}a/(aH)$, and in the third line we used the relation in eq. (1.48) to get $\mathrm{d}a = -\mathrm{d}z/(1+z)^2$. If we want to compute the distance to us, then $t_2 = t_0$, with $a(t_2) = 1$ and by eq. (1.48), $z_2 = 0$. Finally, the comoving distance between an object at redshift z and us is

$$\mathrm{d}_c = c \int_0^z \frac{\mathrm{d}z'}{H(z')} \,, \tag{1.53}$$

which is shown in fig. 1.1. Note that the comoving distance is also not an observable, because we are not available to measure at light speed.

1.3.3 Luminosity Distance

To obtain a measurable distance, we also need to consider the processes in which the light is involved. Light coming from observations in the form of flux, which we define as

$$\mathcal{F} \equiv \frac{L}{S} \,, \tag{1.54}$$

where L is the absolute luminosity of the emission source, and S is the area in which the luminosity is projected as light travels. For a flat Universe, the surface projected is a sphere with area $A = 4\pi r^2$, where r is the proper radius $r = a\chi$. Nevertheless, for a non-flat Universe, the general expression for that surface is the integral over the angular part in

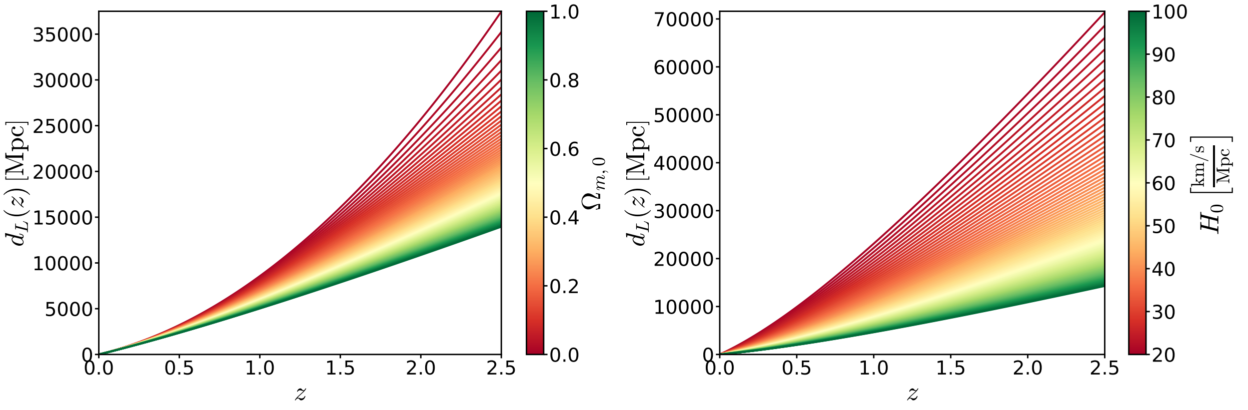


Figure 1.2: The luminosity distance as function of the redshift for the flat ΛCDM model, where we also show its variation as function of the $\Omega_{m,0}$ parameter (*Left*), which is equal to $1-\Omega_{\Lambda,0}$ under this model, and as function of the H_0 parameter (*Right*).

eq. (1.41), *i.e.* $S=\int a^2 f_K^2(\chi)\mathrm{d}\Omega^2 = 4\pi a^2 f_K^2(\chi)$, where we computed it in all the angles. We can write the light flux of eq. (1.54) as

$$\mathcal{F}=\frac{L_O}{4\pi a^2(t_0) f_K^2(\chi)}\,, \tag{1.55}$$

here we called L_O as the luminosity observed at time t_0, for example, in the telescope. That is not the same at the emission point because the expansion causes luminosity to be redshifted, then it is not absolute at all. We proceed to introduce the luminosity distance as

$$\mathrm{d}_L \equiv \sqrt{\frac{L_s}{4\pi\mathcal{F}}}\,, \tag{1.56}$$

where L_s is the luminosity produced by the source. Then, after using eq. (1.55), the luminosity distance is

$$\mathrm{d}_L = a(t_0) f_K(\chi)\sqrt{\frac{L_s}{L_0}}\,. \tag{1.57}$$

Like we mentioned before the expansion also makes the luminosity to be redshifted. Since the luminosity is the energy over amount of time $L=E/\Delta t$ and the energy of the photons depends on the wavelength as $E\propto\lambda^{-1}$, we have that $E_s/E_O=\lambda_O/\lambda_s=a(t_0)/a(t_s)$ because eq. (1.46). We will also use that $\Delta t_0 = a(t_0)\Delta\tau$ and $\Delta t_s = a(t_s)\Delta\tau$ where t_s is the time of emission. Both of them have the same $\Delta\tau$ since it is the same signal as we argued in the derivation of redshift formula (section 1.3). Therefore $L_s/L_O = (E_s/\Delta t_s)/(E_O/\Delta t_0) = a^2(t_0)/a^2(t_s)$, which we can replace in eq. (1.57), and obtain

$$\mathrm{d}_L = \frac{a^2(t_0)}{a(t_s)} f_K(\chi)\,. \tag{1.58}$$

Note that as observers, we defined t_0 at our time and then $a(t_0)=1$ because of the rescaling symmetry. Using eq. (1.45) we can write the luminosity distance as

$$\mathrm{d}_L(a) = \frac{1}{a}\frac{R_0}{\sqrt{-K}}\sinh\left(c\frac{\sqrt{-K}}{R_0}\int_a^1 \frac{\mathrm{d}a'}{a'^2 H(a')}\right)\,, \tag{1.59}$$

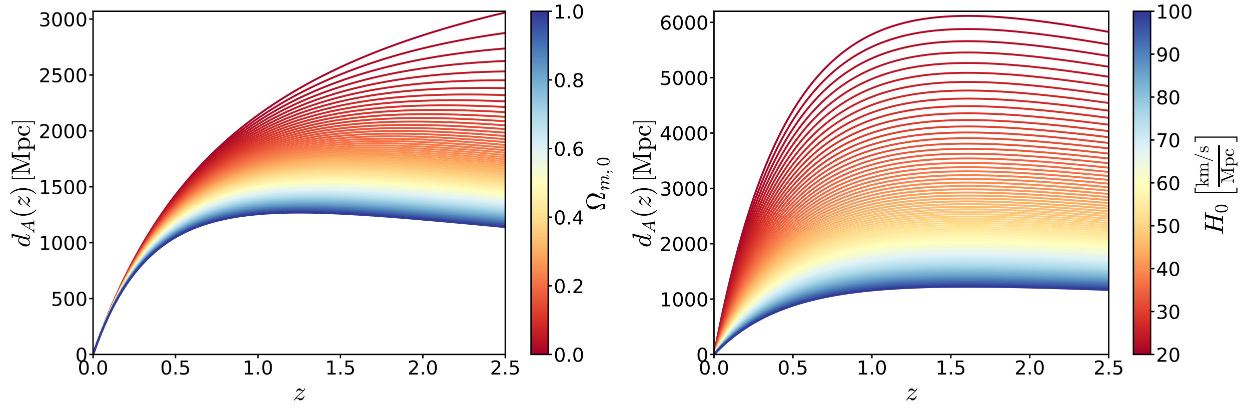


Figure 1.3: The angular diameter distance as function of the redshift for the flat ΛCDM model, where we also show its variation as function of the $\Omega_{m,0}$ parameter (*Left*), which is equal to $1 - \Omega_{\Lambda,0}$ under this model, and as function of the H_0 parameter (*Right*).

where we also used eq. (1.52). In redshift space we can use the eq. (1.53) and therefore eq. (1.59) takes the form

$$\mathrm{d}_L(z) = \frac{R_0}{\sqrt{-K}}(1+z)\sinh\left(c\frac{\sqrt{-K}}{R_0}\int_0^z \frac{\mathrm{d}z'}{H(z')}\right) . \tag{1.60}$$

which is shown in fig. 1.2 for a flat geometry. The luminosity distance is used when we know the luminosity of the source, because it is the same in all the cosmos, at least without considering the redshift. These types of sources are named standard candles, and an example is Supernovae Ia, which has a light curve that can be calibrated and standardized.

1.3.4 Angular Diameter Distance

Another observable distance is that we can measure from the size of a distant object l, which we observe in the sky with an angle $\delta\theta$. It is called the angular diameter distance, and we define it as

$$\mathrm{d}_A \equiv \frac{l}{\delta\theta} , \tag{1.61}$$

where $\delta\theta \ll 1$. The object size in a flat Universe that sustains and angle $\mathrm{d}\Omega$ would be $l = r\mathrm{d}\Omega$, with r the conformal radius, but for a general geometry the angular length of the object is described by the angular part of eq. (1.41), *i.e.* $l = a(t_s)f_K\mathrm{d}\Omega$ where t_s is the time of emission. Setting $\mathrm{d}\Omega = \delta\theta$ in eq. (1.61), we find that the angular distance is

$$\mathrm{d}_A = a(t_s)f_K(\chi) , \tag{1.62}$$

which after using eqs. (1.45) and (1.52) becomes in

$$\mathrm{d}_A(a) = a\frac{R_0}{\sqrt{-K}}\sinh\left(c\frac{\sqrt{-K}}{R_0}\int_a^1 \frac{\mathrm{d}a'}{a'^2H(a')}\right) . \tag{1.63}$$

Now using eq. (1.53) we find the eq. (1.63) as function of redshift

$$\mathrm{d}_A(z) = \frac{R_0}{\sqrt{-K}}(1+z)^{-1}\sinh\left(c\frac{\sqrt{-K}}{R_0}\int_0^z \frac{\mathrm{d}z'}{H(z')}\right) , \tag{1.64}$$

which is shown in fig. 1.3 for a flat geometry. We also note the similitude with eq. (1.60), in fact, the relation between both distances is

$$\mathrm{d}_A(z) = \frac{\mathrm{d}_L(z)}{(1+z)^2} \,. \tag{1.65}$$

The angular diameter distance is used to measure the distances from objects or events with the same size; those are known as standard ruler. Examples are the CMB anisotropies and the BAO peak at large scales in Galaxy Clustering.

1.4 The Epochs

In general, it is impossible to find an analytical solution for $a(t)$ using the Friedmann equation (eq. (1.33)) because, in a complex Universe, many species enter into the energy-momentum tensor. Nonetheless, we can find good approximations if we consider that the Universe is dominated by a single component, which certainly works at some epochs in the history of our Universe. We already saw that there is a form in which we can relate the pressure and the energy density through an equation of state parameter (EoS) defined as

$$w(a) \equiv \frac{P}{\rho} \,. \tag{1.66}$$

Therefore assuming only one single component with an EoS parameter w, we can write eq. (1.39) as

$$\frac{\mathrm{d}\rho}{\rho} = -3(1+w)\frac{\mathrm{d}a}{a} \,, \tag{1.67}$$

which for a constant w has the next solution

$$\rho(a) = \rho_0 a^{-3(1+w)} \,, \tag{1.68}$$

where we are going to set ρ_0 as the energy density today ($t = t_0$ and $a = 1$). Inserting the solution of eq. (1.68) into eq. (1.33) we find a equation for $a(t)$, which solution is

$$a(t) = \left(\frac{t}{t_0}\right)^{\frac{2}{3(1+w)}} \,. \tag{1.69}$$

If we replace eq. (1.69) into eq. (1.68) we can find the solution of $\rho(t)$ that does not depend on the EoS parameter

$$\rho(t) = \rho_0 \left(\frac{t}{t_0}\right)^{-2} \,. \tag{1.70}$$

The eq. (1.70) tells us that the energy density always decreases because the expansion causes the energy to spread in a larger volume. We can see its behavior for different species in fig. 1.4, and for $\Omega(a)$ in fig. 1.6.

If we have more than one component in the Universe, the eq. (1.39) remains the same, but now $\rho = \sum \rho_i$ and $P = \sum w_i \rho_i$ because the energy-momentum is additive. However, this time is not possible to find and exact solution for $\rho(a)$. In these cases, we can approximate

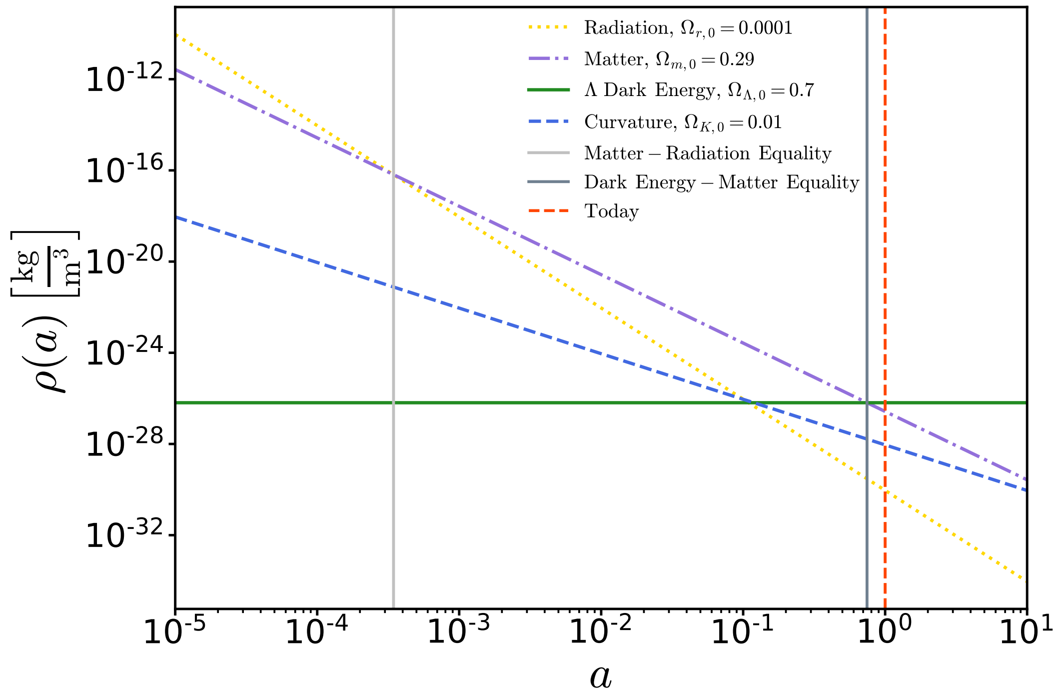


Figure 1.4: The behavior of the energy density as function of the scale factor using eq. (1.68) for different components, as well as the curvature.

$\rho \sim \rho_{\mathrm{d}}$, where ρ_{d} is the specie that dominates in a certain epoch due to eq. (1.68), and this depends exclusively on the EoS parameter w (see fig. 1.4).

If the EoS parameter depends only on the scale factor or time, the eq. (1.67) still works, but now the solution is given by

$$\rho(a) = \rho_0 a^{-3} \exp\left(-3\int_1^a \frac{w(a')}{a'} \mathrm{d}a'\right) . \tag{1.71}$$

Finally if we consider the three most abundant components in our Universe. *i.e.* radiation($w = w_r$), matter ($w = w_m$) and dark energy ($w = w(a)$), we can write the Hubble parameter using eq. (1.31) but with ρ as the sum of our species, whose solutions will be given by eqs. (1.68) and (1.71). Then, the total Hubble parameter is

$$\begin{aligned} H^2(a) = \frac{8\pi G}{3}\Bigg[&\rho_{r,0}a^{-3(1+w_r)} + \rho_{m,0}a^{-3(1+w_m)} + \rho_{K,0}a^{-2} \\ &+ \rho_{DE,0}a^{-3}\exp\left(-3\int_{a_0=1}^{a} \frac{w(a')}{a'}\mathrm{d}a'\right)\Bigg] , \end{aligned} \tag{1.72}$$

where $\rho_{K,0} = -3K/(8\pi G)$.

1.4.1 The Radiation Domination Era (RDE)

We are going to consider as radiation all the species with $p \gg m$, like photons and relativistic neutrinos. In those cases, the EoS parameter is $w = 1/3$ for both bosons and fermions, greater than matter and dark energy. In an expanding Universe, for greater values of w, earlier it is domination due to eq. (1.68). Therefore the radiation was the first specie dominating in the Universe, and during that epoch, the Hubble parameter (see fig. 1.5) can be approximated as

$$H^2(a) \approx H_0^2 \Omega_{r,0} a^{-4} \,, \tag{1.73}$$

where $\Omega_{r,0} = 8\pi G \rho_{r,0} / \left(3H_0^2\right) \sim 8.051 \times 10^{-5}$ [233] is the normalized radiation energy density parameter at nowadays.

1.4.2 The Matter Domination Era (MDE)

For non-relativistic particles, the equation of state parameter is $w \simeq 0$ for both bosons and fermions. Then $w_m < w_r$ making matter the next specie that dominates in the history of our Universe after RDE. During this epoch, the Hubble parameter (see fig. 1.5) can be approximated as

$$H^2(a) \approx H_0^2 \Omega_{m,0} a^{-3} \,, \tag{1.74}$$

where $\Omega_{m,0} = 8\pi G \rho_{m,0} / \left(3H_0^2\right) = 0.311 \pm 0.006$ [59] is the normalized matter energy density parameter at nowadays. Here we do not make the difference between dark matter and baryons, since both behaves as the same at the background level.

1.4.3 The Dark Energy Domination Era (DEDE)

The next "specie" to consider should be the curvature since as we saw in eq. (1.72), it obeys a power law proportional to a^{-2} which due eq. (1.68) is equivalent to have an EoS parameter $w_K = -1/3$. However, the curvature has a minimal normalized energy density Ω_K at all the epochs in the history of our Universe (a problem that we will discuss in section 3.1) and then it can not compete with the other species.

An Universe with accelerating expansion requires a component with $w < -1/3$; as the dark energy is responsible for that acceleration, we can infer that it has that kind of EoS parameter. Since it is a complete mystery, many theories try to explain its nature and propose several models for w. The most accepted is $w = -1$ for dark energy driven by a cosmological constant Λ. Nowadays the dark energy is the dominant specie in our Universe. However, it is not entirely dominant yet to approximate the Hubble parameter only with its density parameter because there is still a lot of matter present. This issue is known as the *why now* problem. Then, the Hubble parameter (see fig. 1.5) can be approximate considering both as

$$H^2(a) = H_0^2 \left(\Omega_{m,0} a^{-3} + \Omega_{DE,0} a^{-3} \exp\left(-3 \int_{a_0=1}^{a} \frac{w(a')}{a'} \mathrm{d}a' \right) \right) \,, \tag{1.75}$$

where $\Omega_{DE,0} = 8\pi G \rho_{DE,0} / \left(3H_0^2\right) \sim 0.7$ [59] is the normalized dark energy density parameter at nowadays.

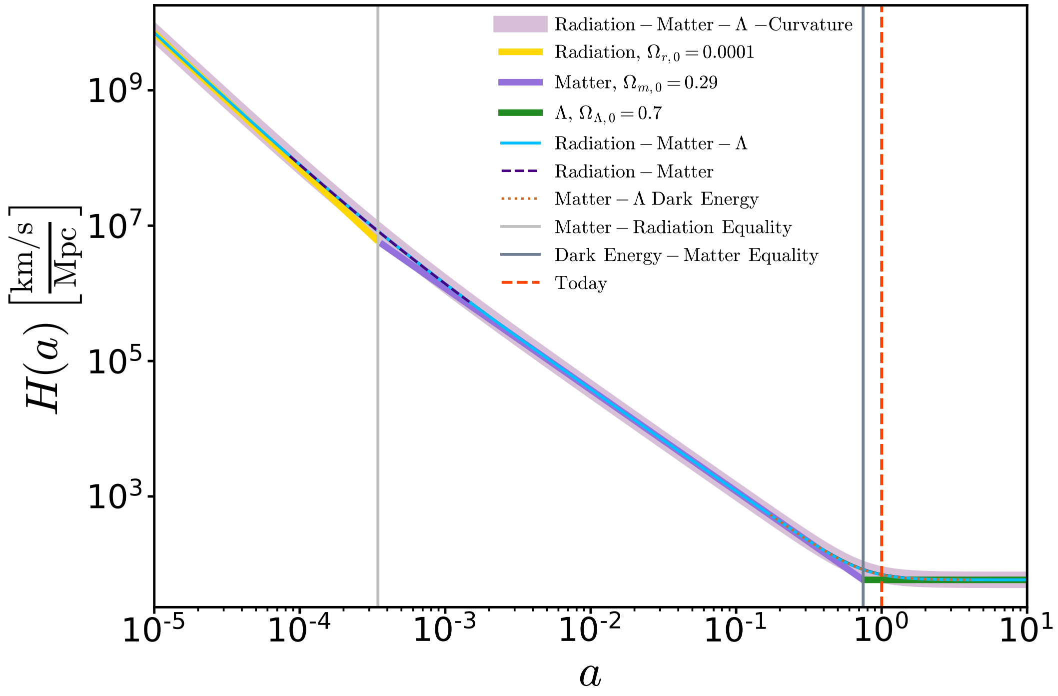


Figure 1.5: The Hubble parameter as function of the scale factor for the ΛCDM model using eqs. (1.31) and (1.72) and considering different regimes and sources -shown in fig. 1.4- as well as the curvature effect.

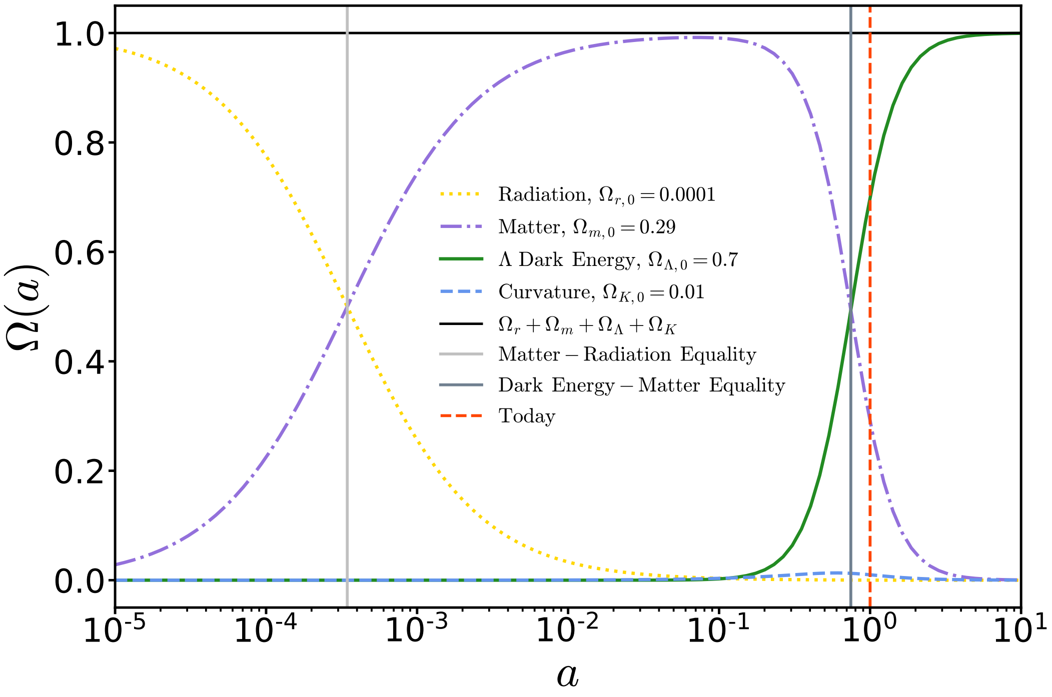


Figure 1.6: The normalized density parameter as function of the scale factor using eqs. (1.35) and (1.36) for different sources -shown in fig. 1.4- as well as the curvature, and using the total Hubble parameter from eq. (1.72), which is shown in fig. 1.5. We also plot the sum of all components shown, which is equal to 1 according to eq. (1.34).

Chapter 2

Basics of Cosmology: Perturbation Theory

Based on the cosmological principle, we have treated our Universe as perfectly homogeneous and isotropic. However, since it is dynamical and governed by gravity and pressure, the perfect homogeneous hypothesis breaks down in its evolution and forces zones of initial overdensity to grow or oscillate. To avoid this, we should have had a perfectly homogeneous Universe in the distant past, but the initial overdensities were of order 1 part in 10^5 [80], which was not small enough. To understand the formation and evolution of large-scale structures, we have to take account of these inhomogeneities, but as long as these perturbations remain small, we can treat them using perturbation theory to first or linear order in Einstein equations. We will base most of the discussions of this chapter on [80, 31, 23].

2.1 The Perturbed Metric

To make perturbation theory in Einstein equations, we need first to consider small perturbations around the FLRW metric described in eq. (1.9). Then a metric that deviates from the FLRW spacetime can be written as the sum of an unperturbed FLRW part $\bar{g}_{\mu\nu}(\tau)$ plus the perturbed part $\delta g_{\mu\nu}(\tau, \vec{x})$:

$$g_{\mu\nu}(\tau, \vec{x}) = \bar{g}_{\mu\nu}(\tau) + \delta g_{\mu\nu}(\tau, \vec{x}) \ . \tag{2.1}$$

The $\delta g_{\mu\nu}$ is also called the perturbed metric, and their entries have to be small compared with those of the zeroth-order part $\bar{g}_{\mu\nu}$ to make perturbation theory. The perturbed metric has 10 independent coefficients since it is a symmetric tensor, and those degrees of freedom can be decomposed into four scalars, two divergenceless vectors, and one traceless, divergenceless rank-2 tensor. This decomposition is known as SVT (scalar, vector, tensor) decomposition, and it is advantageous because the Einstein equations for scalars, vector, and tensors do not mix at linear order, and then we can treat them as separate equations. Note that we are working with the conformal time τ, and from now, we will omit the spacetime $(\tau, \vec{x})$ dependence in the metric and all the perturbed quantities.

To illustrate the STV decomposition we are going to write the matrix form of eq. (2.1) for the case of an FLRW flat background spacetime

$$g_{\mu\nu} = a^2 \begin{pmatrix} -(1+2\Psi) & -B_i \\ -B_j & (1+2\Phi)\delta_{ij} + h_{ij} \end{pmatrix} , \tag{2.2}$$

which defines a line element given by

$$\mathrm{d}s^2 = a^2 \left(-\left(1+2\Psi\right) \mathrm{d}\tau^2 - 2B_i \mathrm{d}x^i \mathrm{d}\tau + \left(\left(1+2\Phi\right)\delta_{ij} + h_{ij}\right) \mathrm{d}x^i \mathrm{d}x^j \right) . \tag{2.3}$$

In eq. (2.2) we have implicitly defined the first two scalar quantities Ψ and Φ with a factor 2 for convenience. We also define a 3-vector B_i and a traceless 3-tensor h_{ij}. From Helmholtz's theorem, we can split any 3-vector into the gradient of a scalar B and a pure divergenceless vector $\hat{B}_i$

$$B_i = \partial_i B + \hat{B}_i , \tag{2.4}$$

with $\partial^i \hat{B}_i = 0$. In the same way, any traceless 3-tensor can be written as

$$h_{ij} = 2 \left(\partial_i \partial_j - \frac{1}{3}\delta_{ij}\nabla^2 \right) E + \left(\partial_i \hat{E}_j + \partial_j \hat{E}_i \right) + 2\hat{E}_{ij} . \tag{2.5}$$

Then we have defined the last scalar quantity E, the other divergenceless vector $\hat{E}_i$ with $\partial^i \hat{E}_i = 0$ and the traceless and divergenceless tensor $\hat{E}_{ij}$ with $\hat{E}^i_i = 0$ and $\partial^i \hat{E}_{ij} = 0$.

The problem with the last decomposition is that the metric perturbations are not uniquely defined because they depend on our choice of coordinates called the gauge choice. For example if we consider the coordinate transformation $X^\mu \longmapsto \tilde{X}^\mu \equiv X^\mu + \xi^\mu \left(\tau, \vec{x}\right)$ where $\xi^0 \equiv T$ and $\xi^i \equiv \partial^i L + \hat{L}^i$ with $\partial^i \hat{L}_i = 0$, the metric in eq. (2.2) will change. To avoid this, we can take two ways; the first is to write the metric using gauge-invariant quantities, which are special combinations of metric perturbations that do not transform under a change of coordinates, an example of these quantities are the Bardeen variables. The second way is to take advantage of the freedom in the gauge functions T and L to set up to four degrees of freedom in the equations.

The vector modes are not of much interest since if they are initially zero, they remain zero, and if they are present initially, they decrease as a^{-1}. The tensor modes are of interest, especially in inflation theories, but since their metric perturbations are gauge invariant, we will not focus on its dynamics. Finally, the scalar modes are of particular interest since they are present in many of the observables that we are going to seek; for that reason, we will focus on them.

2.1.1 Gauge-Invariant Perturbations

To avoid the gauge problem, we will detail the first way that we talked about in the last section. Considering the infinitesimal coordinate transformation

$$x^\mu \longmapsto \tilde{x}^\mu \equiv x^\mu + \xi^\mu \left(\tau, \vec{x}\right) , \tag{2.6}$$

$$\xi^0 \equiv T \left(\tau, \vec{x}\right) , \tag{2.7}$$

$$\xi^i \equiv L^i = \partial^i L \left(\tau, \vec{x}\right) + \hat{L}^i \left(\tau, \vec{x}\right) , \tag{2.8}$$

where we applied SVT decomposition in the vector L^i, then $\hat{L}^i$ is a divergenceless vector $\partial^i \hat{L}_i = 0$. From now we are going to omit the spacetime $(\tau, \vec{x})$ dependency in the variables of coordinate change. If we want to find gauge-invariant quantities we first need to know how the perturbed variables change under that, for that we exploit the invariance of the line element $\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \tilde{g}_{\alpha\beta}\mathrm{d}\tilde{x}^\alpha \mathrm{d}\tilde{x}^\beta$, where we used a different set of dummy variables (α, β) to make clearest the next relation between metrics

$$g_{\mu\nu} = \frac{\partial \tilde{x}^\alpha}{\partial x^\mu}\frac{\partial \tilde{x}^\beta}{\partial x^\nu}\tilde{g}_{\alpha\beta} \,. \tag{2.9}$$

Therefore we can find how the perturbed variables in eq. (2.2) change under the coordinate change in eq. (2.6) component by component using the equations provided by eq. (2.9).

Before developing eq. (2.9) we need to take care about $\tilde{a}(\tilde{\tau}) = a(\tau + T) \simeq a + a'T$, which up to first order can be used in $\tilde{a}^2 \simeq a^2(1 + 2\mathcal{H}T)$. Thus the perturbed variables of eq. (2.2) change like

$$\Psi \longmapsto \tilde{\Psi} = \Psi - T' - \mathcal{H}T \,, \tag{2.10}$$

$$\Phi \longmapsto \tilde{\Phi} = \Phi - \mathcal{H}T - \frac{1}{3}\nabla^2 L \,, \tag{2.11}$$

$$B_i \longmapsto \tilde{B}_i = B_i - \partial_i T + L_i' \,, \tag{2.12}$$

$$h_{ij} \longmapsto \tilde{h}_{ij} = h_{ij} - \left(\partial_i L_j + \partial_j L_i\right) \,, \tag{2.13}$$

where the eqs. (2.12) and (2.13) can be transform using eqs. (2.4), (2.5) and (2.8) to equations in the variables $B, \hat{B}_i, E, \hat{E}_i, \hat{E}_{ij}$

$$B \longmapsto \tilde{B} = B - T + L' \,, \tag{2.14}$$

$$\hat{B}_i \longmapsto \tilde{\hat{B}}_i = \hat{B}_i + \hat{L}_i' \,, \tag{2.15}$$

$$E \longmapsto \tilde{E} = E - L \,, \tag{2.16}$$

$$\hat{E}_i \longmapsto \tilde{\hat{E}}_i = \hat{E}_i - \hat{L}_i \,, \tag{2.17}$$

$$\hat{E}_{ij} \longmapsto \tilde{\hat{E}}_{ij} = \hat{E}_{ij} \,. \tag{2.18}$$

By inspection and a bit of imagination, we can find gauge-invariant variables using the eqs. (2.10) to (2.18), where the main idea is to eliminate the variables of the coordinate change (eqs. (2.7) and (2.8)). The most trivial is $\hat{E}_{ij}$ since it does not change under the change of coordinates, and with a little more effort, we can see that the quantity $\hat{E}_i' + \hat{B}_i$ is also invariant. Bardeen (1980 [28]) identifies two among others, that are called the Bardeen potentials

$$\Phi_A \equiv \Psi - \mathcal{H}\left(B + E'\right) - \left(B + E'\right)' \,, \tag{2.19}$$

$$\Phi_H \equiv -\Phi + \mathcal{H}\left(B + E'\right) + \frac{1}{3}\nabla^2 E \,. \tag{2.20}$$

Note that we can construct other gauge-invariant quantities through linear combinations of the known ones.

2.1.2 The Gauge Fixing

We are going to take the second way of the last section for the gauge problem. For that, we present two gauges; the first is the Synchronous gauge, and the second is the Newtonian gauge in which we are going to work most of the time.

The Synchronous Gauge

In this gauge, the components δg_{00} and δg_{0i} are zero by definition; then, this is equivalent to fix the scalar metric perturbations Ψ, B to zero and also the vector metric perturbation B_i, which completes the four degrees of freedom available. Then the line element is given by

$$\mathrm{d}s^2 = a^2 \left(-\mathrm{d}\tau^2 + (\delta_{ij} + \tilde{h}_{ij})\mathrm{d}x^i \mathrm{d}x^j \right) . \tag{2.21}$$

Note that Φ is absorbed into $\tilde{h}_{ij}$, which is now not traceless anymore. Nevertheless, there are serious disadvantages associated with the synchronous gauge detailed in [160].

The Newtonian Gauge

The Newtonian gauge, also called longitudinal gauge since it is specially used to track scalar modes, is defined by fixing the scalar metric perturbations B, E to zero, which set two of the four degrees of freedom available. In this gauge, we do not worry about the vector and tensor modes, even though we can generalize it to consider them because, as we mentioned, the Einstein equations for scalars, vector, and tensors do not mix at linear order and then they can be treated separately. Thus the line element for the scalar part is given by

$$\mathrm{d}s^2 = a^2 \left(-(1+2\Psi)\mathrm{d}\tau^2 + (1+2\Phi)\delta_{ij}\mathrm{d}x^i \mathrm{d}x^j \right) . \tag{2.22}$$

One advantage of this gauge is that the metric tensor $g_{\mu\nu}$ is diagonal, which simplifies the calculations. Furthermore, Ψ plays the role of the gravitational potential in the Newtonian limit.

The Spatially Flat Gauge

In this gauge, we will set the scalar perturbations in the spatial part to zero *i.e.* $\Phi = E = 0$, using two of the four degrees of freedom. For the moment, we are not concern about vector and tensor perturbations. Thus the line element for the scalar part is given by

$$\mathrm{d}s^2 = a^2 \left(-\left(1+2\Psi\right)\mathrm{d}\tau^2 - 2\partial_i B \mathrm{d}x^i \mathrm{d}\tau + \delta_{ij}\mathrm{d}x^i \mathrm{d}x^j \right) . \tag{2.23}$$

2.2 The Perturbed Energy-Momentum Tensor

In chapter 1, we saw that an essential part of Einstein equations is the energy-momentum tensor, then if we want to make perturbation theory, we need a perturbed expression for it. Since we are considering small perturbations, we would like to write the energy-momentum tensor as

$$T^{\mu}{}_{\nu} = \bar{T}^{\mu}{}_{\nu} + \delta T^{\mu}{}_{\nu} , \tag{2.24}$$

where ${\bar{T}^{\mu}}_{\nu}$ is the unperturbed energy-momentum tensor which for a perfect fluid in a homogeneous and isotropic Universe is given by eq. (1.11) and ${\delta T^{\mu}}_{\nu}$ is the perturbed part that for a perfect fluid can be obtained differentiating eq. (1.11)

$$\delta {T^{\mu}}_{\nu} = (\delta\rho + \delta P)\,\bar{u}^{\mu}\bar{u}_{\nu} + \left(\bar{\rho} + \bar{P}\right)(\delta u^{\mu}\bar{u}_{\nu} + \bar{u}^{\mu}\delta u_{\nu}) + \delta P \delta^{\mu}_{\nu} + {\Pi^{\mu}}_{\nu}\,. \tag{2.25}$$

Now the energy-momentum tensor receives off-diagonal contributions contained in the perturbed four-velocity δu^{μ} and the anisotropic stress ${\Pi^{\mu}}_{\nu}$ which is specially added to take account of these off-diagonal terms in the spatial part and also possible deviations from a perfect fluid at the perturbation level, since in a true perfect fluid it vanishes. Its spatial part trace can always be absorbed into a redefinition of the pressure P, then we can choose the anisotropic stress to be traceless *i.e.* ${\Pi^{i}}_{i} = 0$. We can also choose the tensor ${\Pi^{\mu}}_{\nu}$ orthogonal to u^{μ}, *i.e.* $u^{\mu}\Pi_{\mu\nu} = 0$ and so without loss of generality we can set ${\Pi^{0}}_{0} = {\Pi^{0}}_{i} = 0$.

The perturbations in the four-velocity can induce non-vanishing energy flux $T^{0}_{\ j}$ and momentum density $T^{i}_{\ 0}$. To find these contributions, we need to compute the perturbed four-velocity in the perturbed metric of eq. (2.3). Since $u^{\mu} = \bar{u}^{\mu} + \delta u^{\mu}$, $g_{\mu\nu}u^{\mu}u^{\nu} = -1$, $\bar{g}_{\mu\nu}\bar{u}^{\mu}\bar{u}^{\nu} = -1$ and ignoring non-linear terms, we have

$$\delta g_{\mu\nu}\bar{u}^{\mu}\bar{u}^{\nu} + 2\bar{u}_{\mu}\delta u^{\mu} = 0\,. \tag{2.26}$$

Since in a homogeneous and isotropic universe the relative four-velocity between the fluid and the observer is that for a comoving observer, we have $\bar{u}^{\mu} = a^{-1}(1,0,0,0)$ and $\bar{u}_{\mu} = a(1,0,0,0)$. Then by eq. (2.26) we can find the temporal part of the perturbed four-velocity $\delta u^{0} = -a^{-1}\Psi$ but not the spatial part, which leads us to define $\delta u^{i} \equiv v^{i}/a$. Finally, the four-velocity is written as

$$u^{\mu} = a^{-1}\left(1 - \Psi, v^{i}\right)\,, \tag{2.27}$$

$$u_{\mu} = a\left(-(1+\Psi), v_{i} - B_{i}\right)\,. \tag{2.28}$$

Using the eqs. (2.27) and (2.28) into eqs. (2.24) and (2.25) together with the considerations about the anisotropic stress ${\Pi^{\mu}}_{\nu}$, we find up to linear order

$$T^{0}_{\ 0} = -\left(\bar{\rho} + \delta\rho\right)\,, \tag{2.29}$$

$$T^{i}_{\ 0} = -\left(\bar{\rho} + \bar{P}\right)v^{i}\,, \tag{2.30}$$

$$T^{0}_{\ i} = \left(\bar{\rho} + \bar{P}\right)\left(v_{i} - B_{i}\right)\,, \tag{2.31}$$

$$T^{i}_{\ j} = \left(\bar{P} + \delta P\right)\delta^{i}_{j} + {\Pi^{i}}_{j}\,. \tag{2.32}$$

We can use SVT decomposition in the perturbed quantities that arise as in section 2.1. Since $\delta\rho$ and δP are already scalar, we are left with

$$v_{i} = \partial_{i}v + \hat{v}_{i}\,, \tag{2.33}$$

$$\Pi_{ij} = \left(\partial_{i}\partial_{j} - \frac{1}{3}\delta_{ij}\nabla^{2}\right)\Pi + \frac{1}{2}\left(\partial_{i}\hat{\Pi}_{j} + \partial_{j}\hat{\Pi}_{i}\right) + \hat{\Pi}_{ij}\,, \tag{2.34}$$

with $\partial^{i}\hat{v}_{i} = 0$ and $\partial^{i}\hat{\Pi}_{i} = 0$ since they are divergenceless pure vectors and $\partial^{i}\hat{\Pi}_{ij} = 0$ because it is a divergenceless tensor.

Under the coordinate transformation of eq. (2.6), the energy-momentum also changes. We can see how its components change using the same trick that eq. (2.9), but now with the raised indices

$$T^{\mu}{}_{\nu} = \frac{\partial x^{\mu}}{\partial \tilde{x}^{\alpha}}\frac{\partial \tilde{x}^{\beta}}{\partial x^{\nu}}\tilde{T}^{\alpha}{}_{\beta}\,, \tag{2.35}$$

therefore the perturbed quantities in the energy-momentum tensor change as

$$\delta\rho \longmapsto \delta\tilde{\rho} = \delta\rho - \bar{\rho}' T\,, \tag{2.36}$$

$$\delta P \longmapsto \delta\tilde{P} = \delta P - \bar{P}' T\,, \tag{2.37}$$

$$v_i \longmapsto \tilde{v}_i = v_i + L_i'\,, \tag{2.38}$$

$$\Pi_{ij} \longmapsto \tilde{\Pi}_{ij} = \Pi_{ij}\,, \tag{2.39}$$

where we approximated the background quantities as $\tilde{\bar{\rho}}$ and $\tilde{\bar{P}}$ using $\tilde{\bar{\rho}} \simeq \bar{\rho} + \bar{\rho}' T$ at first order. Using the SVT decomposition we also can find that the variables of v_i (eq. (2.33)) transform as

$$v \longmapsto \tilde{v} = v + L'\,, \tag{2.40}$$

$$\hat{v}_i \longmapsto \tilde{\hat{v}}_i = \hat{v}_i + \hat{L}_i'\,, \tag{2.41}$$

from which we can find another Bardeen or gauge-invariant variables, using eqs. (2.16), (2.36), (2.37) and (2.40)

$$v_s \equiv v + E'\,, \tag{2.42}$$

$$\delta P_{nad} \equiv \delta P - \frac{\bar{P}'}{\bar{\rho}'}\delta\rho\,. \tag{2.43}$$

As we mentioned before we can construct gauge-invariant variables as a linear combination of others, in this case, we present the comoving curvature perturbation which is defined using eqs. (2.20) and (2.42) as

$$\begin{aligned}\mathcal{R} &\equiv \mathcal{H} v_s - \Phi_H \\ &= \Phi + \mathcal{H}\left(v - B\right) - \frac{1}{3}\nabla^2 E\,.\end{aligned} \tag{2.44}$$

2.3 The Perturbed Einstein Tensor

The Einstein tensor $G_{\mu\nu}$, that governs the geometry of spacetime in Einstein equations, depends on the Ricci tensor $R_{\mu\nu}$, which also depends on the connection coefficients or Christoffel symbols $\Gamma^{\alpha}{}_{\mu\nu}$ and these are defined through the metric $g_{\mu\nu}$ as we saw in chapter 1. Then if we want the perturbed Einstein tensor, we first need to compute the Christoffel symbols using a perturbed spacetime. From now on we will work only with scalar quantities and in Newtonian gauge with the metric defined in eq. (2.22), and given by

$$g_{\mu\nu} = a^2 \begin{pmatrix} -(1+2\Psi) & 0 \\ 0 & (1+2\Phi)\delta_{ij} \end{pmatrix}\,. \tag{2.45}$$

The connection coefficients that are defined in eq. (1.8) also require the inverse of the metric $g^{\mu\nu}$. Since the metric in eq. (2.45) is diagonal and considering that $(1+x)^{-1} \simeq 1 - x$ for $x \ll 1$, we have

$$g^{\mu\nu} = a^{-2} \begin{pmatrix} -(1-2\Psi) & 0 \\ 0 & (1-2\Phi)\delta^{ij} \end{pmatrix}\,. \tag{2.46}$$

Substituting eqs. (2.45) and (2.46) into eq. (1.8), gives the perturbed Christoffel symbols for convenient indexes

$$\Gamma^{0}{}_{00} = \mathcal{H} + \Psi' \,, \tag{2.47}$$

$$\Gamma^{0}{}_{0i} = \Gamma^{0}{}_{i0} = \partial_i \Psi \,, \tag{2.48}$$

$$\Gamma^{i}{}_{00} = \partial^i \Psi \,, \tag{2.49}$$

$$\Gamma^{0}{}_{ij} = \mathcal{H}\delta_{ij} + \left(2\mathcal{H}\left\{\Phi - \Psi\right\} + \Phi'\right)\delta_{ij} \,, \tag{2.50}$$

$$\Gamma^{i}{}_{j0} = \Gamma^{i}{}_{0j} = \left(\mathcal{H} + \Phi'\right)\delta^i_j \,, \tag{2.51}$$

$$\Gamma^{i}{}_{jk} = \partial_j \Phi \delta^i_k + \partial_k \Phi \delta^i_j - \partial^i \Phi \delta_{jk} \,. \tag{2.52}$$

Once the connection coefficient are calculated, we have to replace eqs. (2.47) to (2.52) into eq. (1.7) to obtain the Ricci tensor for the desire indexes

$$R_{00} = -3\mathcal{H}' + \nabla^2\Psi - 3\Phi'' + 3\mathcal{H}\left(\Psi' - \Phi'\right) \,, \tag{2.53}$$

$$R_{0i} = -2\partial_i \Phi' + 2\mathcal{H}\partial_i \Psi \,, \tag{2.54}$$

$$\begin{aligned} R_{ij} = &\left\{\mathcal{H}' + 2\mathcal{H}^2 + 2\left(\mathcal{H}' + 2\mathcal{H}^2\right)\left[\Phi - \Psi\right] + 5\mathcal{H}\Phi' - \mathcal{H}\Psi' + \Phi'' - \nabla^2\Phi\right\}\delta_{ij} \\ &- \partial_i\partial_j\left(\Phi + \Psi\right) \,. \end{aligned} \tag{2.55}$$

We can compute the Ricci scalar R using $R = g^{\mu\nu}R_{\mu\nu}$ and the eqs. (2.46) and (2.53) to (2.55):

$$a^2 R = 6\left(\mathcal{H}' + \mathcal{H}^2\right) - 12\left(\mathcal{H}' + \mathcal{H}^2\right)\Psi + 6\mathcal{H}\left(3\Phi' - \Psi'\right) + 6\Phi'' - 2\nabla^2\Psi - 4\nabla^2\Phi \,. \tag{2.56}$$

Finally we can obtain the Einstein tensor $G_{\mu\nu}$ replacing eqs. (2.45) and (2.53) to (2.56) into eq. (1.5) for the same indexes of the Ricci tensor

$$G_{00} = 3\mathcal{H}^2 + 6\mathcal{H}\Phi' - 2\nabla^2\Phi \,, \tag{2.57}$$

$$G_{0i} = 2\partial_i\left(\mathcal{H}\Psi - \Phi'\right) \,, \tag{2.58}$$

$$\begin{aligned} G_{ij} = &\Big\{ -2\mathcal{H}' - \mathcal{H}^2 + 2\left(2\mathcal{H}' + \mathcal{H}^2\right)\left[\Psi - \Phi\right] - 4\mathcal{H}\Phi' + 2\mathcal{H}\Psi' - 2\Phi'' \\ &+ \nabla^2\left(\Phi + \Psi\right)\Big\}\delta_{ij} - \partial_i\partial_j\left(\Phi + \Psi\right) \,. \end{aligned} \tag{2.59}$$

2.4 The Perturbed Einstein Equations

To obtain the first order Einstein equations in our perturbed description we first have to lower the index of eqs. (2.29) to (2.32) through $T_{\mu\nu} = g_{\mu\alpha}T^{\alpha}{}_{\nu}$ and then use them into eq. (1.5), which for the general perturbed metric of eq. (2.2) gives

$$G_{00} = 8\pi G a^2\left(\bar{\rho} + \delta\rho + 2\bar{\rho}\Psi\right) \,, \tag{2.60}$$

$$G_{0i} = G_{i0} = -8\pi G a^2\left\{\left(\bar{\rho} + \bar{P}\right)v_i - \bar{\rho}B_i\right\} \,, \tag{2.61}$$

$$G_{ij} = 8\pi G a^2\left\{\left(\bar{P} + \delta P + 2\bar{P}\Phi\right)\delta_{ij} + \bar{P}h_{ij} + \Pi_{ij}\right\} \,, \tag{2.62}$$

Since we are focused on tracking scalar quantities we are going to set $v_i = \partial_i v$, $B_i = \partial_i B$, $h_{ij} = 2\left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2\right)E$ and $\Pi_{ij} = \left(\partial_i\partial_j - \delta_{ij}/3\nabla^2\right)\Pi$ following the eqs. (2.4), (2.5), (2.33) and (2.34) in eqs. (2.60) to (2.62). However $B = E = 0$ because we are going to work in Newtonian gauge.

To manipulate better the final Einstein equations, we are going to define some quantities:

$$\delta \equiv \frac{\delta\rho}{\bar{\rho}}\,, \tag{2.63}$$

$$\theta \equiv \partial_i v^i = \nabla^2 v\,, \tag{2.64}$$

$$c_s^2 \equiv \frac{\delta P}{\delta\rho} = \frac{\dot{P}}{\dot{\rho}}\,, \tag{2.65}$$

which are respectively the density contrast, the velocity divergence, and the sound speed. The last one is specially defined for barotropic fluids in which P, even when it is perturbed, depends only on ρ. Also, the last equality in the definition of eq. (2.65) is valid only in the FLRW metric because everything at the background level depends only on time. Since c_s, just as w in eq. (1.66), depends at first-order only on backgrounds quantities, the perturbations equations do not introduce any new free function. However, the pressure can depend on the internal degrees of freedom of the fluid, like the entropy s. Then, we need to redefine eq. (2.65) as

$$\begin{aligned} c_s^2 = \frac{\partial P\left(\rho, s\right)}{\partial\rho} &= \frac{\partial P}{\partial\rho} + \frac{\partial P}{\partial s}\frac{\partial s}{\partial\rho} \\ &= c_{s(a)}^2 + c_{s(na)}^2\,, \end{aligned} \tag{2.66}$$

$$c_{s(a)}^2 \equiv \frac{\dot{P}}{\dot{\rho}}\,, \tag{2.67}$$

$$c_{s(na)}^2 \equiv \frac{\partial P}{\partial s}\frac{\partial s}{\partial\rho}\,, \tag{2.68}$$

where $c_{s(a)}$ is known as the adiabatic sound speed and $c_{s(na)}$ as the non-adiabatic sound speed.

Matching the parts of the Einstein tensor from eqs. (2.57) to (2.59) those in eqs. (2.60) to (2.62), we find the perturbed Einstein equations. We start for the (00) component

$$3\mathcal{H}^2 + 6\mathcal{H}\Phi' - 2\nabla^2\Phi = 8\pi G a^2\bar{\rho}\left(1 + \delta + 2\Psi\right)\,, \tag{2.69}$$

where we can identify a zeroth-order part given by

$$\mathcal{H}^2 = \frac{8\pi G}{3}a^2\bar{\rho}\,, \tag{2.70}$$

which is the first Friedman equation, already obtained in eq. (1.31). The first-order part of eq. (2.69) is going to give us the perturbed Einstein equation for this component and after using eq. (2.70) it gives

$$\nabla^2\Phi + 3\mathcal{H}\left(\mathcal{H}\Psi - \Phi'\right) = -4\pi G a^2\bar{\rho}\delta\,. \tag{2.71}$$

The $(0i)$ component has not zeroth-order part since at background it is zero, then its perturbed equation is given by

$$\partial_i\left(\mathcal{H}\Psi - \Phi'\right) = -4\pi G a^2\left(\bar{\rho} + \bar{P}\right)\partial_i v\,, \tag{2.72}$$

where we can leave it in that way or we can integrate it, assuming that all the perturbed quantities decay at infinity, to obtain

$$\mathcal{H}\Psi - \Phi' = -4\pi G a^2\left(\bar{\rho} + \bar{P}\right)v\,, \tag{2.73}$$

but also we can use the operator ∂^i into eq. (2.72) and the definition of eq. (2.64) to obtain

$$\nabla^2 \left(\mathcal{H}\Psi - \Phi'\right) = -4\pi G a^2 \left(\bar{\rho} + \bar{P}\right)\theta\,. \tag{2.74}$$

The reason why we take this procedure will become more clear when we work the equations on Fourier space. Now, we can continue with the (ij) component, and for that we will first take its trace $G^i{}_i = g^{i\mu}G_{\mu i}$, using eq. (2.59) gives

$$\begin{aligned} G^i{}_i = a^{-2}\{&3\left(-2\mathcal{H} - \mathcal{H}^2\right) + 6\left(2\mathcal{H} + \mathcal{H}^2\right)\Psi - 12\mathcal{H}\Phi' + 6\mathcal{H}\Psi' \\ &- 6\Phi'' + 2\nabla^2\left(\Phi + \Psi\right)\}\,, \end{aligned} \tag{2.75}$$

which is equal to $G^i{}_i = 8\pi G T^i_{\ i} = 8\pi G\left(3\left[\bar{P} + \delta P\right]\right)$ according to eq. (2.32) and since $\Pi^i{}_i = 0$. Then the final equations is

$$-\mathcal{H}' - \frac{1}{2}\mathcal{H}^2 + \left(2\mathcal{H}' + \mathcal{H}^2\right)\Psi - 2\mathcal{H}\Phi' + \mathcal{H}\Psi' - \Phi'' + \frac{1}{3}\nabla^2\left(\Phi + \Psi\right) = 4\pi G a^2\left(\bar{P} + \delta P\right)\,, \tag{2.76}$$

where, as before, we can identify a zeroth-order part given by

$$2\mathcal{H}' + \mathcal{H}^2 = -8\pi G a^2 \bar{P}\,, \tag{2.77}$$

that is the second Friedman equation of eq. (1.32). After of using it together with the definitions of eqs. (2.63) and (2.65), the first-order part of eq. (2.76) is

$$\Phi'' - \frac{1}{3}\nabla^2\left(\Phi + \Psi\right) + 2\mathcal{H}\Phi' - \mathcal{H}\Psi' - \left(2\mathcal{H}' + \mathcal{H}^2\right)\Psi = -4\pi G a^2 \bar{\rho} c_s^2 \delta\,. \tag{2.78}$$

Finally, we deal with the off-diagonal terms $(i \neq j)$ which have not zeroth-order part, then the perturbed one is given by

$$-\partial_i\partial_j\left(\Phi + \Psi\right) = 8\pi G a^2\left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2\right)\Pi\,, \tag{2.79}$$

since $i \neq j$ and assuming that all the perturbed quantities decay at infinity we can integrate the eq. (2.79) twice

$$\Phi + \Psi = -8\pi G a^2 \Pi\,, \tag{2.80}$$

which is valid up to a constant that we set to zero since the perturbed quantities have a zero mean by definition.

2.5 The Perturbed Energy-Momentum Conservation

The conservation of energy-momentum tensor is also valid in a perturbed Universe since it is a consequence of the Einstein equations, as we saw in chapter 1. Then we can obtain another set of perturbed equations from eq. (1.38): $\nabla_\mu T^\mu{}_\nu = 0$, starting for $\nu = 0$

$$\nabla_\mu T^\mu{}_0 = \partial_\mu T^\mu{}_0 + \Gamma^\mu{}_{\mu\alpha}T^\alpha{}_0 - \Gamma^\alpha{}_{\mu 0}T^\mu{}_\alpha = 0\,, \tag{2.81}$$

where after of using eqs. (2.29) to (2.32) in Newtonian gauge together with eqs. (2.47) to (2.52) gives the next relation

$$\left(\bar{\rho} + \delta\rho\right)' + \left(\bar{\rho} + \bar{P}\right)\partial_i v^i + 3\left(\bar{\rho} + \bar{P}\right)\left[\mathcal{H} + \Phi'\right] + 3\mathcal{H}\left(\delta\rho + \delta P\right) = 0\,, \tag{2.82}$$

which has a zeroth-order part given by

$$\bar{\rho}' = -3\mathcal{H}\left(\bar{\rho}+\bar{P}\right)\,, \tag{2.83}$$

that is equivalence with eq. (1.39) or also known as continuity equation. The first-order part is then given by

$$\delta\rho' = -\left(\bar{\rho}+\bar{P}\right)\left[\partial_i v^i + 3\Phi'\right] + 3\mathcal{H}\left(\delta\rho+\delta P\right)\,. \tag{2.84}$$

After using the definitions of eqs. (1.66), (2.63) and (2.65) into eq. (2.84), together with eq. (2.83) by noting that $\delta\rho' = (\bar{\rho}\delta)' = \bar{\rho}'\delta + \bar{\rho}\delta'$, we obtain

$$\delta' + 3\mathcal{H}\left(c_s^2 - w\right)\delta + (1+w)\left[\partial_i v^i + 3\Phi'\right] = 0\,, \tag{2.85}$$

which, as we are interested in scalar quantities, is equivalent due eq. (2.33) to

$$\delta' + 3\mathcal{H}\left(c_s^2 - w\right)\delta + (1+w)\left[\nabla^2 v + 3\Phi'\right] = 0\,, \tag{2.86}$$

which is also equivalent, after the definition of eq. (2.64), to

$$\delta' + 3\mathcal{H}\left(c_s^2 - w\right)\delta + (1+w)\left[\theta + 3\Phi'\right] = 0\,. \tag{2.87}$$

All these equations (eqs. (2.84) to (2.87)), depending of your favorite variable choice, are the relativistic version of the continuity equation. The classic version of fluid dynamics can be recovered by taking the limit $P \ll \rho$ and neglecting the Φ' term which is the relativistic correction due to the perturbation to the rate of expansion of space.

Now we can proceed with the component $\nu = i$ in eq. (1.38)

$$\nabla_\mu T^\mu{}_i = \partial_\mu T^\mu{}_i + \Gamma^\mu{}_{\mu\alpha} T^\alpha{}_i - \Gamma^\alpha{}_{\mu i} T^\mu{}_\alpha = 0\,, \tag{2.88}$$

which after setting the Newtonian gauge and using eqs. (2.29) to (2.32), (2.47) to (2.52) and (2.83), we find

$$v_i' + \mathcal{H}\left(1 - 3\frac{\bar{P}'}{\bar{\rho}'}\right) v_i + \frac{\partial_i \delta P}{\bar{\rho}+\bar{P}} + \frac{\partial_j \Pi^j{}_i}{\bar{\rho}+\bar{P}} + \partial_i \Psi = 0\,, \tag{2.89}$$

and writing it with the scalar quantities is

$$\nabla v' + \mathcal{H}\left(1 - 3\frac{\bar{P}'}{\bar{\rho}'}\right)\nabla v + \frac{\nabla \delta P}{\bar{\rho}+\bar{P}} + \frac{2}{3}\frac{\nabla\left(\nabla^2 \Pi\right)}{\bar{\rho}+\bar{P}} + \nabla\Psi = 0\,, \tag{2.90}$$

which, after applying the operator ∂^i to eq. (2.89), is equivalent due eq. (2.64) to

$$\theta' + \mathcal{H}\left(1 - 3\frac{\bar{P}'}{\bar{\rho}'}\right)\theta + \frac{\nabla^2 \delta P}{\bar{\rho}+\bar{P}} + \frac{2}{3}\frac{\nabla^2\left(\nabla^2 \Pi\right)}{\bar{\rho}+\bar{P}} + \nabla^2\Psi = 0\,. \tag{2.91}$$

The eqs. (2.89) to (2.91) are different forms of the relativistic version of Euler equation for a viscous fluid. This equation includes the redshift effect due peculiar velocities in the term $\mathcal{H}v_i$, a small correction for relativistic fluids in the term $\bar{P}'/\bar{\rho}'$ and other for anisotropic fluids in $\partial_j \Pi^j{}_i$. We can recover the classic version in the limit $P \ll \rho$ and ignoring the anisotropic stress.

The eqs. (2.71), (2.74), (2.78), (2.80), (2.87) and (2.91) form a redundant set of equations because the Bianchi identity, but they are consistent and then we can use whichever subsets of our convenience to solve a problem.

2.6 The Perturbed Equations In Fourier Space

The equations found in sections 2.4 and 2.5 are a complex system of partial differential equations (PDEs), and a way to broach them is going to the Fourier space. This means that all the perturbation quantities will be Fourier expanded, turning the system of PDEs into one of ODEs. Since the perturbed Einstein equations and the perturbed energy-momentum tensor conservation equations are linear in the perturbation variables, each plane wave or mode of the expansion will obey the same equation but with a different wavenumber k. The last one is going to define the physical scale of the perturbation λ_p which is defined as

$$\lambda_p \equiv \frac{2\pi}{k} a \,. \tag{2.92}$$

We have to remember that if the perturbation enters into a non-linear regime, this treatment breaks down since it has been developed until first-order perturbations in Einstein equations, where the last ones are still linear.

To go to the Fourier space, we are going to define the Fourier quantities expanding the previous perturbed variables Φ, Ψ, δ and θ or v

$$\Phi\left(\tau, \vec{x}\right) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\, \mathrm{e}^{i\vec{k}\cdot\vec{x}} \phi\big(\tau, \vec{k}\big) \,, \tag{2.93}$$

$$\Psi\left(\tau, \vec{x}\right) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\, \mathrm{e}^{i\vec{k}\cdot\vec{x}} \psi\big(\tau, \vec{k}\big) \,, \tag{2.94}$$

$$\delta\left(\tau, \vec{x}\right) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\, \mathrm{e}^{i\vec{k}\cdot\vec{x}} \delta\big(\tau, \vec{k}\big) \,, \tag{2.95}$$

$$\theta\left(\tau, \vec{x}\right) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\, \mathrm{e}^{i\vec{k}\cdot\vec{x}} \theta\big(\tau, \vec{k}\big) \,, \tag{2.96}$$

$$v\left(\tau, \vec{x}\right) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\, \mathrm{e}^{i\vec{k}\cdot\vec{x}} v\big(\tau, \vec{k}\big) \,. \tag{2.97}$$

We have to note that in Fourier space the potentials Φ and Ψ will be written in lowercase ϕ and ψ, while the perturbed variables δ and θ or v will remain the same and will be identified with a k like δ_k if it is necessary. For the anisotropic stress Π we can do the same expansion

$$\Pi\left(\tau, \vec{x}\right) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\, \mathrm{e}^{i\vec{k}\cdot\vec{x}} \Pi\big(\tau, \vec{k}\big) \,, \tag{2.98}$$

but in Fourier space, another variable is usually used and known as σ, also called anisotropic stress and defined by

$$\left(\bar{\rho} + \bar{P}\right) \sigma\big(\tau, \vec{k}\big) \equiv -\left(\hat{k}_i \hat{k}_j - \frac{1}{3}\delta_{ij}\right) \Pi^i{}_j\big(\tau, \vec{k}\big) \,, \tag{2.99}$$

noting that we do not remove the Fourier space dependence $\vec{k}$ since the definition was made in that space. Then $\Pi^i{}_j$ in eq. (2.99) is actually the Fourier counterpart of our $\Pi^i{}_j$ in eq. (2.34). Finally, the relation between Π and σ in Fourier space is given by

$$\Pi = \frac{3}{2} \frac{\bar{\rho} + \bar{P}}{k^2} \sigma \,. \tag{2.100}$$

Expanding eqs. (2.71), (2.73), (2.74), (2.78) and (2.80) to Fourier space through eqs. (2.93) to (2.99) and using eq. (2.100) make that the perturbed Einstein equations now are respectively

$$k^2\phi + 3\mathcal{H}\left(\phi' - \mathcal{H}\psi\right) = 4\pi G a^2 \bar{\rho}\delta\,, \tag{2.101}$$

$$\phi' - \mathcal{H}\psi = 4\pi G a^2\left(\bar{\rho} + \bar{P}\right) v\,, \tag{2.102}$$

$$k^2\left(\mathcal{H}\psi - \phi'\right) = 4\pi G a^2\left(\bar{\rho} + \bar{P}\right)\theta\,, \tag{2.103}$$

$$\phi'' + \frac{k^2}{3}\left(\phi + \psi\right) + \mathcal{H}\left(2\phi' - \psi'\right) - \left(2\mathcal{H}' + \mathcal{H}^2\right)\psi = -4\pi G a^2 \rho c_s^2 \delta\,, \tag{2.104}$$

$$k^2\left(\phi + \psi\right) = -12\pi G a^2\left(\bar{\rho} + \bar{P}\right)\sigma\,. \tag{2.105}$$

Noting that we drop out the Fourier modes $e^{i\vec{k}\cdot\vec{x}}$ because, like the perturbed Einstein equations are linear, then they decouple between different k-modes.

We continue expanding eqs. (2.87), (2.90) and (2.91) to obtain the perturbed energy-momentum tensor conservation equations in Fourier space

$$\delta' + 3\mathcal{H}\left(c_s^2 - w\right)\delta + (1 + w)\left[\theta + 3\phi'\right] = 0\,, \tag{2.106}$$

$$i\vec{k}\left\{v' + \mathcal{H}\left(1 - 3\frac{\bar{P}'}{\bar{\rho}'}\right)v + \frac{\delta P}{\bar{\rho} + \bar{P}} + \sigma + \psi\right\} = 0\,, \tag{2.107}$$

$$\theta' + \mathcal{H}\left(1 - 3\frac{\bar{P}'}{\bar{\rho}'}\right)\theta - k^2\frac{\delta P}{\bar{\rho} + \bar{P}} - k^2\sigma - k^2\psi = 0\,, \tag{2.108}$$

where the last two equations (eqs. (2.107) and (2.108)) can be rewritten using the relation in eqs. (1.66) and (2.65)

$$i\vec{k}\left\{v' + \mathcal{H}\left(1 - 3c_s^2\right)v + \frac{c_s^2}{1 + w}\delta + \sigma + \psi\right\} = 0\,, \tag{2.109}$$

$$\theta' + \mathcal{H}\left(1 - 3c_s^2\right)\theta - \frac{k^2 c_s^2}{1 + w}\delta - k^2\sigma - k^2\psi = 0\,. \tag{2.110}$$

The final equations to seek are the zeroth-order eqs. (2.70), (2.77) and (2.83), but since they depend only on background quantities, they remain the same in Fourier space.

2.7 Single Perfect Fluid

All the treatments that we made in this chapter have been developed considering only one available source in the energy-momentum tensor; then, if we want to study more that one fluid, we need to modify a bit the equations. However, as the energy-momentum tensor is an additive quantity, we can treat many non-interacting fluids by just sum some of the perturbed variables, as we will see in the next sections. First we are going to try to solve the equations found for only one perfect fluid, so as we mentioned in section 2.2 the anisotropic stress vanishes $\Pi = \sigma = 0$ and this implies by eq. (2.105) that $\psi = -\phi$. Then, and working with the velocity variable θ, the perturbation equations eqs. (2.101), (2.103), (2.104), (2.106) and (2.110) now are

$$k^2\phi + 3\mathcal{H}\phi' + 3\mathcal{H}^2\phi = 4\pi G a^2 \bar{\rho}\delta\,, \tag{2.111}$$

$$k^2\left(\phi'+\mathcal{H}\phi\right)=-4\pi Ga^2\left(\bar{\rho}+\bar{P}\right)\theta\,, \tag{2.112}$$

$$\phi''+3\mathcal{H}\phi'+\left(2\mathcal{H}'+\mathcal{H}^2\right)\phi=-4\pi Ga^2c_s^2\bar{\rho}\delta\,, \tag{2.113}$$

$$\delta'+3\mathcal{H}\left(c_s^2-w\right)\delta+(1+w)\left[\theta+3\phi'\right]=0\,, \tag{2.114}$$

$$\theta'+\mathcal{H}\left(1-3c_s^2\right)\theta-\frac{k^2c_s^2}{1+w}\delta+k^2\phi=0\,. \tag{2.115}$$

We can mix eqs. (2.111) and (2.113) to get an equation only for ϕ

$$\phi''+3\mathcal{H}\left(1+c_s^2\right)\phi'+\left\{\mathcal{H}^2\left(1+3c_s^2\right)+2\mathcal{H}'+c_s^2k^2\right\}\phi=0\,, \tag{2.116}$$

then by eqs. (2.111) and (2.112) we can recover the values for δ and θ. We can also decouple the eqs. (2.114) and (2.115) in the variables δ and θ by deriving them and obtaining second order ODEs but they would still have the ϕ term that can be replace using a combination of eqs. (2.111) and (2.112).

Using the zeroth-order eqs. (2.70) and (2.77) with the definition in eq. (1.66) we find the next relation

$$\mathcal{H}'=-\frac{\mathcal{H}^2}{2}\left(1+3w\right)\,, \tag{2.117}$$

then the eq. (2.116) now can be written as

$$\phi''+3\mathcal{H}\left(1+c_s^2\right)\phi'+\left\{3\mathcal{H}^2\left(c_s^2-w\right)+c_s^2k^2\right\}\phi=0\,. \tag{2.118}$$

If the pressure depends only on the density (barotropic fluids) we have $\mathrm{d}P=w\mathrm{d}\rho+\rho\mathrm{d}w$ by eq. (1.66) and using eq. (2.83) we obtain

$$\mathrm{d}P=\left(w-\frac{w'}{3\mathcal{H}\left(1+w\right)}\right)\mathrm{d}\rho\,, \tag{2.119}$$

then for a constant equation of state parameter we have $c_s^2=w$ due eq. (2.65), which is valid for both matter and radiation. For this case eq. (2.118) becomes

$$\phi''+3\mathcal{H}\left(1+w\right)\phi'+wk^2\phi=0\,. \tag{2.120}$$

2.7.1 Super-Horizon Scales

For large scales $k<<\mathcal{H}$, also known as super-horizon scales or super-horizon limit, the physical wavelength λ_p of perturbations in eq. (2.92) is much larger than the Hubble radius H^{-1}. Then we can approximate the term k^2 to zero in eq. (2.120) and obtain an equation for ϕ due a fluid with constant w

$$\phi''+3\mathcal{H}\left(1+w\right)\phi'=0\,, \tag{2.121}$$

in which $\phi'=0$ is a solution and then $\phi=$ constant. But as eq. (2.121) is a second-order ODE we must have another solution given by

$$\phi=\phi_0\int_{\tau_0}^{\tau}\left(\frac{a(\tau')}{a_0}\right)^{-3(1+w)}\mathrm{d}\tau'\,. \tag{2.122}$$

Since we are tracking growing modes and the solution in eq. (2.122) is a decaying one, at least for $w > -1$, we can ignore it. Then the dominating solution is $\phi = cte, \phi' = 0$ and after of replacing it in eq. (2.111) and using eq. (2.70) we find

$$\delta = \frac{2}{3}\frac{k^2}{\mathcal{H}^2}\phi + 2\phi\,, \tag{2.123}$$

where the first term vanishes since we are working in the super-horizon limit, implying that also $\delta = 2\phi =$ constant on large scales.

2.7.2 Sub-Horizon Scales

In contrast, we have the sub-horizon limit or small scales in which $k >> \mathcal{H}$, *i.e.* scales deep inside the Hubble radius. Then, for a fluid with constant w, we have to work with eq. (2.120), without any approximation yet. We are going to make a change of variables between the conformal time and the number of e-folds N, through

$$N \equiv \ln a\,, \tag{2.124}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau} = \mathcal{H}\frac{\mathrm{d}}{\mathrm{d}N}\,, \tag{2.125}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} = \left(\mathcal{H}^2 + \mathcal{H}'\right)\frac{\mathrm{d}}{\mathrm{d}N} + \mathcal{H}^2\frac{\mathrm{d}^2}{\mathrm{d}N^2}\,. \tag{2.126}$$

Then eq. (2.120) transforms, with the help of eq. (2.117), into

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}N^2} + \frac{1}{2}(3w+7)\frac{\mathrm{d}\phi}{\mathrm{d}N} + \frac{c_s^2k^2}{\mathcal{H}^2}\phi = 0\,, \tag{2.127}$$

which is the equation for an oscillator with a damped term but with variable frequency, so the growing mode is more difficult to track by a general way. We also can find an equation for δ in this regime, since by eq. (2.112) $\phi' + \mathcal{H}\phi = -4\pi Ga^2\left(\bar{\rho} + \bar{P}\right)\theta/k^2 \to 0$, we can replace it into eq. (2.111) and obtain the equation

$$k^2\phi \simeq 4\pi Ga^2\bar{\rho}\delta\,, \tag{2.128}$$

which, by eq. (2.70) becomes in

$$k^2\phi = \frac{3}{2}\mathcal{H}^2\delta\,. \tag{2.129}$$

Then we can replace ϕ into eq. (2.120) and using both eq. (2.117) and its derivative

$$\mathcal{H}'' = \frac{\mathcal{H}^3}{2}(1+3w)^2\,, \tag{2.130}$$

we find the equation for δ

$$\delta'' + \mathcal{H}(1-3w)\delta' + \left[\frac{3}{2}\mathcal{H}^2(w-1)(1+3w) + wk^2\right]\delta = 0\,. \tag{2.131}$$

Note that while eq. (2.117) is valid for any function w, in the derivation of eq. (2.130) we assumed that w is constant. Now, if we study eq. (2.131) taking the Minkowski limit $\mathcal{H} \to 0$, then the equation reduces to the classical fluid wave equation $\delta'' + wk^2\delta = 0$. To obtain growing modes we require that the factor of δ in eq. (2.131) be zero or negative.

2.7.3 Perturbations During RDE

During the radiation domination era, we have $w = 1/3$, as we saw in section 1.4.1. We are going to use the eq. (2.117), which have the exact solution

$$\mathcal{H} = \frac{2}{(1+3w)\,\tau}\,, \tag{2.132}$$

under the right initial conditions. Then, in RDE we have $\mathcal{H} = 1/\tau$ and eq. (2.120) now is

$$\phi'' + \frac{4}{\tau}\phi' + \frac{k^2}{3}\phi = 0\,, \tag{2.133}$$

which has the following exact solution

$$\phi\,(\tau, k) = A_k \frac{j_1(x)}{x} + B_k \frac{n_1(x)}{x}\,, \tag{2.134}$$

where $x \equiv k\tau/\sqrt{3}$, $j_1(x)$ and $n_1(x)$ are the spherical Bessel and Neumann functions of order 1, respectively. The constants A_k and B_k can depend on k since the amplitude can be different for each mode. Anyway we set $B_k = 0$ because the spherical Neumann function $n_1(x)$ blows up for small x (early times). We can also replace the spherical Bessel function using its form with sinusoidal functions

$$\phi\,(\tau, k) = A_k \left(\frac{\sin x - x\cos x}{x^3}\right)\,. \tag{2.135}$$

For super-horizon scales $x \ll 1$, the solution in eq. (2.135) approaches to ϕ =constant (see fig. 2.1) while on sub-horizon scales $x \gg 1$ we get

$$\phi\,(\tau, k) \simeq A_k \frac{\cos x}{x^2}\,, \tag{2.136}$$

meaning that in RDE the sub-horizon modes of ϕ oscillate with frequency $k/\sqrt{3}$ and their amplitudes decay as τ^{-2} (see fig. 2.1).

2.7.4 Perturbations During MDE

In the matter domination era we have $w \simeq 0$, and then by eq. (2.132) $\mathcal{H} = 2/\tau$. The eq. (2.120) now is

$$\phi'' + \frac{6}{\tau}\phi' = 0\,, \tag{2.137}$$

which is an Euler-Cauchy equation with an exact solution given by

$$\phi\,(\tau, k) = A_k + B_k\tau^{-5}\,. \tag{2.138}$$

Then in MDE ϕ has a very quickly decreasing mode and a growing mode, this last one ensures that the potential is frozen on all scales during MDE (see fig. 2.1).

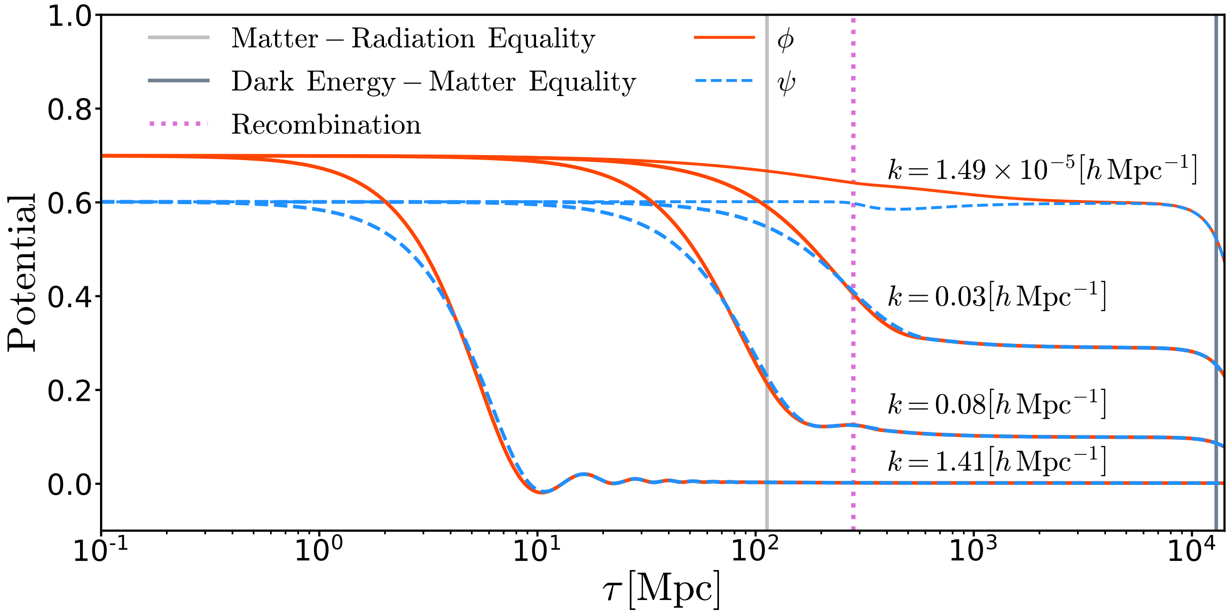


Figure 2.1: The evolution of the potentials ϕ and ψ for four scales computed using the software CLASS [46], which follow the behavior of eqs. (2.135), (2.136) and (2.138). We also plot the time of densities equality to determinate RDE and MDE, and the recombination time (see section 3.4.3). These potentials differ at the beginning because the neutrino decoupling (see section 3.4), but they get close to each other quickly, considering that the $x-$axis is logarithm.

2.8 Multiple Fluids

We know that our Universe is not as simple as we have tried to explain *i.e.* with only one fluid, but it seems that it helps during some phases and conditions. Our Universe is a complex mix of components that can also interact; then, if we want to have a better description at the perturbation level, we should consider how the number of fluids changes the equations found in this chapter. We are going to start with the perturbed Einstein equations and, as we saw, the energy-momentum tensor sources them; then, when we have more that one contribution the total energy-momentum tensor, due to its additivity, is just

$$T_{\mu\nu}^{(total)} = \sum_I T_{\mu\nu}^{(I)}\,, \tag{2.139}$$

which is valid for I independent components. At the background level, the quantities are additive

$$\bar{\rho}^{(total)} = \sum_I \bar{\rho}^{(I)} \tag{2.140}$$

$$\bar{P}^{(total)} = \sum_I \bar{P}^{(I)} = \sum_I w_I \bar{\rho}^{(I)} \tag{2.141}$$

then we can have the total perturbed quantities searching component by component in the energy-momentum tensor and using eqs. (2.29) to (2.32) and (2.100). For the scalar variables we get

$$\delta\rho^{(total)} = \sum_I \delta\rho^{(I)} \tag{2.142}$$

$$\delta P^{(total)} = \sum_I \delta P^{(I)} \tag{2.143}$$

$$\left(\bar{\rho}^{(total)} + \bar{P}^{(total)}\right) v^{(total)} = \sum_I \left(\bar{\rho}^{(I)} + \bar{P}^{(I)}\right) v^{(I)} \tag{2.144}$$

$$\left(\bar{\rho}^{(total)} + \bar{P}^{(total)}\right) \theta^{(total)} = \sum_I \left(\bar{\rho}^{(I)} + \bar{P}^{(I)}\right) \theta^{(I)} \tag{2.145}$$

$$\Pi^{(total)} = \sum_I \Pi^{(I)} \tag{2.146}$$

$$\left(\bar{\rho}^{(total)} + \bar{P}^{(total)}\right) \sigma^{(total)} = \sum_I \left(\bar{\rho}^{(I)} + \bar{P}^{(I)}\right) \sigma^{(I)} \,. \tag{2.147}$$

The eqs. (2.71), (2.73), (2.74), (2.78) and (2.80) in real space and the eqs. (2.101) to (2.105) in Fourier space, now are sourced by the total variables defined in eqs. (2.142) to (2.147). The equations obtained from the energy-momentum conservation can be split into several equations, one for each component, since the eq. (1.38) has to be valid for each contribution as along as they do not interact like we mentioned. Then eqs. (2.87), (2.90) and (2.91) in real space or eqs. (2.106), (2.109) and (2.110) in Fourier space, form a system with $2N$ equations, where N is the number of components.

For the total variable δ we need to take some considerations, by eqs. (2.140) and (2.142) we find

$$\begin{aligned}
\delta^{(total)} = \frac{\delta\rho^{(total)}}{\bar{\rho}^{(total)}} &= \frac{\sum_I \delta\rho^{(I)}}{\sum_I \bar{\rho}^{(I)}} \\
&= \frac{8\pi G}{3H^2\Omega^{(total)}} \sum_I \delta\rho^{(I)} \\
&= \frac{1}{\Omega^{(total)}} \sum_I \frac{8\pi G}{3H^2} \delta\rho^{(I)} \\
&= \frac{\sum_I \Omega^{(I)}\delta^{(I)}}{\Omega^{(total)}} ,
\end{aligned} \tag{2.148}$$

where we used the definition in eq. (1.35), then

$$\Omega^{(total)}\delta^{(total)} = \sum_I \Omega^{(I)}\delta^{(I)} \,, \tag{2.149}$$

which since $\Omega^{(total)} = 1$ by eq. (1.34), we have

$$\delta^{(total)} = \sum_I \Omega^{(I)}\delta^{(I)} \,. \tag{2.150}$$

Doing the same procedure used in eqs. (2.148) to (2.150) we can find an expression for the total w, also called the effective equation of state $w_{\text{eff}} \equiv \bar{P}^{(total)}/\bar{\rho}^{(total)}$

$$w_{\text{eff}} = \sum_I \Omega^{(I)} w^{(I)} \,, \tag{2.151}$$

which can be used into eq. (2.145) to obtain

$$\theta^{(total)} = \frac{\sum_I \left(1 + w^{(I)}\right) \Omega^{(I)}\theta^{(I)}}{1 + w_{\text{eff}}} \,. \tag{2.152}$$

Adiabatic Fluctuations

The principal idea in the definition of adiabatic fluctuations is that the perturbed local state, determined by its energy density and pressure at some spacetime point $(\tau, \vec{x})$ is the same as in the background universe but at some slightly different time $\tau + \delta\tau\left(\vec{x}\right)$. We can see this as some parts in the Universe are late, and others ahead respect to the cosmic evolution or background. If the Universe has multiple fluids, the adiabatic perturbations condition implies that the perturbations are induced by a common and local shift in time $(\delta\tau)$. We can seek this behavior in the density contrast by

$$\begin{aligned}\delta\rho^{(I)}\left(\tau, \vec{x}\right) &= \rho^{(I)}\left(\tau, \vec{x}\right) - \bar{\rho}^{(I)}\left(\tau\right) \\ &= \bar{\rho}^{(I)}\left(\tau + \delta\tau\left(\tau, \vec{x}\right)\right) - \bar{\rho}^{(I)}\left(\tau\right) \\ &= \bar{\rho}'^{(I)}\delta\tau\left(\vec{x}\right)\,,\end{aligned} \tag{2.153}$$

where in the second line, we used the definition of adiabatic perturbations, and in the third line, we used a Taylor expansion assuming a small $\delta\tau$. Like $\delta\tau$ is the same for all species, we have

$$\delta\tau = \frac{\delta\rho^{(I)}}{\bar{\rho}'^{(I)}} = \frac{\delta\rho^{(J)}}{\bar{\rho}'^{(J)}}\,, \tag{2.154}$$

which after of use eq. (2.83) for each component, because we assumed that there is no energy transfer between those fluids, we obtain

$$\frac{\delta^{(I)}}{1 + w^{(I)}} = \frac{\delta^{(J)}}{1 + w^{(J)}}\,. \tag{2.155}$$

Isocurvature Fluctuations

The isocurvature fluctuations are defined as the complement of adiabatic fluctuations. Then, the isocurvature perturbations correspond to those in which the energy density variations for one component do not necessarily correspond with the energy density variations in other components. We can define

$$S_{IJ} \equiv \frac{\delta^{(I)}}{1 + w^{(I)}} - \frac{\delta^{(J)}}{1 + w^{(J)}}\,, \tag{2.156}$$

which for adiabatic perturbations $S_{IJ} = 0$ for all species I and J.

2.8.1 Two Fluids: Matter And Radiation

We are going to study a more realistic case in which two fluids are present, these are going to be matter and radiation. The solution for two fluids problems are in general complicated but under certain assumptions we can make it simpler. For that we are going to introduce the matter perturbation variables δ_m, θ_m $(w_m = c^2_{s_m} = 0)$ and the radiation perturbation variables δ_r, θ_r $(w_r = c^2_{s_r} = 1/3)$. We also are going to assume that there is no interaction between these species (which is true since $z \approx 1000$) and that they behave like perfect fluids which implies that the energy-momentum tensor has not anisotropic shear, then $\sigma_m = \sigma_r = 0$. Finally we obtain, in Fourier space, a set of gravitationally coupled eqs. (2.101) to (2.106), (2.109) and (2.110) using the indications of section 2.8

$$\delta'_m + \theta_m + 3\phi' = 0\,, \tag{2.157}$$

$$\delta_r' + \frac{4}{3}\theta_r + 4\phi' = 0\,, \tag{2.158}$$

$$\theta_m' + \mathcal{H}\theta_m + k^2\phi = 0\,, \tag{2.159}$$

$$\theta_r' - \frac{k^2}{4}\delta_r + k^2\phi = 0\,, \tag{2.160}$$

$$k^2\phi + 3\mathcal{H}\left(\phi' + \mathcal{H}\phi\right) = 4\pi G a^2\left(\bar{\rho}_m + \bar{\rho}_r\right)\delta^{(total)}\,, \tag{2.161}$$

$$k^2\left(\phi' + \mathcal{H}\phi\right) = -4\pi G a^2\left(1 + w_{\rm eff}\right)\left(\bar{\rho}_m + \bar{\rho}_r\right)\theta^{(total)}\,, \tag{2.162}$$

$$\phi'' + 3\mathcal{H}\phi' + \left(2\mathcal{H}' + \mathcal{H}^2\right)\phi = -4\pi G a^2 \delta P^{(total)}\,, \tag{2.163}$$

where the total perturbation variables, according to eqs. (2.143) and (2.150) to (2.152), are

$$\delta^{(total)} = \Omega_m\delta_m + \Omega_r\delta_r\,, \tag{2.164}$$

$$\begin{aligned}\delta P^{(total)} &= \delta P_r + \delta P_m \\ &= c_{s,r}^2\delta\rho_r + c_{s,m}^2\delta\rho_m \\ &= \frac{1}{3}\bar{\rho}_r\delta_r\,,\end{aligned} \tag{2.165}$$

$$\begin{aligned}\theta^{(total)} &= \frac{\left(1 + w_m\right)\Omega_m\theta_m + \left(1 + w_r\right)\Omega_r\theta_r}{1 + w_{eff}} \\ &= \frac{4/3\Omega_r\theta_r + \Omega_m\theta_m}{1 + w_{eff}}\,,\end{aligned} \tag{2.166}$$

$$\begin{aligned}w_{eff} &= \Omega_r w_r + \Omega_m w_m \\ &= \frac{1}{3}\Omega_r\,.\end{aligned} \tag{2.167}$$

In the super-horizon limit, the eq. (2.160) reduces to $\theta_r' = 0$ and then $\theta_r =$constant. Also eq. (2.159) becomes in $\theta_m' + \mathcal{H}\theta = 0$, where we can find a simpler solution after using the definition of $\mathcal{H}$ in eq. (1.4)

$$\begin{aligned}\theta_m &= \theta_{m,0}\mathrm{e}^{-\int \mathcal{H}\mathrm{d}\tau} \\ &= \frac{\theta_{m,0}}{a}\,.\end{aligned} \tag{2.168}$$

While in the sub-horizon limit eq. (2.162) can be approximated as $\phi' + \mathcal{H}\phi = 0$, which result is already known

$$\phi = \frac{\phi_0}{a}\,, \tag{2.169}$$

then we can replace eq. (2.169) into eq. (2.159), integrate it and obtain θ_m. Using the same technique, we can replace ϕ and θ_m into eq. (2.157), integrate it and obtain a solution for δ_m. For δ_r, we find in this regime due eq. (2.160) that $\delta_r \simeq 4\phi$ and by eq. (2.158) we can find directly a solution for $\theta_r \simeq 6\phi'$.

During RDE we have $\Omega_m \sim 0$ and $\Omega_r \sim 1$, then $\delta^{(total)} \sim \delta_r$,$w_{eff} \sim 1/3$ and $\theta^{(total)} \sim \theta_r$ due eqs. (2.164), (2.166) and (2.167). The potential ϕ is sourced only by radiation perturbations, so it is already known from eq. (2.135). Thus the equations for δ_r, θ_r, ϕ form a closed system and the equations for δ_m, θ_m are independent but forced by ϕ, which implies that the matter perturbations are forced by radiation perturbations. The real behavior during this epoch is more complicated since now we are not considering that a fraction of the matter is interacting strongly with the radiation.

For MDE, we have the opposite behavior since now the potential is sourced mostly by matter perturbations. Then, it is constant due to the solution in eq. (2.138). Now the radiation perturbations are forced by the matter perturbations.

2.9 Interacting Fluids

In the last sections, we consider the equations that govern our Universe until first-order perturbations, but always assuming one or more fluids with null anisotropic shear $\Pi^{i}{}_{j} = 0$, which is true for species that do not interact, except gravitationally. For the late Universe and on cosmological scales, we can assume that the species are coupled purely with the metric and ignore the anisotropic stress in the Einstein general relativity framework. Nevertheless, our Universe was more complicated, and we know that there were phases where its components had interactions, such as the process of creation/annihilation and particle collision in the early Universe. These processes need a better description that takes us away from the ideal fluids that we have characterized, mainly because components like photons and neutrinos develop anisotropic shear under certain circumstances. Then, if we want to seek those effects we need a more general form of the energy-momentum tensor (eq. (1.11)) or its first-order expansion (eq. (2.24)), which for a fluid in the full general relativistic framework is

$$T^{\mu}{}_{\nu} = g_i \int \frac{\mathrm{d}P_1 \mathrm{d}P_2 \mathrm{d}P_3}{(2\pi)^3\sqrt{-g}} \frac{P^{\mu} P_{\nu}}{P^0} f\left(P_j, \vec{x}, t\right) \,, \tag{2.170}$$

where $P^{\mu} \equiv \mathrm{d}x^{\mu}/\mathrm{d}\lambda$ is the four-momentum and λ the parameter that characterizes the particle path. While $f\left(P_j, \vec{x}, t\right)$ is the distribution function which is going to take in consideration all the interactions, also it is a scalar quantity, invariant under canonical transformations and phase space-dependent, but it only depends in the spacial part of the four-momentum P_j since the mass constraint $P^2 \equiv P^{\mu}P_{\mu} = -m^2$. Finally, g_i is the number of spin states available for each specie and $g \equiv \det g_{\alpha\beta}$ is the determinant of the metric, which we will work in Newtonian gauge (eq. (2.22)) but using cosmic time instead conformal time

$$\mathrm{d}s^2 = -\left(1 + 2\Psi\right)\mathrm{d}t^2 + a^2\left(1 + 2\Phi\right)\delta_{ij}\mathrm{d}x^i \mathrm{d}x^j \,, \tag{2.171}$$

noting that for the moment, we are working in real space.

From eq. (2.170) we can note that the anitropic shear terms $T^{i}_{\ j}$ come from

$$T^{i}_{\ j} \propto \frac{P^i P_j}{\left(P^0\right)^2} = v^i v_j \,, \tag{2.172}$$

which is a second-order quantity for massive particles, and then it is excluded in our treatment. Nevertheless, for massless particles or relativistic ones, we should consider it, which is the case of photons and neutrinos. Then, we are going to study them and also their behavior when processes of Compton scattering with electrons are present. For that, we need to take account of them into the distribution function which obeys the Boltzmann equation

$$\frac{\mathrm{d}f}{\mathrm{d}t} = C[f] \,, \tag{2.173}$$

where $C[f]$ is the called collision functional term.

2.9.1 Photons

The photons obey $P^2 = 0$ since they are massless particles and by definition we have $P^2 = g_{00}\left(P^0\right)^2 + g_{ij}P^iP^j = -\left(1+2\Psi\right)\left(P^0\right)^2 + p^2$, where $p^2 \equiv g_{ij}P^iP^j$. Then we can find an expression for P^0

$$\begin{aligned} P^0 &= \frac{p}{\sqrt{1+2\Psi}} \\ &= p\left(1+\Psi\right) , \end{aligned} \tag{2.174}$$

valid until first-order. We can also write the spatial part as $P^i = A\hat{p}^i$, where A is a proportionally constant to determinate and $\hat{p}^i$ is an unit vector with the same direction of the momentum. Thus we have the relation $p^2 = g_{ij}\hat{p}^i\hat{p}^jA^2 = a^2\left(1+2\Phi\right)\delta_{ij}\hat{p}^i\hat{p}^jA^2 = a^2\left(1+2\Phi\right)A^2$, and then

$$A = \frac{p}{a}\left(1-\Phi\right) . \tag{2.175}$$

The final four-momentum for massless particles is

$$P^\mu = \left(p\left(1+\Psi\right), \frac{p}{a}\left(1-\Phi\right)\hat{p}^i\right) . \tag{2.176}$$

As we saw, the four-momentum can be expressed only in terms of p and $\hat{p}^i$, then the l.h.s. of eq. (2.173) is

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x^i}\cdot\frac{\mathrm{d}x^i}{\mathrm{d}t} + \frac{\partial f}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\partial f}{\partial \hat{p}^i}\cdot\frac{\mathrm{d}\hat{p}^i}{\mathrm{d}t} , \tag{2.177}$$

where the last term is a second-order quantity since $\partial f/\partial\hat{p}^i$ and $\mathrm{d}\hat{p}^i/\mathrm{d}t$ are of first-order, then it does not contribute at first order in perturbation theory. The second term in eq. (2.177) can be expressed using $P^i \equiv \mathrm{d}x^i/\mathrm{d}\lambda$, $P^0 \equiv \mathrm{d}t/\mathrm{d}\lambda$ and eq. (2.176)

$$\begin{aligned} \frac{\mathrm{d}x^i}{\mathrm{d}t} &= \frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}\lambda}{\mathrm{d}t} \\ &= \frac{P^i}{P^0} \\ &= \frac{\hat{p}^i}{a}\left(1+\Psi+\Phi\right) , \end{aligned} \tag{2.178}$$

and since $\partial f/\partial x^i$ is a first-order term, eq. (2.177) now can be written as

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \frac{\hat{p}^i}{a}\cdot\frac{\partial f}{\partial x^i} + \frac{\partial f}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} . \tag{2.179}$$

The last term in eq. (2.179) can be rewritten using the geodesic equation

$$\frac{\mathrm{d}P^\mu}{\mathrm{d}\lambda} = -\Gamma^\mu{}_{\alpha\beta}P^\alpha P^\beta , \tag{2.180}$$

together with eqs. (1.8), (2.47), (2.48), (2.50) and (2.178) to obtain

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -p\left[H + \frac{\partial\Phi}{\partial t} + \frac{\hat{p}^i}{a}\cdot\frac{\partial\Psi}{\partial x^i}\right] , \tag{2.181}$$

then eq. (2.179) now is

$$\frac{\mathrm{d}f}{\mathrm{d}t}=\frac{\partial f}{\partial t}+\frac{\hat{p}^i}{a}\cdot\frac{\partial f}{\partial x^i}-p\frac{\partial f}{\partial p}\left[H+\dot{\Phi}+\frac{\hat{p}^i}{a}\cdot\frac{\partial\Psi}{\partial x^i}\right]\,. \tag{2.182}$$

The first two terms in eq. (2.179) are familiar from hydrodynamics, given the continuity and Euler equation for photons when they are integrated with no collision term ($C[f]=0$). The third term says that photons lose energy in an expanding Universe, and the last two terms encode how the overdense regions affect the photon distribution function.

Now we have to treat the distribution function of photons $f\left(t,\vec{x},p,\hat{p}\right)$, which its zero-order part is given by the Bose-Einstein distribution since photons are bosons

$$f^{(0)}(t,p)=\frac{1}{\exp\{(\varepsilon-\mu)/T\}-1}\,, \tag{2.183}$$

where μ is the chemical potential that we will not use $\mu=0$. Also as photons are massless particles we have $\varepsilon^2=p^2$. We are going to expand eq. (2.183) through the temperature, since in a smooth Universe it is a function of time only, then to consider the inhomogeneities and anisotropies we will write it as $T\left(t,\vec{x},p\hat{p}\right)=\bar{T}(t)+\delta T\left(t,\vec{x},p\hat{p}\right)=\bar{T}(t)\left(1+\Theta\left(t,\vec{x},p\hat{p}\right)\right)$, where we have defined the temperature perturbation Θ as

$$\Theta\equiv\frac{\delta T}{\bar{T}}\,. \tag{2.184}$$

Expanding eq. (2.183) we have $f\simeq\bar{f}+\frac{\partial f}{\partial T}\delta T$, which is equal to

$$f\simeq\bar{f}-p\frac{\partial\bar{f}}{\partial p}\Theta\,, \tag{2.185}$$

since $T\frac{\partial\bar{f}}{\partial T}=-p\frac{\partial\bar{f}}{\partial p}$ and where $\bar{f}$ is

$$\bar{f}\equiv\frac{1}{\exp\left(p/T\right)-1}\,. \tag{2.186}$$

Computing the time derivative of the zeroth-order part of eq. (2.183) through eqs. (2.179) and (2.186) we find

$$\begin{aligned}\left.\frac{\mathrm{d}f}{\mathrm{d}t}\right|_{0\text{th}}&=\frac{\partial\bar{f}}{\partial t}-Hp\frac{\partial\bar{f}}{\partial p}\\&=\left(-\frac{\mathrm{d}T/\mathrm{d}t}{T}-\frac{\mathrm{d}a/\mathrm{d}t}{a}\right)\frac{\partial\bar{f}}{\partial p}\,,\end{aligned} \tag{2.187}$$

which for collisionless processes, it is zero since $C[f]=0$; then, in that case, the solution of eq. (2.187) is

$$T=\frac{a_0}{a}\,, \tag{2.188}$$

as we expect in the smooth background Universe. Note that set $C[f]=0$ in the zeroth-order part is not trivial, but as we will see, the collision term for the processes of interest is proportional to Θ and other perturbed quantities, then it will not enter in the zeroth-order equation.

Now we are going to compute the time derivative of the first-order photon distribution function. We will use the eq. (2.179)

$$\left.\frac{\mathrm{d}f}{\mathrm{d}t}\right|_{\mathrm{1st}} = -p\frac{\partial \bar{f}}{\partial p}\left(\dot{\Theta} + \frac{\hat{p}^i}{a}\cdot\frac{\partial\Theta}{\partial x^i} + \dot{\Phi} + \frac{\hat{p}^i}{a}\cdot\frac{\partial\Psi}{\partial x^i}\right) \,, \tag{2.189}$$

where the first two terms take account for the "free streaming" of photons and the last two for gravity effects.

One process of interest is the Compton scattering in which electrons scatter photons

$$\mathrm{e}^-\left(\vec{q}\,\right) + \gamma\left(\vec{p}\,\right) \leftrightarrow \mathrm{e}^-\left(\vec{q}\,'\right) + \gamma\left(\vec{p}\,'\right) \,, \tag{2.190}$$

where $\vec{q}, \vec{p}, \vec{q}\,', \vec{p}\,'$ are the momentum of each particle. Then if we are interested in the collision term for the photon with momentum $\vec{p}$ we must sum over all other momenta. Schematically it is given by

$$\begin{aligned} C\left[f\left(\vec{p}\,\right)\right] = \sum_{\vec{q},\vec{q}\,',\vec{p}\,'} \Big\{ & |\mathcal{M}|^2_{\rightarrow} f_{\mathrm{e}}\left(\vec{q}\,'\right) f\left(\vec{p}\,'\right)\left[1 \pm f\left(\vec{p}\,\right)\right]\left[1 \pm f_{\mathrm{e}}\left(\vec{q}\,\right)\right] \\ & - |\mathcal{M}|^2_{\leftarrow} f_{\mathrm{e}}\left(\vec{q}\,\right) f\left(\vec{p}\,\right)\left[1 \pm f\left(\vec{p}\,'\right)\right]\left[1 \pm f_{\mathrm{e}}\left(\vec{q}\,'\right)\right] \Big\} \,, \end{aligned} \tag{2.191}$$

where the $[1 \pm f]$ terms represent the phenomena of Bose enhancement and Pauli blocking; then for bosons, we must use the $+$ sing and for fermions the $-$ sign. However, in our treatment, since we are doing first-order perturbation theory, we will consider the factor 1 only in those terms. The f_{e} is the distribution function for electrons and as eq. (2.190) is reversible, the amplitudes are equal $|\mathcal{M}|^2_{\rightarrow} = |\mathcal{M}|^2_{\leftarrow} = |\mathcal{M}|^2$. Replacing the sum by integrals in phase space, the collision term takes the form

$$\begin{aligned} C\left[f\left(\vec{p}\,\right)\right] = \frac{1}{p}\int\frac{\mathrm{d}^3q}{(2\pi)^3 2\varepsilon_{\mathrm{e}}(q)}\int\frac{\mathrm{d}^3q'}{(2\pi)^3 2\varepsilon_{\mathrm{e}}(q')}\int\frac{\mathrm{d}^3p'}{(2\pi)^3 2\varepsilon(p')}|\mathcal{M}|^2(2\pi)^4\delta^3\left(\vec{p} + \vec{q} - \vec{p}\,' - \vec{q}\,'\right) \\ \delta\left(\varepsilon(p) + \varepsilon_{\mathrm{e}}(q) - \varepsilon(p') - \varepsilon_{\mathrm{e}}(q')\right)\left[f_{\mathrm{e}}\left(\vec{q}\,'\right) f\left(\vec{p}\,'\right) - f_{\mathrm{e}}\left(\vec{q}\,\right) f\left(\vec{p}\,\right)\right] \,. \end{aligned} \tag{2.192}$$

To explain how the terms in eq. (2.192) appear, we are going to start with the $(2\pi)^3$ factors or $(2\pi\hbar)^3$ with $\hbar = 1$. These represent the volume of one unit in the phase space, and then we want to sum over all such units. The factors 2ε or $2\varepsilon_{\mathrm{e}}$ arise because we have to integrate over a four-dimensional phase space due to our general relativity framework. However, these dimensions are constrained on shell, making appear

$$\int \mathrm{d}^3p \int \mathrm{d}\varepsilon\,\delta(\varepsilon^2 - p^2 - m^2) = \int \mathrm{d}^3p \int \mathrm{d}\varepsilon\frac{\delta\left(\varepsilon - \sqrt{p^2 + m^2}\right)}{2\varepsilon} \,. \tag{2.193}$$

Also, due to the general relativity treatment, we should have to define the eq. (2.173) using the path parameter λ instead of the cosmic time. Moreover, that change of variables makes appear the factor $1/p$ at the beginning of the r.h.s. of eq. (2.192). Finally, the Dirac delta functions enforce energy and momentum conservation, and the factor $(2\pi)^4$ is the result of moving from discrete Kronecker deltas to continuous Dirac deltas.

Computing the eq. (2.192) in the non-relativistic limit for electrons and considering that under these assumptions the energy transferred between electrons and photons is little (both valid in the cosmological process involved), the integral of eq. (2.192) becomes in

$$C[f(\vec{p})] = \frac{\pi}{4m_e^2 p}\int \mathrm{d}^3q \frac{f_e(\vec{q})}{(2\pi)^3}\int \frac{\mathrm{d}^3p'}{(2\pi)^3 p'}|\mathcal{M}|^2 \left\{\delta(p-p') + \frac{(\vec{p}-\vec{p}\,')\cdot\vec{q}}{m_e}\frac{\partial\delta(p-p')}{\partial p'}\right\} \\ [f(\vec{p}\,') - f(\vec{p})]\,, \tag{2.194}$$

where m_e is the electron mass. We need the amplitude of the Compton scattering $|\mathcal{M}|^2$, which can be computed using fundamental physics through Feynman rules [40]. Its value is given by

$$|\mathcal{M}|^2 = 8\pi\sigma_T m_e^2\,, \tag{2.195}$$

here σ_T is the Thomson cross-section. The expression in eq. (2.195) is not totally true because the amplitude has both angular and polarization dependence and its real value is given by $|\mathcal{M}|^2 = 6\pi\sigma_T m_e^2\left(1+\cos^2(\hat{p}\cdot\hat{p}')\right)$, but we are not interested in those effects now. Nevertheless they have to be considered if we want to study anisotropies. Replacing eq. (2.195) into eq. (2.194) and developing the integrals, up to first-order in energy transfer, gives us the final collision term

$$C[f(\vec{p})] = -p\frac{\partial\bar{f}}{\partial p}n_e\sigma_T\left(\Theta_0 - \Theta(\hat{p}) + \hat{p}\cdot\vec{v}_b\right)\,, \tag{2.196}$$

where n_e is the electron density and $\vec{v}_b$ is the velocity field of electrons defined as

$$n_e \equiv \int\frac{\mathrm{d}^3q}{(2\pi)^3}f_e\,, \tag{2.197}$$

$$\vec{v}_e \equiv \frac{1}{n_e}\int\frac{\mathrm{d}^3q}{(2\pi)^3}\frac{\vec{q}}{\varepsilon_e}f_e\,, \tag{2.198}$$

noting that for non-relativistic electrons $\varepsilon_e \simeq m_e$. We also assumed that electrons have the same velocity of baryons, and then we use the subscript b. Finally, Θ_0 is the monopole part of temperature perturbations, defined as

$$\Theta_0(\vec{x},t) \equiv \frac{1}{4\pi}\int \mathrm{d}\Omega'\Theta(\hat{p}\,',\vec{x},t)\,. \tag{2.199}$$

The monopole Θ_0 does not depend on the direction vector $\hat{p}'$, since we are integrating over all the directions, but it has spacial dependence and then it is not a zero-order quantity. In the absence of bulk velocity ($v_b = 0$), eq. (2.196) tells us that the collision term serves to drive Θ to Θ_0, which for a very efficient Compton scattering means that only the monopole survives. When electrons carry a bulk velocity, the photons will also have a dipole moment, fixed by the electron velocity, but all higher moments will vanish. This is equivalent to saying that the photons behave like a fluid. In conclusion, the strong scattering or tight coupling produces that the photons and electrons behave as a single fluid.

Since eq. (2.196) contains only first-order quantities, we can link it with eq. (2.189) through eq. (2.173) to obtain

$$\dot{\Theta} + \frac{\hat{p}^i}{a}\cdot\frac{\partial\Theta}{\partial x^i} + \dot{\Phi} + \frac{\hat{p}^i}{a}\cdot\frac{\partial\Psi}{\partial x^i} = n_e\sigma_T\left(\Theta_0 - \Theta + \hat{p}\cdot\vec{v}_b\right)\,, \tag{2.200}$$

or its counterpart using conformal time

$$\Theta' + \hat{p}^i \cdot \frac{\partial \Theta}{\partial x^i} + \Phi' + \hat{p}^i \cdot \frac{\partial \Psi}{\partial x^i} = n_{\mathrm{e}} \sigma_T a \left(\Theta_0 - \Theta + \hat{p} \cdot \vec{v}_b \right) . \tag{2.201}$$

Now we are going to pass to the Fourier space, and then we need the Fourier expansion of the temperature perturbations

$$\Theta \left(\tau, \vec{x} \right) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \mathrm{e}^{i\vec{k}\cdot\vec{x}} \Theta\big(\tau, \vec{k}\,\big) , \tag{2.202}$$

while the perturbed quantities $\vec{v}_b = v_b \hat{v}$, Ψ and Φ have their Fourier expansion in eqs. (2.93), (2.94) and (2.97) respectively. However, we must take care from now, because there could be a confusion between the $v \to v_{\mathrm{now}} = ||\vec{v}||$ presented above and $v \to v_{\mathrm{before}}$ defined in eq. (2.33), the relation between both in real and Fourier space is

$$\vec{v} = v_{\mathrm{now}} \hat{v} = \partial_i v_{\mathrm{before}} \Leftrightarrow \vec{v}_f = v_{\mathrm{now};f} \hat{v}_f = ik v_{\mathrm{before};f} . \tag{2.203}$$

We also define the cosine of the angle α between $\vec{k}$ and $\hat{p}$ as

$$\mu \equiv \frac{\vec{k} \cdot \hat{p}}{k} , \tag{2.204}$$

which for irrotational fluids, we can assume that is the same as the angle between the velocity field $\vec{v}_b$ and the direction of the photon momentum $\hat{p}$, then $\vec{v}_b \cdot \hat{p} = \mu v_b$. Finally, the eq. (2.201) in Fourier space takes the form

$$\Theta' + ik\mu\Theta + \phi' + ik\mu\psi = -\tau'_{op} \left(\Theta_0 - \Theta + \mu v_b \right) , \tag{2.205}$$

where τ_{op} is defined as

$$\tau_{op} \equiv \int_{\tau}^{\tau_0} n_{\mathrm{e}} \sigma_T a \mathrm{d}\tau' . \tag{2.206}$$

The eq. (2.205) is not entirely correct, since we ignored the angular dependence of Compton scattering and the polarization effects that also influence the anisotropies in the photon distribution. If we had considered the first, eq. (2.205) would have a dependency on Θ_2, which is the quadrupole or the second moment. In general, we can define the lth multipole moment of the temperature field as

$$\Theta_l \equiv i^l \int_{-1}^{1} \frac{\mathrm{d}\mu}{2} \mathcal{P}_l \left(\mu \right) \Theta , \tag{2.207}$$

where $\mathcal{P}_l$ is the Legendre polynomial of order l and in contrast with eq. (2.199), we omitted the spacetime variables in eq. (2.207) since it can be used in real or Fourier space. If we had also incorporated the polarization effects, an equation for the strength of the polarization Θ_P would have appeared, which describes the change in the polarization field. Considering those effects, the equations for the photon distribution are

$$\Theta' + ik\mu\Theta + \phi' + ik\mu\psi = -\tau'_{op} \left(\Theta_0 - \Theta + \mu v_b - \frac{1}{2} \mathcal{P}_2(\mu) \Pi \right) , \tag{2.208}$$

$$\Pi = \Theta_2 + \Theta_{P2} + \Theta_{P0} , \tag{2.209}$$

$$\Theta'_P + ik\mu\Theta_P = -\tau'_{op} \left(-\Theta_P + \frac{1}{2} \left[1 - \mathcal{P}_2(\mu) \right] \Pi \right) . \tag{2.210}$$

2.9.2 Baryons

The baryons also require a set of Boltzmann equations since, as we saw, the photons interact with electrons. These last ones are not baryons (in fact, they are leptons), but in Cosmology, they are often called using that nomenclature and grouped with protons.

In section 2.9.1, we studied a process in which the electrons are coupled with photons because we will see that it is crucial at some epoch of our Universe. Nonetheless, we did not mention that during that period the electrons were also coupled with the protons by Coulomb scattering

$$\mathrm{e}^-\left(\vec{q}\right)+p^+\left(\vec{Q}\right)\leftrightarrow \mathrm{e}^-\left(\vec{q}\,'\right)+p^+\left(\vec{Q}\,'\right)\,, \tag{2.211}$$

where $\vec{Q},\vec{Q}\,'$ are the momentum of the protons before and after the collision, respectively. The rate in this process is much larger than the expansion rate, at least at the epoch of interest [80]. This tight coupling forces that the electron and proton overdensities take a common "baryon" value $\delta_b=\delta_\mathrm{e}=\delta_p$. The same happens with the velocities of these two species $\vec{v}_b=\vec{v}_\mathrm{e}=\vec{v}_p$, fact that we already used in section 2.9.1 when we called the electron velocity field as $\vec{v}_b$.

Since baryons are massive particles they obey $P^2=-m_l^2$ and $\varepsilon_l^2=p_l^2+m_l^2$, where $l=\mathrm{e},p$ indicates if we are talking about a proton or an electron. We are going to drop out that index for the moment but later we will recover it, specially in equations where the two species mix. Now we have $P^2=g_{00}\left(P^0\right)^2+p^2=-\left(1+2\Psi\right)\left(P^0\right)^2=-m^2$, which lead us to $P^0=\varepsilon\left(1-\Psi\right)$. The spatial part of the four-momentum is the same as eq. (2.176) since the treatment is independent of the relation $P^\mu P_\mu=-m^2$, so the final four-momentum for baryons is

$$P^\mu=\left(\varepsilon\left(1-\Psi\right),\frac{p}{a}\left(1-\Phi\right)\hat{p}^i\right)\,. \tag{2.212}$$

Now, we are going to use ε as an independent variable. Thus, the total time derivative of some baryon distribution (similar to eq. (2.177)) can be written as

$$\frac{\mathrm{d}f}{\mathrm{d}t}=\frac{\partial f}{\partial t}+\frac{\partial f}{\partial x^i}\cdot\frac{\mathrm{d}x^i}{\mathrm{d}t}+\frac{\partial f}{\partial\varepsilon}\frac{\mathrm{d}\varepsilon}{\mathrm{d}t}+\frac{\partial f}{\partial\hat{p}^i}\cdot\frac{\mathrm{d}\hat{p}^i}{\mathrm{d}t}\,, \tag{2.213}$$

where once again, the last term vanishes since it is the product of two first-order terms. Following the same procedure of eqs. (2.179) and (2.182), the eq. (2.213) becomes in

$$\frac{\mathrm{d}f}{\mathrm{d}t}=\frac{\partial f}{\partial t}+\frac{p}{\varepsilon}\frac{\hat{p}^i}{a}\cdot\frac{\partial f}{\partial x^i}-\frac{p^2}{\varepsilon}\frac{\partial f}{\partial\varepsilon}\left[H+\dot{\Phi}+\frac{\varepsilon}{p}\frac{\hat{p}^i}{a}\cdot\frac{\partial\Psi}{\partial x^i}\right]\,, \tag{2.214}$$

which in the massless limit $\varepsilon\to p$ we recover the eq. (2.182).

The eq. (2.214) is the l.h.s. of eq. (2.173), which is valid for both protons and electrons as long as we change their masses. For protons, the next step is to compute the collision term $C[f_p]$ of the Coulomb scattering. For the electrons, the collision term $C[f_\mathrm{e}]$ now depends on two processes: Coulomb scattering and Compton scattering. Since they are independent we simply add them as $C[f_\mathrm{e}]=C_{\mathrm{e}^-p^+}[f_\mathrm{e}]+C_{\mathrm{e}^-\gamma}[f_\mathrm{e}]$. These assumptions are not trivial because two reasons, the first is that in principle, we should take account of the scattering between protons and photons in the collision term for protons, but in practice, the cross-section of that process is much smaller than for Compton scattering. The second reason is that we should include ionization and recombination terms in both species, but we are going to treat all electrons as ionized [80]. Thus the Boltzmann equations for protons and electrons are

$$\frac{\mathrm{d}f_p}{\mathrm{d}t} = C_{\mathrm{e}^-p^+}[f_p]\,, \tag{2.215}$$

$$\frac{\mathrm{d}f_\mathrm{e}}{\mathrm{d}t} = C_{\mathrm{e}^-p^+}[f_\mathrm{e}] + C_{\mathrm{e}^-\gamma}[f_\mathrm{e}]\,. \tag{2.216}$$

Considering the processes in eqs. (2.190) and (2.211), the collision terms in eqs. (2.215) and (2.216) are given by (which are obtained in a similar way to the eq. (2.192))

$$\begin{aligned} C_{\mathrm{e}^-\gamma}\left[f_\mathrm{e}\left(\vec{q}\right)\right] = \frac{1}{q}\int\frac{\mathrm{d}^3p}{(2\pi)^3 2\varepsilon(p)}\int\frac{\mathrm{d}^3q'}{(2\pi)^3 2\varepsilon_\mathrm{e}(q')}\int\frac{\mathrm{d}^3p'}{(2\pi)^3 2\varepsilon(p')}|\mathcal{M}|^2(2\pi)^4\delta^3\left(\vec{p}+\vec{q}-\vec{p}\,'-\vec{q}\,'\right) \\ \delta\left(\varepsilon(p)+\varepsilon_\mathrm{e}(q)-\varepsilon(p')-\varepsilon_\mathrm{e}(q')\right)\left[f_\mathrm{e}\left(\vec{q}\,'\right)f\left(\vec{p}\,'\right)-f_\mathrm{e}\left(\vec{q}\right)f\left(\vec{p}\right)\right]\,, \end{aligned} \tag{2.217}$$

$$\begin{aligned} C_{\mathrm{e}^-p^+}\left[f_\mathrm{e}\left(\vec{q}\right)\right] = \frac{1}{q}\int\frac{\mathrm{d}^3Q}{(2\pi)^3 2\varepsilon_p(Q)}\int\frac{\mathrm{d}^3q'}{(2\pi)^3 2\varepsilon_\mathrm{e}(q')}\int\frac{\mathrm{d}^3Q'}{(2\pi)^3 2\varepsilon_p(Q')}|\mathcal{M}|^2_{\mathrm{e}^-p^+}(2\pi)^4 \\ \delta^3\left(\vec{Q}+\vec{q}-\vec{Q}\,'-\vec{q}\,'\right)\delta\left(\varepsilon_p(Q)+\varepsilon_\mathrm{e}(q)-\varepsilon_p(Q')-\varepsilon_\mathrm{e}(q')\right) \\ \left[f_\mathrm{e}\left(\vec{q}\,'\right)f_p\left(\vec{Q}\,'\right)-f_\mathrm{e}\left(\vec{q}\right)f_p\left(\vec{Q}\right)\right]\,, \end{aligned} \tag{2.218}$$

$$\begin{aligned} C_{\mathrm{e}^-p^+}\left[f_p\left(\vec{Q}\right)\right] = \frac{1}{Q}\int\frac{\mathrm{d}^3q}{(2\pi)^3 2\varepsilon_\mathrm{e}(q)}\int\frac{\mathrm{d}^3q'}{(2\pi)^3 2\varepsilon_\mathrm{e}(q')}\int\frac{\mathrm{d}^3Q'}{(2\pi)^3 2\varepsilon_p(Q')}|\mathcal{M}|^2_{\mathrm{e}^-p^+}(2\pi)^4 \\ \delta^3\left(\vec{Q}+\vec{q}-\vec{Q}\,'-\vec{q}\,'\right)\delta\left(\varepsilon_p(Q)+\varepsilon_\mathrm{e}(q)-\varepsilon_p(Q')-\varepsilon_\mathrm{e}(q')\right) \\ \left[f_\mathrm{e}\left(\vec{q}\,'\right)f_p\left(\vec{Q}\,'\right)-f_\mathrm{e}\left(\vec{q}\right)f_p\left(\vec{Q}\right)\right]\,, \end{aligned} \tag{2.219}$$

where $|\mathcal{M}|^2_{\mathrm{e}^-p^+}$ is the amplitude of the Coulomb scattering.

Now, instead of give some form of the distribution function for electrons or protons, we are going to integrate eq. (2.216) in the $q-$space. Then using eq. (2.214), the integrated l.h.s. of eq. (2.216) after replace p by q is

$$\begin{aligned} \int\frac{\mathrm{d}^3q}{(2\pi)^3}\frac{\mathrm{d}f_\mathrm{e}}{\mathrm{d}t} =& \frac{\partial}{\partial t}\int\frac{\mathrm{d}^3q}{(2\pi)^3}f_\mathrm{e} + \frac{1}{a}\frac{\partial}{\partial x^i}\cdot\int\frac{\mathrm{d}^3q}{(2\pi)^3}\frac{q\hat{q}^i}{\varepsilon_\mathrm{e}}f_\mathrm{e} - \left[H+\dot{\Phi}\right]\int\frac{\mathrm{d}^3q}{(2\pi)^3}\frac{q^2}{\varepsilon_\mathrm{e}}\frac{\partial f_\mathrm{e}}{\partial\varepsilon_\mathrm{e}} \\ &- \frac{1}{a}\frac{\partial\Psi}{\partial x^i}\cdot\int\frac{\mathrm{d}^3q}{(2\pi)^3}q\hat{q}^i\frac{\partial f_\mathrm{e}}{\partial\varepsilon_\mathrm{e}}\,, \end{aligned} \tag{2.220}$$

noting that the integral over q passes through $H, \dot{\Phi}$ and $\nabla\Psi$, and the partial derivatives with respect to x^i and t since they are independent variables. The last term of eq. (2.220) vanishes in first-order perturbation theory since the integral over the direction vector is nonzero only for the perturbed part, then the integral is a first-order term, and it multiplies the first-order variable $\nabla\Psi$ giving a second-order quantity. Thus eq. (2.220) becomes in

$$\int\frac{\mathrm{d}^3q}{(2\pi)^3}\frac{\mathrm{d}f_\mathrm{e}}{\mathrm{d}t} = \frac{\partial n_\mathrm{e}}{\partial t} + \frac{1}{a}\frac{\partial\left(n_\mathrm{e}v^i_\mathrm{e}\right)}{\partial x^i} + 3\left[H+\dot{\Phi}\right]n_\mathrm{e}\,, \tag{2.221}$$

where n_e and v_e are defined in eqs. (2.197) and (2.198) respectively, and the last term was computed using spherical coordinates together with the relation $\varepsilon_\mathrm{e}\mathrm{d}\varepsilon_\mathrm{e} = q\mathrm{d}q$ and integrating by parts.

To construct the zeroth-order and first order-part of the perturbed equations we are going to expand the density number n in its zeroth-order part $\bar{n}\,(t)$ and its perturbed part $\delta n\,(t, \vec{x})$

$$\begin{aligned} n &\equiv \bar{n} + \delta n \\ &= \bar{n}\left[1 + \frac{\delta n}{\bar{n}}\right] \\ &= \bar{n}\left[1 + \delta\right] \,, \end{aligned} \tag{2.222}$$

where the last equality holds because $\rho = mn$, $\bar{\rho} = m\bar{n}$ and using the definition of eq. (2.63). Collecting the zeroth-order terms in eq. (2.221) and setting $C[f] = 0$ at this order as we argued in section 2.9.1, we find that the integrated zeroth-order part of eq. (2.216) is

$$\frac{\partial \bar{n}}{\partial t} + 3H\bar{n} = 0\,, \tag{2.223}$$

which solution is known $\bar{n} = \bar{n}_0 a^{-3}$ because it is the same as $\bar{\rho}$ in eq. (1.71) for matter. Then the equation eq. (2.223) is also valid for protons and matter in general as along as the collision term has not zeroth-order part.

Now collecting the terms up to first-order in eq. (2.221), we find

$$\int \frac{\mathrm{d}^3 q}{(2\pi)^3}\frac{\mathrm{d}f_\mathrm{e}}{\mathrm{d}t}\bigg|_{\mathrm{1st}} = \dot{\bar{n}}\delta + \bar{n}\dot{\delta} + \frac{\bar{n}}{a}\frac{\partial v^i}{\partial x^i} + 3\bar{n}\dot{\Phi} + 3H\bar{n}\delta\,, \tag{2.224}$$

where we can use the eq. (2.223) to replace $\dot{\bar{n}}$ and then obtain

$$\int \frac{\mathrm{d}^3 q}{(2\pi)^3}\frac{\mathrm{d}f}{\mathrm{d}t}\bigg|_{\mathrm{1st}} = \bar{n}\left(\dot{\delta} + \frac{1}{a}\frac{\partial v^i}{\partial x^i} + 3\dot{\Phi}\right)\,, \tag{2.225}$$

where we omitted the index e since it can also be used for protons (l.h.s. of eq. (2.215)) using the appropriate subscripts, this treatment is known as the extraction of the zeroth moment of the Boltzmann equation.

Instead of treat the collision term to complete the eq. (2.225), we are going back to the unintegrated eq. (2.214) and we will extract its first moment. For electrons, we multiply that equation by $\vec{q}/m_\mathrm{e} = q\hat{q}^j/m_\mathrm{e}$ to later integrate it in the $q-$space

$$\begin{aligned} \int \frac{\mathrm{d}^3 q}{(2\pi)^3}\frac{\mathrm{d}f_\mathrm{e}}{\mathrm{d}t}\frac{\vec{q}}{m_\mathrm{e}} =& \frac{\partial}{\partial t}\int \frac{\mathrm{d}^3 q}{(2\pi)^3} f_\mathrm{e}\frac{q\hat{q}^j}{m_\mathrm{e}} + \frac{1}{a}\frac{\partial}{\partial x^i}\cdot\int \frac{\mathrm{d}^3 q}{(2\pi)^3}\frac{q^2\hat{q}^i\hat{q}^j}{m_\mathrm{e}\varepsilon_\mathrm{e}} f_\mathrm{e} - \left[H + \dot{\Phi}\right]\int \frac{\mathrm{d}^3 q}{(2\pi)^3}\frac{q^3\hat{q}^j}{m_\mathrm{e}\varepsilon_\mathrm{e}}\frac{\partial f_\mathrm{e}}{\partial \varepsilon_\mathrm{e}} \\ & - \frac{1}{a}\frac{\partial \Psi}{\partial x^i}\cdot\int \frac{\mathrm{d}^3 q}{(2\pi)^3}\frac{q^2\hat{q}^i\hat{q}^j}{m_\mathrm{e}}\frac{\partial f_\mathrm{e}}{\partial \varepsilon_\mathrm{e}}\,. \end{aligned} \tag{2.226}$$

Since electrons behaves as non-relativistic matter, we have that $\varepsilon_\mathrm{e} \simeq m_\mathrm{e}$, and also, this allows us to neglect all the terms of order $v^2 = (q/\varepsilon_\mathrm{e})^2$ and higher in first-order perturbation theory. Thus the second term in the r.h.s. of eq. (2.226) vanishes since it has a term of order $q^2/(\varepsilon_{\rm e} m_{\rm e}) \sim (q/\varepsilon_{\rm e})^2$, while the first term is just the time derivative of $n_{\rm e} v_{\rm e}^j$ as we saw in section 2.9.1. The integral in the third term has to be done by parts, and after neglecting some terms, its value is $-4 n_{\rm e} v_{\rm e}^j$, thus the only term that multiplies H survives. Finally, with the last term, we must do the same as with the third, obtaining $-n_{\rm e}\delta^{ij}$. So eq. (2.226) becomes in

$$\int \frac{{\rm d}^3 q}{(2\pi)^3}\frac{{\rm d}f_{\rm e}}{{\rm d}t}\frac{\vec{q}}{m_{\rm e}} = \frac{\partial (n_{\rm e} v_{\rm e}^j)}{\partial t} + 4H n_{\rm e} v_{\rm e}^j + \frac{n_{\rm e}}{a}\frac{\partial \Psi}{\partial x^j}\,, \tag{2.227}$$

which has not zeroth-order part since it depends purely of first-order quantities. This also implies that we need to replace $n_{\rm e}$ by $\bar{n}_{\rm e}$ and $\dot{\bar{n}}_{\rm e}$ by eq. (2.223), which give us

$$\int \frac{{\rm d}^3 q}{(2\pi)^3}\frac{{\rm d}f}{{\rm d}t}\frac{\vec{q}}{m} = \bar{n}\left(\dot{v}^j + H v^j + \frac{1}{a}\frac{\partial \Psi}{\partial x^j}\right)\,, \tag{2.228}$$

where once again, we omitted the index e as in eq. (2.225) since it can also be applied for protons because they behave as non-relativistic particles too.

Using the conformal time, the eqs. (2.225) and (2.228) take the form

$$\int \frac{{\rm d}^3 q}{(2\pi)^3}\frac{{\rm d}f}{{\rm d}\tau}\bigg|_{\rm 1st} = \bar{n}\left(\delta' + \frac{\partial v^i}{\partial x^i} + 3\Phi'\right)\,, \tag{2.229}$$

$$\int \frac{{\rm d}^3 q}{(2\pi)^3}\frac{{\rm d}f}{{\rm d}\tau}\frac{\vec{q}}{m} = \bar{n}\left(v'^j + \mathcal{H} v^j + \frac{\partial \Psi}{\partial x^j}\right)\,. \tag{2.230}$$

In Fourier space the eqs. (2.229) and (2.230) are

$$\int \frac{{\rm d}^3 q}{(2\pi)^3}\frac{{\rm d}f}{{\rm d}\tau}\bigg|_{\rm 1st} = \bar{n}\left(\delta' + ikv + 3\phi'\right)\,, \tag{2.231}$$

$$\int \frac{{\rm d}^3 q}{(2\pi)^3}\frac{{\rm d}f}{{\rm d}\tau}\frac{\vec{q}}{m} = \bar{n}\left(v' + \mathcal{H} v + ik\psi\right)\,, \tag{2.232}$$

with δ, v, ϕ, ψ defined in eqs. (2.93) to (2.95) and (2.97) respectively. Note that if we set the collision term to zero in both eqs. (2.231) and (2.232), *i.e.* non-interacting matter, their l.h.s. vanish and we recover the eqs. (2.106) and (2.109) for matter ($w = c_s = 0$).

It is time to face the collision terms of eqs. (2.217) to (2.219), which need to be integrated following the treatment that we made to the l.h.s. of eq. (2.214). We are going to start integrating completely the eq. (2.216), for that we rewrite eq. (2.221) using the definition of $n_{\rm e}$ and the eqs. (2.217) and (2.218)

$$\frac{{\rm d}n_{\rm e}}{{\rm d}t} = \frac{\partial n_{\rm e}}{\partial t} + \frac{1}{a}\frac{\partial\left(n_{\rm e} v_{\rm e}^i\right)}{\partial x^i} + 3\left[H + \dot{\Phi}\right] n_{\rm e} = \langle C_{{\rm e}^- p^+}\rangle_{\vec{q}} + \langle C_{{\rm e}^- \gamma}\rangle_{\vec{q}}\,, \tag{2.233}$$

where $\langle ... \rangle_{\vec{q}}$ indicates that the collision terms of eqs. (2.217) and (2.218) are being integrated in $q-$ space. Nevertheless, note that in eq. (2.233) both collision terms vanish, intuitively this is because the processes that we are considering conserve the electron number and therefore ${\rm d}n_{\rm e}/{\rm d}t$ must be zero. A more careful explanation lies in the fact that the integrated collision terms are completely symmetric under momentum interchange, but their intergrands are antisymmetric. Thus the l.h.s. of eqs. (2.225), (2.229) and (2.231) is zero, and the equation for electrons, protons or baryons, since they have the same density contrast and velocity, in Fourier space is

$$\delta_b' + ikv_b + 3\phi' = 0\,. \tag{2.234}$$

We can find another equation, which is obtained extracting the first moment, as we did in eq. (2.228). For that, we are going to multiply by $\vec{Q}$ and $\vec{p}$ the eqs. (2.215) and (2.216) respectively, then we are going to integrate them in the respective momentum space. Since we multiply by the momentum instead of the momentum over the mass, we will obtain in the l.h.s. the eq. (2.228) multiplied by the respective mass, thus the integrated eqs. (2.215) and (2.216) are

$$m_p\frac{\mathrm{d}n_p}{\mathrm{d}t} = m_p\bar{n}_p\left(\dot{v}_p^j + Hv_p^j + \frac{1}{a}\frac{\partial\Psi}{\partial x^j}\right) = \langle C_{\mathrm{e}^-p^+}Q^j\rangle_{\vec{Q}}\,, \tag{2.235}$$

$$m_\mathrm{e}\frac{\mathrm{d}n_\mathrm{e}}{\mathrm{d}t} = m_\mathrm{e}\bar{n}_\mathrm{e}\left(\dot{v}_\mathrm{e}^j + Hv_\mathrm{e}^j + \frac{1}{a}\frac{\partial\Psi}{\partial x^j}\right) = \langle C_{\mathrm{e}^-p^+}q^j\rangle_{\vec{q}} + \langle C_{\mathrm{e}^-\gamma}q^j\rangle_{\vec{q}}\,. \tag{2.236}$$

We proceed to sum them considering that they have the same terms inside the parenthesis, because the velocity is equal for electrons and protons. We also consider that $m\bar{n} = \bar{\rho}$ and that the proton mass is predominant, then $\bar{\rho}_p + \bar{\rho}_\mathrm{e} \simeq \bar{\rho}_b$. Finally, we can group the collision terms for Coulomb scattering as $\langle C_{\mathrm{e}^-p^+}Q^j\rangle_{\vec{Q}} + \langle C_{\mathrm{e}^-p^+}q^j\rangle_{\vec{q}} = \langle C_{\mathrm{e}^-p^+}(q^j + Q^j)\rangle_{\vec{q}} = \langle C_{\mathrm{e}^-p^+}(q^j + Q^j)\rangle_{\vec{Q}}$ because they are terms integrated in all momenta. The sum of eqs. (2.235) and (2.236) is

$$\rho_b\left(\dot{v}_b^j + Hv_b^j + \frac{1}{a}\frac{\partial\Psi}{\partial x^j}\right) = \langle C_{\mathrm{e}^-p^+}(q^j + Q^j)\rangle_{\vec{q}} + \langle C_{\mathrm{e}^-\gamma}q^j\rangle_{\vec{q}}\,, \tag{2.237}$$

where the first term of the r.h.s. vanishes due to momentum conservation since we are doing an integral over all momenta. Then only the Compton scattering term survives and eq. (2.237) can be written as

$$\dot{v}_b^j + Hv_b^j + \frac{1}{a}\frac{\partial\Psi}{\partial x^j} = \frac{1}{\rho_b}\langle C_{\mathrm{e}^-\gamma}q^j\rangle_{\vec{q}}\,. \tag{2.238}$$

Using the same momentum conservation for the Compton scattering we have that $\langle C_{\mathrm{e}^-\gamma}\vec{p} + \vec{q}\rangle = 0$ and then $\langle C_{\mathrm{e}^-\gamma}\vec{q}\rangle = -\langle C_{\mathrm{e}^-\gamma}\vec{p}\rangle$, where we omitted the momentum subscript because it loses its purpose since these quantities are being integrated in all momenta. Therefore, we are going to write eq. (2.238) using the conformal time and the indication above

$$v'^j_b + \mathcal{H}v_b^j + \frac{\partial\Psi}{\partial x^j} = -\frac{a}{\rho_b}\langle C_{\mathrm{e}^-\gamma}p^j\rangle\,, \tag{2.239}$$

which in Fourier space is

$$\vec{v}_b' + \mathcal{H}\vec{v}_b + i\vec{k}\psi = -\frac{a}{\rho_b}\langle C_{\mathrm{e}^-\gamma}\vec{p}\rangle\,, \tag{2.240}$$

then we proceed to made the dot product by $\hat{k}$ in both sides of the equation. In section 2.9.1 we saw that $\vec{v}_b$ is in the same direction that $\hat{k}$, so we will recover its modulus, while in the r.h.s. we saw that $\vec{p}\cdot\hat{k} = p\mu$. Then eq. (2.240) becomes in

$$v_b' + \mathcal{H}v_n + ik\psi = -\frac{a}{\rho_b}\langle C_{\mathrm{e}^-\gamma}p\mu\rangle\,. \tag{2.241}$$

We have already computed $C_{\mathrm{e}^-\gamma}$ in eq. (2.196), thus the r.h.s. of eq. (2.241) is

$$\begin{aligned} -\frac{a}{\rho_b}\langle C_{\mathrm{e}^-\gamma} p\mu\rangle &= -\frac{a}{\rho_b}\left[-n_\mathrm{e}\sigma_T\int\frac{\mathrm{d}^3p}{(2\pi)^3}\frac{\partial\bar{f}}{\partial p}p^2\mu\left(\Theta_0-\Theta(\hat{p})+\hat{p}\cdot\vec{v}_b\right)\right] \\ &= \frac{an_\mathrm{e}\sigma_T}{\rho_b}\int_0^\infty\frac{\mathrm{d}p}{2\pi^2}p^4\frac{\partial\bar{f}}{\partial p}\int_{-1}^1\frac{\mathrm{d}\mu}{2}\mu\left(\Theta_0-\Theta(\hat{p})+\mu v_b\right) \\ &= \tau'_{op}\frac{4\bar{\rho}_\gamma}{\bar{\rho}_b}\left(i\Theta_1+3v_b\right)\,, \end{aligned} \tag{2.242}$$

where in the second line, we used spherical coordinates and separated integrals for p and μ. We have to remember that, besides spacetime coordinates, Θ_0 does not depend on the momentum, and Θ does not depend on the momentum modulus. Then the integral in p can be done by parts which gives $-\int_0^\infty\frac{\mathrm{d}p}{2\pi^2}4p^3\bar{f}=-\int\frac{\mathrm{d}^3p}{(2\pi)^3}4p\bar{f}$ and since for photons $\varepsilon=p$ we can recognize that its value is $-4\bar{\rho}_\gamma$, but its negative sign is used in the definition of τ_{op} in eq. (2.206). Θ_0 and v_b are independent for the integral in μ, so its first term vanishes since it is an even function in a symmetric interval, and the third term gives $v_b/3$. Finally, we have defined the first moment or dipole of the temperature perturbation as

$$\Theta_1\equiv i\int_{-1}^1\frac{\mathrm{d}\mu}{2}\mu\Theta(\hat{p})\,, \tag{2.243}$$

and according to eq. (2.207), where we only made explicit the dependency in the momentum direction since its definition can be used in both real and Fourier space. The eq. (2.241) then is

$$v_b'+\mathcal{H}v_n+ik\psi=\frac{\tau'_{op}}{R_s}\left(3i\Theta_1+v_b\right)\,, \tag{2.244}$$

in which we have defined the baryon-to-photon density ratio R_s as

$$R_s\equiv\frac{3\bar{\rho}_b}{4\bar{\rho}_\gamma}\,. \tag{2.245}$$

2.9.3 Cold Dark Matter

To consider the dark matter in the mix of baryons and photons, we just have to add the equations for its perturbed quantities. Since the dark matter in principle only interacts gravitationally, it has no collision term, at least one known, because its effects are encoded in the potentials Φ and Ψ. Hence we can use the same equations for the matter (baryons) without a collision term, which in Fourier space is equivalent to set eqs. (2.231) and (2.232) to zero. But when we obtained those equations we dropped terms of order $v^2\sim(p/\varepsilon)^2$ and higher, then to preserve the equations we have to assume the same for dark matter, which is the reason why we called it as cold dark matter. Once done, the equations for cold dark matter in Fourier space are

$$\delta'_\mathrm{cdm}+-ikv_\mathrm{cdm}+3\phi'=0\,, \tag{2.246}$$

$$v'_\mathrm{cdm}+\mathcal{H}v_\mathrm{cdm}+ik\psi=0\,, \tag{2.247}$$

which are the same that we obtained for a general matter fluid ($w=c_s^2=\sigma=0$) using the conservation of the energy-momentum tensor in eqs. (2.106) and (2.109).

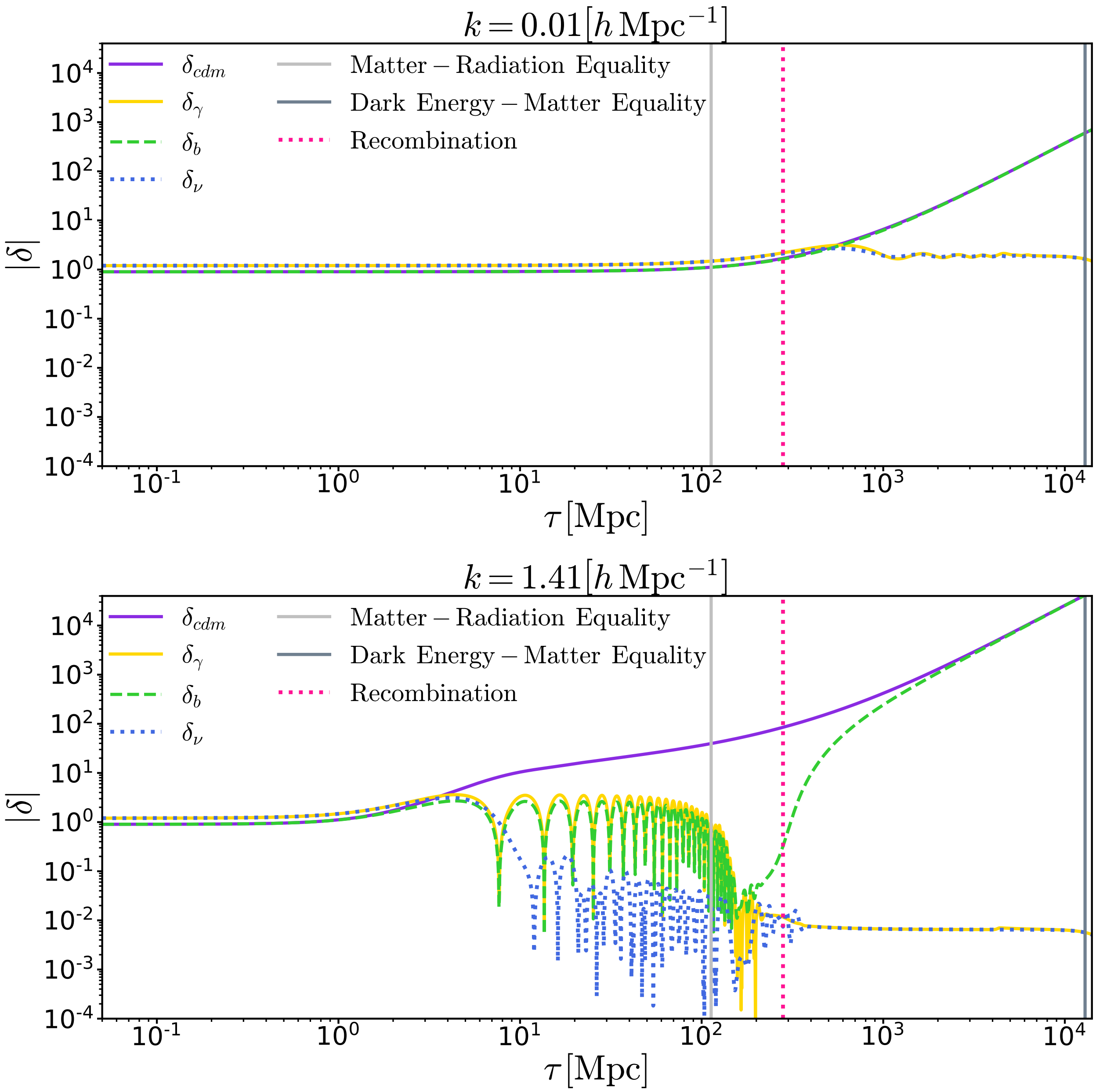


Figure 2.2: The absolute value of the density contrast (eq. (2.63)) as function of the conformal time for two characteristic scales -super-horizon limit at the *top* and sub-horizon limit at the *bottom*- and different components, computed using the software CLASS [46]. We consider the effects of the interacting fluids studied in sections 2.9.1 to 2.9.4, which take place near recombination. We also include the time of densities equality to determinate RDE and MDE, and the recombination time (see section 3.4.3).

2.9.4 Neutrinos

The neutrinos can be included in our treatment; in fact, as they interact only very weakly with the other species, we can add them in the same way we did for cold dark matter. Nevertheless, we are going to consider that they are massless or at least that they have not appreciable mass, then we will directly take their temperature perturbations $\mathcal{N}$ instead of the energy density perturbations δ_ν. Thus neutrinos behaves like radiation $\rho_\nu \propto a^{-4}$ ($w = 1/3$ in eq. (1.71)) and due to eq. (2.188) $\rho_\nu \propto T^4$, which after differentiating becomes in $\mathrm{d}\rho_\nu = 4\rho_\nu \Theta \Rightarrow \delta_\nu = 4\mathcal{N}_0$ where we integrated over the μ dependency which makes appear the monopole. The same happens with photons, and therefore the total radiation energy density perturbations can be written as

$$\begin{aligned}\bar{\rho}_r \delta_r &= 4\left(\bar{\rho}_\gamma \Theta_0 + \bar{\rho}_\nu \mathcal{N}_0\right) \\ &= 4\bar{\rho}_r\left(\left[1 - f_\nu\right]\Theta_0 + f_\nu \mathcal{N}_0\right) ,\end{aligned} \tag{2.248}$$

where we defined f_ν as

$$f_\nu \equiv \frac{\bar{\rho}_\nu + \bar{\rho}_\gamma}{\bar{\rho}_r} . \tag{2.249}$$

As neutrinos interact very weakly, we can use the same equation for photons (eq. (2.205) in Fourier space) but without scattering terms coming from Compton scattering, which gives

$$\mathcal{N}' + ik\mu\mathcal{N} + \phi' + ik\mu\psi = 0 . \tag{2.250}$$

Nonetheless, we also can probe that other multipoles source the perturbed potentials. If we apply the same operator of eq. (2.99) in Fourier space, but now on the general energy-momentum tensor of eq. (2.170), we can note that its spatial part obeys

$$\left(\hat{k}_i \hat{k}^j - \frac{1}{3}\delta_i^j\right) T^i_{\ j} = g \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{p^2\mu^2 - \frac{1}{3}p^2}{\varepsilon(p)} f\left(\vec{p}, \vec{x}, t\right) , \tag{2.251}$$

where we used eq. (2.204), and the combination $\mu^2 - 1/3$ suggests us that we should write it as $2\mathcal{P}_2(\mu)/3$. Then, expanding up to first-order the distribution in eq. (2.251) according to eq. (2.185) and after integrating by parts, we obtain

$$\left(\hat{k}_i + \hat{k}^j - \frac{1}{3}\delta_i^j\right) T^i_{\ j} = -\frac{8}{3}\bar{\rho}\Theta_2 , \tag{2.252}$$

in which we recognized the definition of quadrupole according to eq. (2.207). Thus, by eq. (2.100), we can rewrite eq. (2.105) using eq. (2.252) as

$$k^2\left(\phi + \psi\right) = -32\pi G a^2 \bar{\rho}\Theta_2 , \tag{2.253}$$

meaning that the quadrupole causes a difference in the potentials. Since for non-relativistic fluids, the quadrupole vanishes, they do not develop anisotropic stress, and then, the potentials are opposite $\phi = -\psi$. Like the energy-momentum is additive, we can rewrite eq. (2.253) considering all relativistic particles *i.e.* photons and neutrinos, as

$$k^2\left(\phi + \psi\right) = -32\pi G a^2\left(\bar{\rho}_\gamma \Theta_2 + \bar{\rho}_\nu \mathcal{N}_2\right) . \tag{2.254}$$

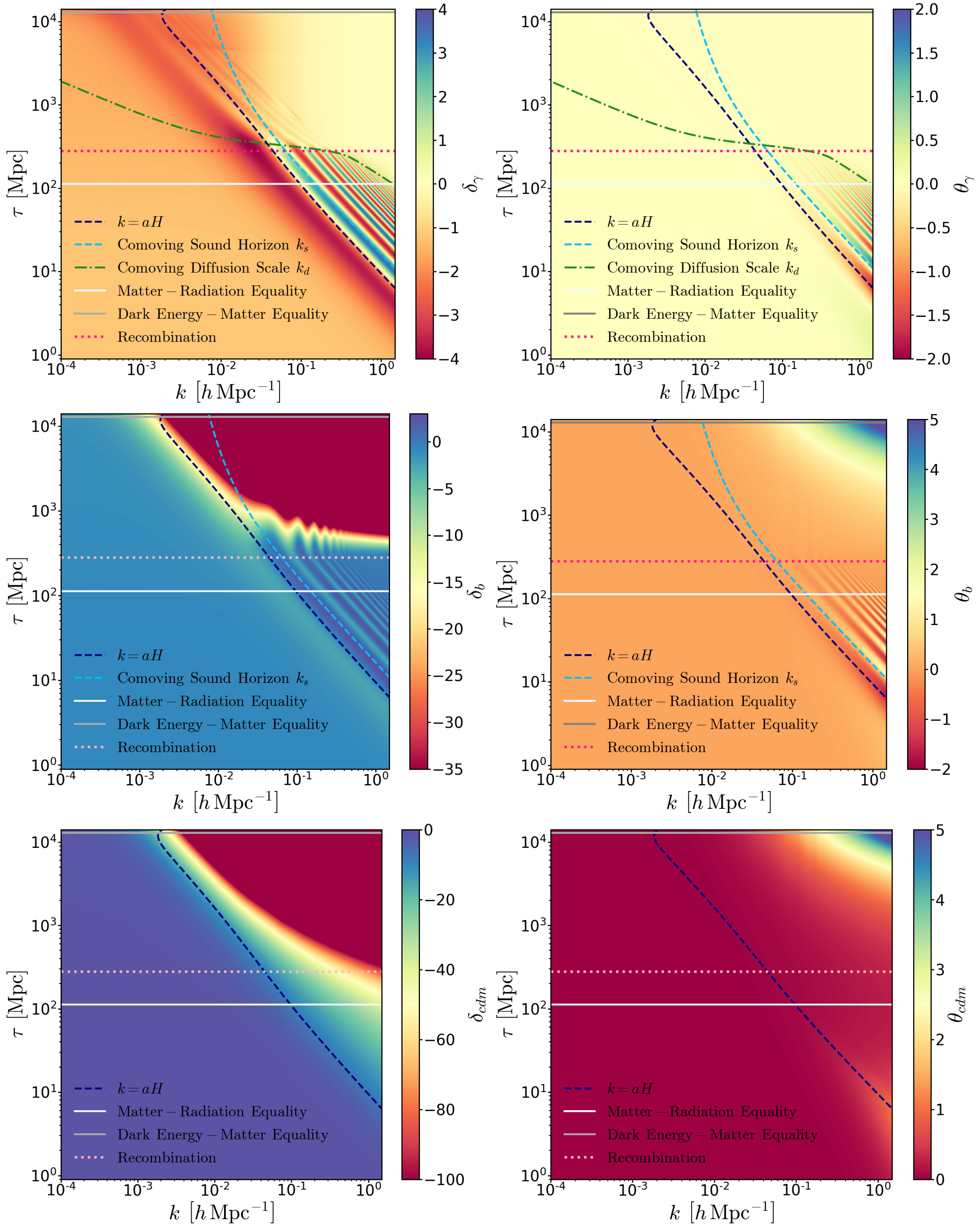


Figure 2.3: The fields $\delta(\tau, k)$ and $\theta(\tau, k)$ from eqs. (2.63) and (2.64), respectively, for photons (*upper panel*), baryons (*central panel*) and cold dark matter (*lower panel*) computed using the software CLASS [46]. We consider the interacting fluids effects from section 2.9. We also show the time of densities equality to determinate RDE and MDE, the recombination time (see section 3.4.3), the horizon $k = aH$ which set the limit between sub-horizon and super-horizon scales, the comoving sound horizon k_s (see eq. (3.127)), and the comoving diffusion scale k_d that affects photons.

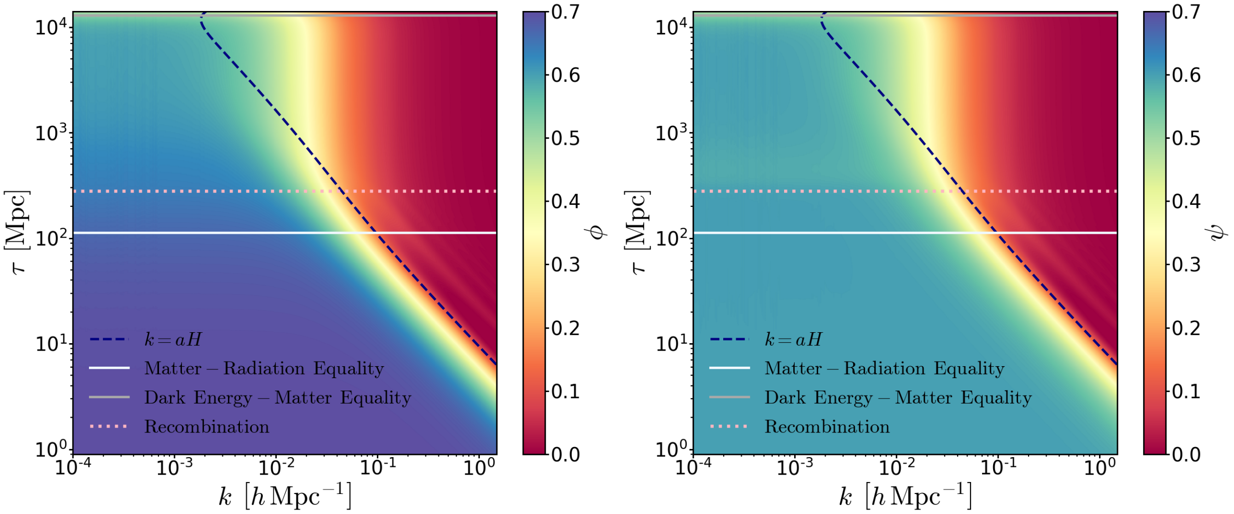


Figure 2.4: The potentials fields $\phi(\tau, k)$ (*left*) and $\psi(\tau, k)$ (*right*) computed using the software CLASS [46]. We consider the effects of the interacting fluids studied in sections 2.9.1 to 2.9.4. We also include the time of densities equality to determinate RDE and MDE, the recombination time (see section 3.4.3) and the horizon (blue dashed line).

Chapter 3

History Of The Universe

In the last chapters, we studied the equations that govern the background and perturbed quantities, at least for scalar variables. Nonetheless, to solve them, we need initial conditions and knowledge of the processes from the beginning until nowadays. The quest for initial conditions led to physicians of the 20th century to propose the theory of Inflation, which also explained some additional problems of the epoch. The expansion also forecasts that our primordial Universe was small, where processes of high energy physics arise, explaining why the early Universes was radiation, and latter matter, dominated. All these effects change how we observe our Universe now, and then, to prepare the field for some observations, we are going to make a quick summary of them, following [80, 31, 32, 212].

3.1 Inflation

One of the recurrent arguments that we used to describe our Universe is the cosmological principle, but we have not described a mechanism that explains it. The fact that nowadays we measure disjoint patches of sky with almost the same densities and temperatures is called the horizon problem. Also, our Universe looks flat nowadays according to precise measures [59], which does not have a straightforward explanation, forcing the called flatness problem. Those problems together with the apparent lack of magnetic monopoles necessary in some Grand Unified Theories [74] -the monopole problem- can be explained by a period early in the history of the Universe, when the expansion was extremely accelerating called Inflation.

To explain the horizon problem, we first have to define the particle horizon, which is the maximal comoving distance that light can travel between two times t_1 and t_2. Using eq. (1.52) with $c = 1$, it is just

$$\Delta\chi = \int_{t_1}^{t_2} \frac{\mathrm{d}t}{a(t)} = \tau_2 - \tau_1 \,, \tag{3.1}$$

where we used the conformal time in eq. (1.3). Then, eq. (3.1) can be rewrite using again eq. (1.52) as

$$\Delta\chi = \int_{a_1}^{a_2} \frac{\mathrm{d}a}{a} \frac{1}{aH(a)} = \int_{\ln a_1}^{\ln a_2} \mathrm{d}\ln a \frac{1}{aH(a)} \,. \tag{3.2}$$

That means that the particle horizon is the logarithmic integral of the comoving Hubble radius $1/aH(a)$. For an Universe dominated by a fluid with constant equation of state parameter w we can write the Hubble radius using eqs. (1.71) and (1.72)

$$\frac{1}{aH(a)} = H_0^{-1} a^{\frac{1}{2}(1+3w)}\,, \tag{3.3}$$

then the particle horizon $\chi_p(a)$ can be found inserting eq. (3.3) into eq. (3.2)

$$\begin{aligned} \chi_p(a) &= \frac{2}{H_0(1+3w)}\left[a^{\frac{1}{2}(1+3w)} - a_i^{\frac{1}{2}(1+3w)}\right] \\ &= \frac{2}{1+3w}\left[(aH(a))^{-1} - (a_iH(a_i))^{-1}\right]\,, \end{aligned} \tag{3.4}$$

where, after using the correspondence of eq. (3.1), it is possible to write

$$\tau(a) = \frac{2a^{\frac{1}{2}(1+3w)}}{H_0(1+3w)}\,. \tag{3.5}$$

If our fluid has $w > -1/3$ like all familiar sources, we can note that $\tau_i \equiv \tau(a \to 0) \to 0$ and $\chi_p \propto (aH)^{-1}$. These constraints mean that the comoving horizon of eq. (3.4) does not depend on early times, and it is finite. How is possible then, that photons coming from two opposite directions on the sky that were emitted at CMB epoch (that is sufficiently close to the Big Bang singularity [80]) have almost the same temperature if they were never in causal contact? A solution for this problem could be a phase of decreasing Hubble radius in the early Universe,

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(aH)^{-1} < 0\,. \tag{3.6}$$

Under this assumption, the Hubble radius may have been large enough to communicate different regions in the sky that we see similar nowadays. But eq. (3.3) says that a decreasing Hubble radius requires that $1+3w < 0$ and then $\tau_i \equiv \tau(a \to 0) \to -\infty$ by eq. (3.5). Therefore the Big Bang singularity is at a negative conformal time instead of zero like we thought, meaning that there was more conformal time between the singularity and CMB.

To explain the flatness problem, we must take eq. (1.31) and divide it by H^2,

$$1 - \Omega(t) = -\frac{K}{(aH)^2}\,, \tag{3.7}$$

where we also use the definition of eq. (1.35). Setting $t = t_{now}$ and therefore $a = a_0 = 1$, $H = H_0$ and $\Omega = \Omega_0$ we find $1 - \Omega_0 = -K/H_0^2$. Replacing the last expression in order to eliminate K from eq. (3.7), give us

$$1 - \Omega(t) = H_0^2(1-\Omega_0)\frac{1}{(aH)^2}\,. \tag{3.8}$$

Precise measures [59, 1] have constrained the value of $|1 - \Omega_0| = 0.001 \pm 0.002$, meaning that our Universe noways is mostly flat. However, as we saw previously, the Hubble radius is an increasing function of time and by eq. (3.8), $1 - \Omega$ too. Meaning that if our Universe is flat today, in the past, it was extremely flat, for example, $|1-\Omega(a_{GUT})| \leq 10^{-55}$ [31][1]. In classical Cosmology, we have not explained how this happens, but taking the solution of eq. (3.6), we can also resolve this problem because, during this period, the Hubble radius decreases and then our Universe is driven towards flatness. We can demonstrate it using the derivative of eq. (3.7) with respect to the logarithm scale factor

$$\begin{aligned} \frac{\mathrm{d}\Omega}{\mathrm{d}\ln a} &= \frac{\mathrm{d}\Omega}{\mathrm{d}a}\frac{\mathrm{d}a}{\mathrm{d}\ln a} \\ &= \frac{K(1+3w)}{a^3H^2}a \\ &= (\Omega-1)(1+3w) \ , \end{aligned} \tag{3.9}$$

where we used the Hubble radius definition of eq. (3.3). In eq. (3.9) we can see that $\Omega = 1$ is an unstable fixed point if we have a classical fluid with $w > -1/3$, but it becomes in an attractor for $w < -1/3$, which is the same condition required for the horizon problem.

Now we can explain why this period is called Inflation, for that, we have to take our condition in eq. (3.6) and develop it

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(aH)^{-1} &= \frac{\mathrm{d}}{\mathrm{d}t}(\dot{a})^{-1} \\ &= -\frac{\ddot{a}}{(\dot{a})^2} \ , \end{aligned} \tag{3.10}$$

then to satisfy eq. (3.6), $\ddot{a}$ has to be positive. This constraint is why we call this phase as Inflation and also as a period of accelerated expansion. We can also develop the derivative in another way

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(aH)^{-1} &= -\frac{\dot{a}H + a\dot{H}}{(aH)^2} \\ &= -\frac{1}{a}(1-\epsilon) \ , \end{aligned} \tag{3.11}$$

where we defined the first slow-roll parameter ϵ as

$$\epsilon \equiv -\frac{\dot{H}}{H^2} \ , \tag{3.12}$$

and then $\epsilon < 1$ in order to satisfy eq. (3.6). We also can write ϵ as

$$\epsilon = \frac{\mathrm{d}\ln H}{\mathrm{d}N} \ , \tag{3.13}$$

with $N = \ln a$ as the number of e-folds of inflationary expansion, also defined in eq. (2.124). Then, the condition in eq. (3.6) implies that the fractional change of the Hubble parameter per e-folds is small. But to solve the cosmological problems, Inflation must hold between 40 to 60 e-folds [31], for that we introduce the second slow-roll parameter

$$\eta \equiv \frac{\dot{\epsilon}}{H\epsilon} = \frac{\mathrm{d}\ln\epsilon}{\mathrm{d}N} \ , \tag{3.14}$$

---

[1] Grand Unified Theories.

then to achieve that ϵ remains small for a large number of e-folds we have to demand that $|\eta| < 1$.

For a perfect Inflation $\epsilon = 0$ and then by the definition in eq. (3.12), $H =$ constant. Thus, the metric of eq. (1.9) becomes in de Sitter spacetime

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{e}^{2Ht}\mathrm{d}x^i\mathrm{d}x^j\,, \tag{3.15}$$

but we know that Inflation has to end, and then the spacetime can not be a perfect de Sitter metric. However, for a small ϵ, the line element in eq. (3.15) is still a good approximation, and that is why we can also refer to Inflation as the quasi-de Sitter period.

To explain the accelerated expansion at Inflation, we have to rewrite eq. (1.33) using eq. (1.31) with $K = 0$ (because the flatness in this period) as

$$\begin{aligned} \dot{H} + H^2 &= -\frac{4\pi G}{3}\left(\rho + 3P\right) \\ &= -\frac{H^2}{2}\left(1 + \frac{3P}{\rho}\right)\,, \end{aligned} \tag{3.16}$$

then using eq. (3.16) in the definition of eq. (3.12), we can obtain ϵ

$$\begin{aligned} \epsilon &= -\frac{\dot{H}}{H^2} \\ &= \frac{3}{2}\left(1 + \frac{P}{\rho}\right) \qquad \leq \qquad 1\,, \\ \Rightarrow &\qquad\qquad\quad \frac{P}{\rho} \leq \; -\frac{1}{3}\,, \end{aligned} \tag{3.17}$$

where we also used the condition on ϵ. Meaning that Inflation requires a negative pressure or an EoS parameter $w < -1/3$, as we mentioned before. Now if we consider eq. (1.39) with the value of ϵ given in eq. (3.17), we obtain

$$\begin{aligned} \frac{\mathrm{d}\rho}{\rho} &= -\,3H\left(1 + \frac{P}{\rho}\right)\mathrm{d}t \\ \mathrm{d}\ln\rho &= -\,3H\left(1 + \frac{P}{\rho}\right)\frac{\mathrm{d}\ln a}{H}\,, \\ \Rightarrow \frac{\mathrm{d}\ln\rho}{\mathrm{d}\ln a} &= -\,2\epsilon\,, \\ \Rightarrow \left|\frac{\mathrm{d}\ln\rho}{\mathrm{d}\ln a}\right| &= 2\epsilon\,. \end{aligned} \tag{3.18}$$

The eq. (3.18) tells us that for a small ϵ, as the condition in eq. (3.6) claims it, also the change of the logarithm density over e-folds is small. Then, the energy density is nearly constant, but conventional matter sources dilute with expansion. Therefore this fluid has to be something more unusual. Many theories have emerged [32] to try to explain this kind of fluid, but we are going to explain the most common and simple in which a scalar field called the *inflaton* drives Inflation. To begin, we are going to write the energy-momentum tensor for a generically scalar field $\phi(\vec{x}, t)$

$$T^{\mu}{}_{\nu} = g^{\mu\alpha}\frac{\partial\phi}{\partial x^{\alpha}}\frac{\partial\phi}{\partial x^{\nu}} - \delta^{\mu}_{\nu}\left[\frac{1}{2}g^{\alpha\beta}\frac{\partial\phi}{\partial x^{\alpha}}\frac{\partial\phi}{\partial x^{\beta}} + V(\phi)\right]\,, \tag{3.19}$$

where $V(\phi)$ is the potential for the field. Take care of not confuse the scalar field with the gravitational potential in Fourier space (eq. (2.93)). To preserve the symmetries of the FLRW metric, we require that the value of the inflaton only depends on time, at least to zeroth-order

$$\phi(\vec{x}, t) = \bar{\phi}(t) + \delta\phi(\vec{x}, t)\,. \tag{3.20}$$

For the background part of the field $\bar{\phi}(t)$, only time derivatives are relevant in eq. (3.19). We can find expressions for the energy density and pressure for the homogeneous part of the field by taking eq. (1.11) and noting that $T^0_{\ 0} = -\rho_\phi$ and $T^i_{\ j} = P_\phi \delta^i_j$, then

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi)\,, \tag{3.21}$$

$$P_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi)\,, \tag{3.22}$$

where we omitted the time dependence for simplicity, and also we wrote the background part of the field just as ϕ. The expressions in eqs. (3.21) and (3.22) tell us that the condition in eq. (3.17) is satisfied if the field has more potential energy than kinetic, this situation is called slow-roll inflation and it is satisfied by the slow-roll parameters ϵ, η (eqs. (3.12) and (3.14)). It is also possible to find an equation of movement for ϕ replacing those expressions in eqs. (1.31) and (1.32) for $K = 0$

$$H^2 = \frac{8\pi G}{3}\left(\frac{1}{2}\dot{\phi}^2 + V\right)\,, \tag{3.23}$$

$$\begin{aligned} \dot{H} &= -8\pi G P - 3H^2 \\ &= -4\pi G \dot{\phi}^2\,, \end{aligned} \tag{3.24}$$

where we also replaced the eq. (3.23) into eq. (3.24). Using the eq. (3.24) into the definition of eq. (3.12) we find that

$$\epsilon = 4\pi G \frac{\dot{\phi}^2}{H^2}\,. \tag{3.25}$$

Now we take the time derivative of eq. (3.23) and insert the result of eq. (3.24) that leads us the Klein-Gordon evolution equation for the scalar field

$$\ddot{\phi} + 3H\dot{\phi} + \frac{\partial V}{\partial \phi} = 0\,, \tag{3.26}$$

which tells us that in the inflationary period, the inflaton potential acts as a force, while the expansion adds friction.

Now we are going to describe the perturbed part of the inflaton field, for that, it is convenient to write eq. (3.20) using the conformal time as

$$\phi(\vec{x}, \tau) = \bar{\phi}(\tau) + \frac{f(\vec{x}, \tau)}{a(\tau)}\,, \tag{3.27}$$

noting that we are now going to work over the function f instead $\delta\phi$ to simplify future calculations. We begin writing the action for a generically scalar field

$$S = \int \mathrm{d}\tau \mathrm{d}^3x \sqrt{-g}\left[\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right]\,, \tag{3.28}$$

where $g \equiv \det(g_{\alpha\beta})$ and $g_{\alpha\beta}$ is the unperturbed FLRW metric of eq. (1.9) for flat cosmologies in cartesian coordinates. Using eq. (3.27) into eq. (3.28) the action takes the form

$$S = \int \mathrm{d}\tau \mathrm{d}^3x \left[-\frac{a^2}{2} \left\{ \bar{\phi}' + \frac{f'}{a} - \mathcal{H}\frac{f}{a} \right\}^2 + \frac{1}{2} \left(\nabla f\right)^2 - a^4 V\left(\bar{\phi} + \frac{f}{a}\right) \right] . \tag{3.29}$$

To work up to first-order, it is convenient to expand the potential until the second-order in perturbations as

$$V\left(\bar{\phi} + \frac{f}{a}\right) \simeq V\left(\bar{\phi}\right) + \frac{\partial V}{\partial \phi}\frac{f}{a} + \frac{1}{2}\frac{\partial^2 V}{\partial \phi^2}\left(\frac{f}{a}\right)^2 , \tag{3.30}$$

which we are going to replace into eq. (3.29). Using that expansion and developing the parenthesis in the kinetic part of eq. (3.29) allow us to write the action as

$$\begin{aligned} S \simeq \int \mathrm{d}\tau \mathrm{d}^3x \Bigg[& -\frac{a^2}{2}\left(\bar{\phi}'\right)^2 - \frac{1}{2}f'^2 - \frac{\mathcal{H}^2}{2}f^2 - a\bar{\phi}'f' + a\mathcal{H}\bar{\phi}'f + \mathcal{H}f'f \\ & + \frac{1}{2}\left(\nabla f\right)^2 - a^4 V\left(\bar{\phi}\right) - a^3 f\frac{\partial V}{\partial \phi} - \frac{a^2}{2}f^2\frac{\partial^2 V}{\partial \phi^2} \Bigg] . \end{aligned} \tag{3.31}$$

The action in eq. (3.31) can be separated into powers of the perturbed field f or its derivatives

$$S^{(0)} = \int \mathrm{d}\tau \mathrm{d}^3x \left[-\frac{a^2}{2}\left(\bar{\phi}'\right)^2 - a^4 V\left(\bar{\phi}\right) \right] , \tag{3.32}$$

$$S^{(1)} = \int \mathrm{d}\tau \mathrm{d}^3x \left[-a\bar{\phi}'f' + a\mathcal{H}\bar{\phi}'f - a^3 f\frac{\partial V}{\partial \phi} \right] , \tag{3.33}$$

$$S^{(2)} = \int \mathrm{d}\tau \mathrm{d}^3x \left[-\frac{1}{2}f'^2 - \frac{\mathcal{H}^2}{2}f^2 + \mathcal{H}f'f + \frac{1}{2}\left(\nabla f\right)^2 - \frac{a^2}{2}f^2\frac{\partial^2 V}{\partial \phi^2} \right] . \tag{3.34}$$

The eq. (3.32) does not give us information since their terms disappear after to demand the least action principle $\delta S = 0$. While the first term of eq. (3.33) can be integrated by parts, dropping the boundary term, which allows us to write eq. (3.33) as

$$S^{(1)} = \int \mathrm{d}\tau \mathrm{d}^3x \left[\bar{\phi}'' + 2\mathcal{H}\bar{\phi}' + a^2\frac{\partial V}{\partial \phi} \right] af , \tag{3.35}$$

which after imposing $S^{(1)} = 0$ for all f gives the same Klein-Gordon equation of eq. (3.26) but now as a function of conformal time

$$\phi'' + 2\mathcal{H}\phi' + a^2\frac{\partial V}{\partial \phi} = 0 . \tag{3.36}$$

Finally writing the term $\mathcal{H}f'f = \frac{1}{2}\mathcal{H}(f^2)'$ in eq. (3.34) and integrating it by parts, dropping again the boundary term, the quadratic action of eq. (3.34) can be written as

$$S^{(2)} = -\frac{1}{2}\int \mathrm{d}\tau \mathrm{d}^3x \left[f'^2 - \left(\nabla f\right)^2 + f^2 \left\{ \mathcal{H}' + \mathcal{H}^2 + a^2\frac{\partial^2 V}{\partial \phi^2} \right\} \right] , \tag{3.37}$$

then and after of considering $S^{(2)} = \int \mathrm{d}\tau \mathrm{d}^3x\mathcal{L}$, we can apply the Euler-Lagrange equation to the quadratic action

$$\frac{\partial \mathcal{L}}{\partial f} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu f \right)} \right) = 0 \,, \tag{3.38}$$

which gives us the motion equation for the field f

$$f'' - \nabla^2 f - f \left[\mathcal{H}' + \mathcal{H}^2 + a^2 \frac{\partial^2 V}{\partial \phi^2} \right] - \frac{1}{2} a f^2 \frac{\partial^3 V}{\partial \phi^3} = 0 \,, \tag{3.39}$$

where we are going to drop the last term in eq. (3.39) since it is of second-order in f. We will do the same with the second derivative of the potential since it is a first-order perturbation variable multiplied by f, which gives a second-order quantity. Finally and writing $\mathcal{H}' + \mathcal{H}^2 = a''/a$, gives the Mukhanov-Sasaki equation

$$f'' - \nabla^2 f - \frac{a''}{a} f = 0 \,, \tag{3.40}$$

which in Fourier space is[2]

$$f_k'' + \left(k^2 - \frac{a''}{a} \right) f_k = 0 \,. \tag{3.41}$$

Since in the Sitter spacetime $a''/a = 2(aH)^2 = 2/\tau^2$, the Mukhanov-Sasaki equation also takes the form

$$f_k'' + \left(k^2 - \frac{2}{\tau^2} \right) f_k = 0 \,, \tag{3.42}$$

which exact solution is given by

$$f_k(\tau) = C_1 \frac{\mathrm{e}^{-ik\tau}}{\sqrt{2k}} \left(1 - \frac{i}{k\tau} \right) + C_2 \frac{\mathrm{e}^{ik\tau}}{\sqrt{2k}} \left(1 + \frac{i}{k\tau} \right) \,. \tag{3.43}$$

To complete our results, we are going to make a quantum treatment of the inflaton fluctuations $f = a\delta\phi$. For that we define its momentum conjugate $\pi(\vec{x}, \tau)$ using the quadratic action of eq. (3.37)

$$\pi(\vec{x}, \tau) \equiv \frac{\partial \mathcal{L}}{\partial f'} = f' \,, \tag{3.44}$$

then we proceed to promote the fields $f(\vec{x}, \tau)$ and $\pi(\vec{x}, \tau)$ to quantum operators $\hat{f}$ and $\hat{\pi}$, which satisfy the standard equal time commutation relation

$$\left[\hat{f}(\vec{x}, \tau), \hat{\pi}(\vec{y}, \tau) \right] = i\delta \left(\vec{x} - \vec{y} \right) \,, \tag{3.45}$$

which in Fourier space is

$$\left[\hat{f}_{\vec{k}}(\tau), \hat{\pi}_{\vec{k}'}(\tau) \right] = i\delta \big(\vec{k} + \vec{k}' \big) \,. \tag{3.46}$$

Then we can expand the modes $\hat{f}_{\vec{k}}$ through

$$\hat{f}_{\vec{k}}(\tau) = f_k(\tau) \hat{a}_{\vec{k}} + f_k^*(\tau) \hat{a}_{\vec{k}}^\dagger \,, \tag{3.47}$$

---

[2] Taking the Fourier transform of f with the convention of eq. (2.93).

where f_k and its complex conjugate f_k^* are the solutions given by eq. (3.43). While $\hat{a}_{\vec{k}}$ and its Hermitian conjugate $\hat{a}^\dagger_{\vec{k}}$ are time-independent operators that satisfy

$$\left[\hat{a}_{\vec{k}}, \hat{a}^\dagger_{\vec{k}'}\right] = \delta\big(\vec{k} + \vec{k}'\big)\,, \tag{3.48}$$

which is valid as long as the solutions in eq. (3.43) are normalized as they are. These operators can be interpreted as the creation and annihilation operators, as in the quantum harmonic oscillator. Hence

$$\hat{a}_{\vec{k}}|0\rangle = 0\,, \tag{3.49}$$

$$|m_{\vec{k}_1}\rangle = \frac{1}{\sqrt{m!}}\big(\hat{a}^\dagger_{\vec{k}_1}\big)^m|0\rangle\,, \tag{3.50}$$

where $|0\rangle$ is the vacuum state. Nevertheless, for general time-dependent background, the choice of this state can be ambiguous, and there is a preferred option for Inflation. In early times ($\tau \to -\infty$) all the modes of cosmological interest were deep inside the horizon, which changes eq. (3.42) into

$$f_k'' + k^2 f_k = 0\,, \tag{3.51}$$

which means that modes had time-independent frequencies in the remote past. Thus the eq. (3.51) is the free field equation in Minkowski space, and their solutions are $f_k \propto \mathrm{e}^{\pm ik\tau}$. The factor $1/\sqrt{m!}$ in eq. (3.48) is because we set the Wronskian of the mode functions as $W[f_k, f_k^*] \equiv 1$, that also set the normalization in the solutions of eq. (3.43) and allows that only the positive frequency modes $f_k \propto \mathrm{e}^{-ik\tau}$ can be the minimal excitation state. For that reason, it is going to be chosen to define the vacuum state in Inflation, and therefore we are going to resolve the Mukhanov-Sasaki equation in eq. (3.42) with initial conditions given by

$$\lim_{\tau\to-\infty} f_k(\tau) = \frac{1}{\sqrt{2k}}\mathrm{e}^{-ik\tau}\,. \tag{3.52}$$

This state defines an unique physical vacuum called the Bunch-Davies vacuum, which sets the constants $C_1 = 1$ and $C_2 = 0$ in the solution of eq. (3.43).

Finally we aim to compute the quantum statistics of the operator $\hat{f}$ using eqs. (3.43) and (3.47) with $C_1 = 1$ and $C_2 = 0$ because the above discussion. Thus the operator $\hat{f}$ is

$$\begin{aligned}\hat{f}(\vec{x},\tau) &= \int \frac{\mathrm{d}^3k}{(2\pi)^3}\left[f_k(\tau)\hat{a}_{\vec{k}} + f_k^*(\tau)\hat{a}^\dagger_{\vec{k}}\right]\mathrm{e}^{i\vec{k}\cdot\vec{x}} \\ &= \int \frac{\mathrm{d}^3k}{(2\pi)^3}\left[\frac{\mathrm{e}^{-ik\tau}}{\sqrt{2k}}\left(1 - \frac{i}{k\tau}\right)\hat{a}_{\vec{k}} + \frac{\mathrm{e}^{ik\tau}}{\sqrt{2k}}\left(1 + \frac{i}{k\tau}\right)\hat{a}^\dagger_{\vec{k}}\right]\mathrm{e}^{i\vec{k}\cdot\vec{x}}\,,\end{aligned} \tag{3.53}$$

which we are going to use to compute its expectation value $\langle\hat{f}\rangle$

$$\begin{aligned}\langle\hat{f}\rangle \equiv& \langle 0|\hat{f}|0\rangle \\ =& \int \frac{\mathrm{d}^3k}{(2\pi)^3}\left[\frac{\mathrm{e}^{-ik\tau}}{\sqrt{2k}}\left(1 - \frac{i}{k\tau}\right)\langle 0|\hat{a}_{\vec{k}}|0\rangle + \frac{\mathrm{e}^{ik\tau}}{\sqrt{2k}}\left(1 + \frac{i}{k\tau}\right)\langle 0|\hat{a}^\dagger_{\vec{k}}|0\rangle\right]\mathrm{e}^{i\vec{k}\cdot\vec{x}} = 0\,.\end{aligned} \tag{3.54}$$

where we use $\hat{a}_{\vec{k}}|0\rangle = \langle 0|\hat{a}^{\dagger}_{\vec{k}} = 0$ from eq. (3.49). As we expected the expectation value of the perturbation vanishes, however its variance receive non-zero quantum fluctuations

$$\begin{aligned}
\langle|\hat{f}|^2\rangle \equiv& \langle 0|\hat{f}^{\dagger}(\vec{x},\tau)\hat{f}(\vec{x},\tau)|0\rangle \\
=& \int \frac{\mathrm{d}^3k}{(2\pi)^3}\int \frac{\mathrm{d}^3k'}{(2\pi)^3}\left\langle 0\left|\left(f^*_k(\tau)\hat{a}^{\dagger}_{\vec{k}} + f_k(\tau)\hat{a}_{\vec{k}}\right)\left(f_{k'}(\tau)\hat{a}_{\vec{k}'} + f^*_{k'}(\tau)\hat{a}^{\dagger}_{\vec{k}'}\right)\right|0\right\rangle \mathrm{e}^{i\vec{x}\cdot(\vec{k}'-\vec{k})} \\
=& \int \frac{\mathrm{d}^3k}{(2\pi)^3}\int \frac{\mathrm{d}^3k'}{(2\pi)^3} f_k(\tau)f^*_{k'}(\tau)\delta\big(\vec{k}+\vec{k}'\big)\mathrm{e}^{i\vec{x}\cdot(\vec{k}'-\vec{k})} \\
=& \int \frac{\mathrm{d}^3k}{(2\pi)^3}|f_k(\tau)|^2 \\
=& \int \frac{\mathrm{d}k}{k}\frac{k^3}{2\pi^2}|f_k(\tau)|^2\,,
\end{aligned} \tag{3.55}$$

noting that in the last line we write the integral using spherical coordinates to define the dimensionless power spectrum of fluctuations as

$$\begin{aligned}
\Delta^2_f(k,\tau) \equiv& \frac{k^3}{2\pi^2}|f_k(\tau)|^2 \\
=& \left(\frac{k}{2\pi}\right)^2\left(1+\frac{1}{(k\tau)^2}\right)\,,
\end{aligned} \tag{3.56}$$

where we used the solutions of eq. (3.43) with the known initial conditions.

In super-horizon scales, the second term in eq. (3.56) predominates over the first, and the k-dependence is lost for now. Using the de Sitter spacetime approximation $\tau^{-2} = (aH)^2$ the power spectrum of inflaton perturbations at those scales is given by

$$\begin{aligned}
\Delta^2_{\delta\phi}(k,\tau) =& a^{-2}\Delta^2_f(k,\tau) \\
=& \left(\frac{H}{2\pi}\right)^2\,.
\end{aligned} \tag{3.57}$$

It seems that both mode and time dependencies were lost in eq. (3.57), but we have to remember that H was assumed constant when we derived this result. More generally, we can compute the power spectrum at a specific time by evaluating H when the mode of interest leaves the horizon, and it implicitly extends the result for the pure de Sitter background to a slowly time-evolving quasi-de Sitter space. With this, the time is set, and we recover the k dependence at the horizon crossing $k = aH$ in the power spectrum

$$\Delta^2_{\delta\phi}(k) = \left(\frac{H}{2\pi}\right)^2\Bigg|_{k=aH}\,, \tag{3.58}$$

which also has the added benefit that the error caused by ignoring the metric fluctuations does not accumulate over time.

If we had considered metric perturbations using the spatially flat gauge[3] of eq. (2.23), the

---

[3] The same results can be obtained using the comoving gauge in which $\delta\phi = 0$, for more details, consult [32]

eq. (3.41) would have changed to

$$f_k'' + \left(k^2 - \frac{z''}{z}\right) f_k = 0\,, \tag{3.59}$$

$$z^2 \equiv 2a^2\varepsilon\,, \tag{3.60}$$

where in the de Sitter space, we can still make the approximation $z''/z \simeq 2/\tau^2$ for super-horizon perturbations and then obtain the same power spectrum of eq. (3.58). Also, there is another approximation to first-order using the slow-roll parameters

$$\frac{z''}{z} \simeq \frac{1}{\tau^2}\left\{2 + 3\left(\varepsilon + \frac{\eta}{2}\right)\right\} = \frac{1}{\tau^2}\left(\nu^2 - \frac{1}{4}\right)\,, \tag{3.61}$$

$$\nu \equiv \frac{3}{2} + \varepsilon + \frac{\eta}{2}\,, \tag{3.62}$$

which for a constant ν, it has an exact solution given by Hankel functions of the first and second kind, but after imposing the Bunch-Davies vacuum condition (eq. (3.52)), only the first kind remains.

In chapter 2, we computed the comoving curvature perturbation, which is a gauge-invariant quantity given in eq. (2.44). Like it does not change from gauge to gauge, we can track how the power spectrum of eq. (3.58) behaves after Inflation. We will compare it in two different gauges, where the first is the spatially flat gauge where the power spectrum was computed. Thus $\mathcal{R}$ in that gauge is

$$\mathcal{R} = \mathcal{H}\left(v - B\right)\,, \tag{3.63}$$

because $\Phi = E = 0$ for that choice (eq. (2.23)). Comparing it with the general form of the component $T^0_{\ i}$ in eq. (2.31), we note that in both appears the factor $B - v$ since for scalar quantities we have $T^0_{\ i} = \left(\bar\rho + \bar P\right)\left(v_i - B_i\right) = \left(\bar\rho + \bar P\right)\partial_i\left(v - B\right)$. Then, we are going to compute that component for the inflaton field using the eq. (3.19) and obtaining

$$T^0_{\ i} = -\frac{\bar\phi'}{a^2}\partial_i\delta\phi\,, \tag{3.64}$$

where we used the trick $\left(\mathbb{1} + \varepsilon\right)^{-1} \simeq \mathbb{1} - \varepsilon$ for the inverse of the metric in that gauge. Using again eq. (2.31) we make the direct relation $-\bar\phi'\delta\phi/a^2 = \left(\bar\rho + \bar P\right)\left(v - B\right)$, we also know that $\bar\phi' = a\dot{\bar\phi}$ and $\bar\rho + \bar P = \dot{\bar\phi}^2$ for the inflaton field according to eqs. (3.21) and (3.22). Therefore $v - B = -\delta\phi/\bar\phi' = -\delta\phi/(a\dot{\bar\phi})$, and the comoving curvature perturbation takes the form

$$\mathcal{R} = -\frac{\mathcal{H}}{\bar\phi'}\delta\phi = -\frac{H}{\dot{\bar\phi}}\delta\phi\,. \tag{3.65}$$

With the eq. (3.65) we can relate the inflaton perturbations directly with the comoving curvature perturbation variable, and in Fourier space, we can also connect their power spectrum

$$\Delta^2_{\mathcal{R}}(k) = \left(\frac{\mathcal{H}}{\bar\phi'}\right)^2 \Delta^2_{\delta\phi}(k) = \left(\frac{H}{\dot{\bar\phi}}\right)^2 \Delta^2_{\delta\phi}(k)\,, \tag{3.66}$$

then the power spectrum for $\mathcal{R}$ after using eqs. (3.25) and (3.58) is

$$\Delta^2_{\mathcal{R}}(k) = \frac{4\pi G}{\epsilon}\left(\frac{H}{2\pi}\right)^2\Bigg|_{k=aH} . \tag{3.67}$$

The other gauge in which we will compute $\mathcal{R}$ is the Newtonian gauge (eq. (2.22)) since it was where we mostly worked in chapter 2. The comoving curvature perturbation of eq. (2.44) in that gauge takes the form

$$\mathcal{R} = \Phi + \mathcal{H}v\,, \tag{3.68}$$

which after using eq. (2.73) becomes in

$$\mathcal{R} = \Phi + \frac{\mathcal{H}\left(\Phi' - \mathcal{H}\Psi\right)}{4\pi G a^2\left(\bar{\rho} + \bar{P}\right)}\,, \tag{3.69}$$

that only depends of the perturbed potentials. Taking the derivative of eq. (3.69) and using the eqs. (1.66), (2.43), (2.67), (2.70), (2.71), (2.77), (2.78) and (2.83), we find

$$4\pi G a^2\left(\bar{\rho} + \bar{P}\right)\mathcal{R}' = -4\pi G a^2\mathcal{H}\delta P_{\text{nad}} + \mathcal{H}c^2_{s(a)}\nabla^2\Phi + \frac{\mathcal{H}}{3}\nabla^2\left(\Phi + \Psi\right)\,, \tag{3.70}$$

in which we have $\delta P_{nad} = 0$ for barotropic fluids and $\Phi + \Psi = 0$ from eq. (2.80) for a fluid with null anisotropic stress. Then, the r.h.s. of eq. (3.70) is proportional to $\mathcal{H}\nabla^2\Phi$ and in Fourier space it is to $\mathcal{H}k^2\phi$. Meanwhile the l.h.s. of eq. (3.70) is proportional to $\mathcal{H}^2\mathcal{R}' = \mathcal{H}^3\mathcal{R}\left(\mathrm{d}\ln\mathcal{R}/\mathrm{d}\ln a\right)$. Also for super-horizon scales we saw that $\phi' = 0$ from eq. (2.121), thus $\mathcal{R}$ is proportional to Φ by eq. (3.69). Finally, for that limit we can conclude that the logarithm derivative of $\mathcal{R}$ scales as

$$\frac{\mathrm{d}\ln\mathcal{R}}{\mathrm{d}\ln a} \sim \left(\frac{k}{\mathcal{H}}\right)^2 \to 0\,, \tag{3.71}$$

which vanishes because we are seeking super-horizon scales, also it tells us that the comoving curvature perturbation remains unaltered until a later time in this limit. Hence $\mathcal{R}$ is a gauge-invariant and also constant in the super-horizon, turning it essential to track quantities from Inflation to nowadays.

The spectrum in eqs. (3.58) and (3.67) can be only a purely function of k since it is evaluated at $k = aH$, but as we saw it is a scale-free spectrum. Which it is bad because a scale-free or Harrison-Zel'dovich-Peebles spectrum is what one might have expected even without the Inflation theory. However, since H is a slow-varying function of time, we expect that the power spectrum will deviate slightly from the scale-invariant form. In general, to quantify those deviations for some power spectrum Δ^2 we define the spectral index

$$n - 1 \equiv \frac{\mathrm{d}\ln\Delta^2}{\mathrm{d}\ln k}\,, \tag{3.72}$$

then the power spectrum can be written as

$$\Delta^2 = A\left(\frac{k}{k_*}\right)^{n-1}\,, \tag{3.73}$$

where A is the amplitude set at some k_*. Once obtained, the power spectrum of eq. (3.73) can be related with the one in eq. (3.67). Besides, for the power spectrum in eqs. (3.58) and (3.67), the spectral index is called the scalar spectral index n_s, which to first order in the slow-roll parameters takes the value

$$n_s - 1 = -2\varepsilon - \eta\,. \tag{3.74}$$

Thus it shows that for de Sitter spacetime, the power spectrum is perfectly scale-invariant. On the other hand, n_s measures the deviations from the perfect de Sitter.

Finally, we can also compute the original power spectrum instead of the dimensionless one. For that, we go back to eq. (3.56), defining

$$P(k) \equiv \frac{2\pi^2}{k^3}\Delta^2(k)\,, \tag{3.75}$$

which for the comoving curvature perturbation reads

$$P_{\mathcal{R}}(k) = \frac{2\pi^2}{k^3} A \left(\frac{k}{k_*}\right)^{n_s-1}\,, \tag{3.76}$$

where we used eq. (3.73). We also are going to set k_* at horizon crossing nowadays *i.e.* $k_* = a_0H_0 = H_0$ for eq. (1.10), this also sets the amplitude as we argued before, where we will named it as A_s. Therefore, the power spectrum of eq. (3.76) now is

$$P_{\mathcal{R}}(k) = \frac{2\pi^2}{k^3} A_s \left(\frac{k}{H_0}\right)^{n_s-1}\,. \tag{3.77}$$

3.2 Reheating

As we mentioned before, the inflationary epoch has to end to produce the Hot Big Bang scenario and then the Universe that we observe. Nevertheless, at the end of Inflation, the Universe is typically in a highly non-thermal state because the ability of Inflation to homogenize it, then our Universe is left at effectively zero temperature. That means that it has to be reheated or defrosted to the high temperatures that we require for the standard Hot Big Bang. This process is known as Reheating, and any successful theory of Inflation must also explain it.

After Inflation, the inflaton field ϕ begins to oscillate at the bottom of the potential $V(\phi)$ since it transforms the initial potential into kinetic energy. Near the minimum of $V(\phi)$, we can approximate it as its quadratic form $V(\phi) = \frac{1}{2}m^2\phi^2$, where ϕ has to be assumed small in amplitude. Replacing this approximation into eq. (3.26) gives

$$\ddot{\phi} + 3H\dot{\phi} + m^2\phi = 0\,, \tag{3.78}$$

which for $H \sim cte$, the eq. (3.78) is just the equation for a damped harmonic oscillator in which if the damping coefficient $(3H)^{-1}$ is much shorter than the frequency m, we can neglect the friction term. Using the continuity eq. (1.39) for the inflaton field and its pressure in eq. (3.22) we can describe its behavior as

$$\begin{aligned} \dot{\rho}_\phi + 3H\rho_\phi =& -3HP_\phi \\ =& \frac{3}{2}\left(m^2\phi - \dot{\phi}^2\right)\,, \end{aligned} \tag{3.79}$$

where the term in parenthesis averages to zero over one oscillation period considering $H^{-1} \gg m^{-1}$, therefore the oscillating field behaves like pressureless matter.

The inflaton must decay into Standard Model fields to avoid that our Universe ends up empty and to explain the components that we know. If the inflaton decays into bosons, the process will be very rapid and will be far from thermal equilibrium, a process known as preheating. But if the decay is slow, in which case the inflaton would decay into fermions, the energy density behaves as

$$\dot{\rho}_\phi + 3H\rho_\phi = -\Gamma_\phi \rho_\phi \,, \tag{3.80}$$

where Γ_ϕ is the inflaton decay rate. However, we need both bosons and fermions in our Universe, and then the decay is more complicated that we explained here. Also, the particles created will interact and create other particles, and this soup will eventually reach thermal equilibrium at some temperature T_{rh} in a process called thermalization, at least for the particles whose interactions are not so weak. The energy density determines this reheating temperature at the end of reheating, which necessarily has to be less than the inflaton energy density at the end of Inflation, but if reheating takes a long time, the difference between the two could be large and affect the value of T_{rh}. In the end, the energy density of this soup behaves like radiation as long as the momenta of the particles are much higher than their masses, and finally, the Hot Big Bang era begins after the complete thermalization of at least the baryons, photons, and neutrinos.

3.3 Initial Conditions

After preheating ends, we can start to use the system of equations studied in chapter 2, which in order to find solutions need initial conditions. Since we are going to track eight variables $(\Theta, \mathcal{N}, \delta_{\rm cdm}, v_{\rm cdm}, \delta_b, v_b, \Phi, \Psi)$ we need eight initial conditions, but due to the inflationary epoch in practice we only need one because Inflation lefts our Universe in a very homogeneous initial state, then all the variables are connected at the beginning with the potential Φ.

At very early times, we have $k\tau \ll 1$, which implies that all the equations must be worked in the super-horizon limit. Physically, this means that all the perturbations have wavelengths much larger than the distance over which causal physics operates. Therefore, the eq. (2.250) takes the form

$$\mathcal{N}_0' + \phi' = 0 \,, \tag{3.81}$$

while for the eq. (2.208) we need to consider that τ_{op}' (eq. (2.206)) is extremely large at that time, then the parenthesis that it multiplies must to be zero in order to preserve the order of magnitude in eq. (2.208). Thus it takes the form

$$\Theta_0' + \phi' = 0 \,, \tag{3.82}$$

where only the monopole survives in both photons and neutrinos because the perturbations have wavelengths much larger than the causal horizon, making higher multipoles much smaller. Implying that for a hypothetical observer who, for example, sees only photons within his causal horizon, will see a uniform sky. Hence the monopoles are directly connected with the potential ϕ and the same happens with the density contrast for baryons and cold dark matter, since eqs. (2.234) and (2.246) at early times become in

$$\delta_b' + 3\phi' = 0 \,, \tag{3.83}$$

$$\delta'_{\mathrm{cdm}} + 3\phi' = 0\,. \tag{3.84}$$

Since in this period radiation dominates, we can use the Einstein equations considering only it as a single fluid even when it is compound by photons and neutrinos. Therefore, we will go back until the result that we discovered studying the equation of potential for super-horizon scales (eq. (2.122)), which says us that in this regime $\phi' = 0$ and then all the quantities in eqs. (3.81) to (3.84) also vanish. Taking the eq. (2.111) in the super-horizon limit and replacing $\bar{\rho}\delta = \bar{\rho}_r\delta_r$ we have

$$\frac{\phi'}{\mathcal{H}} - \psi = 2\left(\left[1 - f_\nu\right]\Theta_0 + f_\nu\mathcal{N}_0\right)\,, \tag{3.85}$$

where we also used eqs. (2.70) and (2.249). In general, the anisotropic stress of radiation is proportional to the quadrupole at the first order of neutrinos $\mathcal{N}_2$ and photons Θ_2, but as we are neglecting higher multipoles at the early time we can consider that they vanish and by eq. (2.105) set that $\psi = -\phi$, which implies that

$$\phi = 2\left(\left[1 - f_\nu\right]\Theta_0 + f_\nu\mathcal{N}_0\right)\,. \tag{3.86}$$

We also know that both Θ_0 and $\mathcal{N}_0$ are constant at early times, but also they are equal in this regime because the models of structure formation do not predict differences in the perturbations of photons and neutrinos. Therefore eq. (3.86) becomes in

$$\phi = 2\Theta_0 = 2\mathcal{N}_0\,. \tag{3.87}$$

For the initial conditions of the density contrasts we need to make an assumption, since if we integrate eqs. (3.83) and (3.84) we will obtain $\delta_b = 3\Theta_0 + cte$, where we also used eq. (3.82) before integrating. To determinate that constant, we are going to assume that the primordial perturbations are adiabatic, then by eqs. (2.155) and (2.248) $\delta_b = \delta_{\mathrm{cdm}} = 3\delta_r/4 = 3\Theta_0$. Thus primordial adiabatic fluctuations make that the constant vanishes and set the initial conditions for the density contrast

$$\delta_b = \delta_{\mathrm{cdm}} = 3\Theta_0\,. \tag{3.88}$$

Of course, we can reject that assumption and use its opposite, the isocurvature fluctuations; however, the models based on that type of perturbations have not been very successful.

Finally, we need the initial conditions for the velocities, and we are going to start using again the argument that the optical depth is huge at early times but now in eq. (2.244), then the parenthesis that is being multiplied by τ'_{op} must vanishes to preserve the order of magnitude in that equation, which lead to us

$$v_b = v_{\mathrm{cdm}} = -3i\Theta_1\,; \tag{3.89}$$

but, as we are considering that higher multipoles that the monopole are considerably small, we also can neglect the initial conditions for the velocities.

3.4 Thermal History

Many events took place in the three minutes after the Hot Bing Bang starts, due principally to the high temperatures of our primordial Universe. In this section, we will make a brief description of those processes and how they shaped our Universe.

In the first picosecond in the history of our Universe we talk about the very early Universe, it was at the Planck epoch in which the currently understood laws of physics might not apply. During this period, it is believed that a Grand Unified Theory (GUT) governs at a temperature of 10^{27}-10^{28}eV, from which the four fundamental interactions emerged. The first was the gravity approximately at 10^{-43}s and then the electromagnetic, weak, and strong interactions. It is also believed that Inflation started when the symmetry of GUT was spontaneously broken, and the Universe had a temperature of $10^{22} - 10^{23}$eV; then, it is cooled until the absolute zero by the ability of Inflation at the 10^{-33}s. Finally, the reheating process increased the temperature again until 10^{22}-10^{23}eV.

When the Big Bang scenario begins, we talk about the early Universe, in which the expansion of our Universe also started to cool it. The key to understanding the processes of the early Universe is to compare the rate of expansion given by the Hubble parameter H with the rate of the interactions involved Γ. When $\Gamma \gg H$, the time scale of particle interactions is much smaller than the time scale of expansion, hence we can say that local equilibrium is reached before the effect of expansion becomes relevant. Nevertheless, when $\Gamma \sim H$, the particles decouple from the thermal bath, which happens at different times since they have different interaction rates.

As the primordial temperatures were high, all known particles were ultra-relativistic. Also, we can write their interaction rate as $\Gamma = n_2 \sigma v$, where n_2 in the number density of the other particles involved in the bath, v is the average relative velocity, and σ is the interaction cross-section which can be computed from the fundamental physics of the process. Therefore, as the different particles start to decouple from the thermal bath, we have less and less relativistic particles every time. Since radiation dominates (RDE), the Hubble parameter is proportional to the squared temperature $H \sim T^2$ (eqs. (1.69), (2.132) and (2.188))[4]; then, the decoupling is characterized exclusively by the dependence of the cross-section on the temperature.

At approximately 20ps, the electroweak phase transition happens at 100GeV, from which the gauge bosons of weak interactions $W^{\pm}$ and Z received their masses through the Higgs mechanism. Then, it was followed by the QCD phase transition at 20μs and 150MeV, where the interactions between quarks and gluons become important due to the strong force, forming bound systems like baryons (three quarks) and mesons (quark-antiquark pairs). When the temperature drops to 1MeV at 1s, the particles that interact only due to the weak force decouple, for example, neutrinos. This causes that higher moments, that the dipole, describe the neutrino distribution $\mathcal{N}$, which makes that neutrinos develop anisotropic stress since it is proportional to the quadrupole according to eq. (2.252). Therefore, by eqs. (2.253) and (2.254), the potentials Φ and Ψ differ between each other after neutrino decoupling, noting that the difference is proportional to the radiation density, which decreases on time.

When the temperature drops until 500KeV at 6s, electrons and positrons annihilate each other; then, the energy of that process is transferred to photons due to electromagnetic interactions such

$$\mathrm{e}^- + \mathrm{e}^+ \leftrightarrow \gamma + \gamma\,. \tag{3.90}$$

This also means that the photons are heated relative to the neutrinos. Then, the last ones did not receive any energy from this process because they were already decoupled. However,

---

[4] This is valid at least until the process of Big Bang nucleosynthesis.

a small neutrino fraction was not decoupled at that time, and a bit of energy is transferred to them. This changes the effective number of neutrinos from $N_{\mathrm{e}ff} = 3$ to $N_{\mathrm{e}ff} = 3.046$. Nonetheless, precise measurements from Planck [59] have constrained its value to $N_{eff} = 3.36$, which is a problem to explain nowadays. Finally, when the Universe is cooled until 100keV in the third minute, the light elements like Hydrogen and Helium were formed in a process called Big Band Nucleosynthesis, which we are going to discuss in the next section.

The relativistic quantum field theory requires particles and anti-particles, but this poses a slight puzzle because they annihilate each other like electrons with positrons. If the Universe was filled with an equal abundance of matter and anti-matter, we expect that this lead to a Universe dominated by radiation. However, we observe an overabundance of matter over anti-matter in our Universe today. For electrons and positrons, we already saw that they share the same abundance, but in the case of baryons, we need some mechanism to explain the overabundance, known as Baryogenesis. There are many models of Baryogenesis that try to explain this mystery through a dynamical mechanism or assuming a primordial matter-antimatter asymmetry as an initial condition, but not of them have been already tested.

Another mystery is the time when Dark Matter decouples from the thermal bath. As we know, Dark Matter is weakly interacting with ordinary matter, so we expect that it decouples relatively early on. Candidates for Dark Matter particles are the WIMPs that are weakly interacting massive particles, and it is expected that they freeze out around 1MeV.

3.4.1 Big Bang Nucleosynthesis

The end of this section is to determine how the baryons produced after the QCD phase transition end up. We need to trace the baryonic matter, which is much fewer in number than relativistic species. However, it obeys the baryon number conservation law, which says that the total number of nucleons (protons and neutrons) stays constant. Then we need to track them solving a Boltzmann equation similar to eq. (2.173), which for a process of type $1+2 \leftrightarrow 3+4$ can be written as

$$\begin{aligned} a^{-3}\frac{\mathrm{d}\left(n_1 a^3\right)}{\mathrm{d}t} = & \int \frac{\mathrm{d}^3 p_1}{(2\pi)^3\varepsilon_1} \int \frac{\mathrm{d}^3 p_2}{(2\pi)^3\varepsilon_2} \int \frac{\mathrm{d}^3 p_3}{(2\pi)^3\varepsilon_3} \int \frac{\mathrm{d}^3 p_4}{(2\pi)^3\varepsilon_4} \\ & (2\pi)^4 \delta^3\left(p_1+p_2-p_3-p_4\right) \delta\left(\varepsilon_1+\varepsilon_2-\varepsilon_3-\varepsilon_4\right) \\ & |\mathcal{M}|^2 \left\{ f_3 f_4 \left[1 \pm f_1\right]\left[1 \pm f_2\right] - f_1 f_2 \left[1 \pm f_3\right]\left[1 \pm f_4\right] \right\} , \end{aligned} \tag{3.91}$$

which is the integrated form of eq. (2.173), and where its terms are explained in sections 2.9.1 and 2.9.2. For nucleosynthesis and other processes, we will typically be interested in temperatures $T \ll \varepsilon - \mu$, where μ is the chemical potential. This allows us to ignore the complications of quantum statistics and just use the Boltzmann distribution $f \sim \exp\left(-(\varepsilon-\mu)/T\right)$ instead of the Bose-Einstein or Fermi-Dirac statistics. Besides, we can neglect the Pauli blocking or the Bose enhancement factors in eq. (3.91) (those in square brackets). Finally, we have to deal with

$$a^{-3}\frac{\mathrm{d}\left(n_1 a^3\right)}{\mathrm{d}t} = n_1^{\mathrm{eq}} n_2^{\mathrm{eq}} \langle \sigma v \rangle \left\{ \frac{n_3 n_4}{n_3^{eq} n_4^{\mathrm{eq}}} - \frac{n_1 n_2}{n_1^{\mathrm{eq}} n_2^{eq}} \right\} , \tag{3.92}$$

in which the n_i are defined in the same way that for electrons in eq. (2.197), but now considering the Boltzmann distribution

$$n_i = g_i \mathrm{e}^{\mu_i/T} \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \mathrm{e}^{-\varepsilon_i/T} \,, \tag{3.93}$$

where g_i is the degeneracy of the specie i, while $\langle \sigma v \rangle$ is the thermally averaged cross-section defined as

$$\begin{aligned} \langle \sigma v \rangle \equiv \frac{1}{n_1^{\mathrm{eq}} n_2^{\mathrm{eq}}} \int \frac{\mathrm{d}^3 p_1}{(2\pi)^3 \varepsilon_1} \int \frac{\mathrm{d}^3 p_2}{(2\pi)^3 \varepsilon_2} \int \frac{\mathrm{d}^3 p_3}{(2\pi)^3 \varepsilon_3} \int \frac{\mathrm{d}^3 p_4}{(2\pi)^3 \varepsilon_4} \mathrm{e}^{-(\varepsilon_1+\varepsilon_2)/T} \\ (2\pi)^4 \delta^3 (p_1 + p_2 - p_3 - p_4)\, \delta(\varepsilon_1 + \varepsilon_2 - \varepsilon_3 - \varepsilon_4) \left|\mathcal{M}\right|^2 , \end{aligned} \tag{3.94}$$

and the n_i^{eq} are the equilibrium number density[5], which are given by

$$n_i^{eq} \equiv g_i \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \mathrm{e}^{-\varepsilon_i/T} = \begin{cases} g_i \left(\frac{m_i T}{2\pi}\right)^{3/2} \mathrm{e}^{-m_i/T} & m_i \gg T \\ g_i \frac{T^3}{\pi^2} & m_i \ll T \end{cases} . \tag{3.95}$$

We can note that the l.h.s. of eq. (3.91) if proportional to the expansion rate *i.e.* H, while the r.h.s. is proportional to the rate of interactions Γ, which in the processes that we want to seek is much larger than H because they happened before that species decoupled. Thus, to preserve the order of magnitude in eq. (3.91) its r.h.s. parenthesis must vanishes, which leads us with the Saha equation of chemical equilibrium

$$\frac{n_3 n_4}{n_3^{eq} n_4^{eq}} = \frac{n_1 n_2}{n_1^{eq} n_2^{eq}} \,. \tag{3.96}$$

Therefore we must solve eq. (3.96) for all the nuclei interactions, which gives us a broad set of coupled ODEs. However, we can make two simplifications:

1. The first simplification is to consider elements that are not heavier than Helium, because they are not produced at appreciable levels in the early Universe[6]. So the only nuclei that we need to track are Hydrogen, Helium, and their stable isotopes (Deuterium, Tritium, and ^{3}He).
2. The second simplification is to solve the neutron/proton ratio first, since above 0.1MeV only free protons and neutrons exits, while other light nuclei have not been formed yet. Finally, we can use this ratio as input for the synthesis of Helium and the isotopes.

For the second simplification, we need first to determinate the relative abundance of neutrons and protons, which in the early Universe are coupled by weak interactions like the β-decay and inverse β-decay

$$n + \nu_\mathrm{e} \leftrightarrow p^+ + \mathrm{e}^- \tag{3.97}$$

$$n + \mathrm{e}^+ \leftrightarrow p^+ + \bar{\nu}_\mathrm{e} \,. \tag{3.98}$$

Assuming that the chemical potential for neutrinos and electrons are small compare to neutrons and protons we can make the approximation $\mu_n = \mu_p$, which used in eq. (3.96) gives

$$\left(\frac{n_n}{n_p}\right)_{\mathrm{eq}} = \left(\frac{m_n}{m_p}\right)^{3/2} \mathrm{e}^{-(m_n - m_p)/T} \propto \mathrm{e}^{-Q/T} \,, \tag{3.99}$$

---

[5] We talk about chemical equilibrium.

[6] An exception is the Lithium, because its trace abundance may be observable today.

where $Q \equiv m_n - m_p = 1.30\text{MeV}$, concluding that for $T \gg 1\text{MeV}$ there are as many neutrons as protons. However, for $T < 1\text{MeV}$, the neutron fraction gets smaller, which would drop to zero the neutron abundance for very efficiently weak interaction. Luckily, the weak interactions are not so efficient, at least in this context.

We are going to consider the production of Deuterium (one proton and one neutron), through the following reaction:

$$n + p^+ \leftrightarrow D + \gamma \,, \tag{3.100}$$

where we have $\mu_n + \mu_p = \mu_D$ since $\mu_\gamma = 0$. Using again the Saha equation (eq. (3.96)) and the eq. (3.95) with $g_D = 3$ for the Deuterium, we get the next ratio

$$\begin{aligned} \left(\frac{n_D}{n_n n_p}\right)_{\text{eq}} &= \frac{3}{4}\left(\frac{2\pi m_D}{m_n m_p T}\right)^{3/2} \mathrm{e}^{-(m_D - m_n - m_p)/T} \\ \Rightarrow \left(\frac{n_D}{n_p}\right)_{\text{eq}} &= \frac{3}{4} n_n^{\text{eq}} \left(\frac{2\pi m_D}{m_n m_p T}\right)^{3/2} \mathrm{e}^{B_D/T} \,, \end{aligned} \tag{3.101}$$

in which $B_D \equiv m_n + m_p - m_D = 2.22\text{MeV}$. Nonetheless, the parenthesis in the r.h.s. of eq. (3.101) is too small, and then the temperature has to drop enough $T \ll B_D$ if we want that the exponential factor can compete. The same applies to all the other nuclei, but at temperatures around 0.1meV, the exponential factor is competitive, starting with the production of Deuterium and Helium. However, their reaction rates are too low even to think to produce heavier elements.

Before the neutron freeze-out at $T \sim 0.8\text{MeV}$, neutrons are in equilibrium with protons due to weak interactions, which also ends at neutrino decoupling. After that, we must solve the Boltzmann equation (eq. (3.91)) instead of the Saha equation (eq. (3.96)) to track the neutron abundance. For that we first are going to define the neutron fraction as

$$X_n \equiv \frac{n_n}{n_n + n_p} \,, \tag{3.102}$$

which in equilibrium and following eq. (3.99), takes the form

$$X_n^{\text{eq}}(T) = \frac{\mathrm{e}^{-Q/T}}{1 + \mathrm{e}^{-Q/T}} \,. \tag{3.103}$$

At neutrino decoupling the neutron fraction is $X_n(T = 0.8\text{MeV}) = 0.17$, but we also need to consider that the neutron decay exists because below 0.2MeV ($t > 100\text{s}$) it is $\tau_n = 806.7\text{s}$ which is in the same order of magnitude. We can consider it easily, just multiplying the freeze-out abundance by a damping exponential term

$$X_n(t) = X_n^{\infty} \mathrm{e}^{-t/\tau_n} \,, \tag{3.104}$$

where $X_n^{\infty} \equiv X_n(Q/T \to \infty)$ is a constant of the freeze-out neutron abundance that does not change a lot in comparison with the value at $T = 0.8\text{MeV}$[7].

Once we used the first simplification, we can consider the second to compute the production of Helium. At this time, our Universe is mostly protons and neutrons, which could form

---

[7] The complete deduction with full of details can be found in [80, 31].

Helium directly through reactions with three or four incoming nuclei, but the density is too low, and the time available is too short for that type of process. Then, the heavier nuclei have to be built sequentially from lighter nuclei in two-particles reactions, which are

$$n + p^+ \longleftrightarrow D + \gamma\,, \tag{3.105}$$

$$D + p^+ \longleftrightarrow {}^3H\mathrm{e} + \gamma\,, \tag{3.106}$$

$$D + {}^3H\mathrm{e} \longleftrightarrow {}^4H\mathrm{e} + p^+\,. \tag{3.107}$$

The reactions in eqs. (3.106) and (3.107) happen only when the Deuterium produced by eq. (3.105) is available. The Deuterium follows the equilibrium abundance of neutrons and protons as long as enough free neutrons are available, but as we saw in eq. (3.101), the Deuterium abundance becomes large rather late. Then, even that heavier nuclei have larger equilibrium abundances since they have large binding energies, they can not be formed until sufficient Deuterium has become available. This effect is known as the Deuterium bottleneck.

To estimate roughly the time of nucleosynthesis, we need first to determine the temperature when the Deuterium fraction of eq. (3.101) is of order one, which corresponds to 0.06MeV. The time corresponding to that temperature is approximately 330s because the Hubble parameter scales as T^2, and it obeys $H \propto t^{-1}$ during RDE. Finally and by eq. (3.104), the neutron fraction at that time is $X_n(t_{NucleoSynth}) \sim 1/8$.

Since the bending energy of Helium is larger than Deuterium, the Boltzmann factor $\mathrm{e}^{B/T}$ favors the first. Helium is produced almost immediately after Deuterium. Now, if we consider that all the neutrons at $t \sim t_{NucleoSynth}$ are processed into ${}^4H\mathrm{e}$, the final ${}^4H\mathrm{e}$ abundance is equal to half of the neutron abundance at $t_{NucleoSynth}$, because there are two neutrons in one ${}^4H\mathrm{e}$ atom. Finally, we get

$$\frac{n_{H\mathrm{e}}}{n_H} = \frac{n_{H\mathrm{e}}}{n_p} \simeq \frac{\frac{1}{2}X_n(t_{NucleoSynth})}{1 - X_n(t_{NucleoSynth})} \sim \frac{1}{2}X_n(t_{NucleoSynth}) = \frac{1}{16}\,. \tag{3.108}$$

Forgetting all the approximations used in the result of eq. (3.108), we have to remark that it is consistent with the exact solution and the observed Helium fraction in our Universe. Nonetheless, we can not forget that this solution depends on several input parameters that take account of the physics behind the process, so measuring the Helium fraction also could tell us about possible new physics beyond the Standard Model.

Finally, to determine the abundance of the other light elements (as Lithium), the coupled Boltzmann equations of eq. (3.91) have to be solved numerically, whose solutions have a reasonably good quantitative agreement with observations, making a great triumph for the Big Bang model. The elements that are heavier than Beryllium and its isotopes are not produced in the early Universe, but they are in the core of stars only up to Fe. The heaviest ones are produced in the explosion of stars or labs, as long as we know.

3.4.2 Matter-Radiation Equality

In section 1.4, we saw that there are periods in which one of the different components of our Universe dominates and dictates the behavior of the background quantities. The first epoch is RDE since many of the components are relativistic, as we tracked in the last section.

Nonetheless, the primordial plasma left by the early Universe suffers a transition from RDE to MDE, which implies that it is no longer modeled by the relativistic degrees of freedom, even when they are not completely decoupled. We can find the time when the transition happens matching the energy density of radiation with that of matter

$$a_{\rm eq} = \frac{1}{1+z_{\rm eq}} = \frac{\rho_{0,r}}{\rho_{m,0}} = \frac{\Omega_{0,r}}{\Omega_{m,0}}\,, \tag{3.109}$$

where we used eqs. (1.35), (1.48) and (1.68). Those values [59] are approximately $a_{\rm eq} \approx 2,9 \times 10^{-4}$ and $z_{\rm eq} \approx 3411$. Note that the compute made in eq. (3.109) is just a reference, since we ignored the other species. Also the transition is not instantaneously and then we also need to consider that for values near $a_{\rm eq}$ there is not a domination of one component from another.

One of the quantities that are affected by this transition is the potential Φ. To see this, we are going to write the comoving curvature perturbation in Newtonian gauge (eq. (3.69)) at the super-horizon limit, in which $\Phi' = 0$, as we derived in section 2.7.1. Then, eq. (3.69) can be written as

$$\mathcal{R} = \frac{5+3w}{3\,(1+w)}\Phi\,, \tag{3.110}$$

where we used eq. (2.70) and $\Psi = -\Phi$ because none of the species involved has anisotropic stress. Since $\mathcal{R}$ remains constant on super-horizon scales due eq. (3.71), we have that $\mathcal{R}_{RDE} = \mathcal{R}_{MDE}$ and then

$$\frac{3}{2}\Phi_{RDE} = \frac{5}{3}\Phi_{MDE} \Rightarrow \Phi_{MDE} = \frac{9}{10}\Phi_{RDE}\,, \tag{3.111}$$

meaning that the potential decreases 9/10 in the transition from RDE to MDE at super-horizon scales.

We also noted that the decoupling of neutrinos causes a difference between Φ and Ψ during RDE, but once entered in MDE, the radiation density becomes insignificant. Like matter has not anisotropic stress; after matter-radiation equality, both potentials rapidly approach each other.

3.4.3 Recombination

Before the temperature drops to 1eV, our Universe is still a coupled plasma of photons, free electrons, and nuclei. The photons were tightly coupled to electrons via Compton scattering, and the protons and electrons were strongly interacting via Coulomb scattering. There was very little neutral Hydrogen produced by the reaction

$$\mathrm{e}^- + p^+ \leftrightarrow H + \gamma\,, \tag{3.112}$$

whose production is favored because its binding energy is 13.6eV, but the high photon to baryon ratio ensures that any atom of Hydrogen produced will be ionized. As long as the reaction in eq. (3.112) remains in equilibrium the eq. (3.96) ensures that

$$\left(\frac{n_H}{n_{\rm e} n_p}\right)_{\rm eq} = \frac{g_H}{g_{\rm e} g_p}\left(\frac{m_H}{m_{\rm e} m_p}\frac{2\pi}{T}\right)^{3/2} \mathrm{e}^{-(m_H - m_p - m_{\rm e})/T}\,, \tag{3.113}$$

since $\mu_\gamma = 0$ and the $m_i > T$. Also we can define $B_H \equiv m_p + m_{\mathrm{e}} - m_H = 13.6\mathrm{eV}$ and the free electron fraction as

$$X_{\mathrm{e}} \equiv \frac{n_{\mathrm{e}}}{n_{\mathrm{e}} + n_H} \,. \tag{3.114}$$

We can make the approximations $n_{\mathrm{e}} \sim n_p$ and $m_H \sim m_p$, where in the first we are assuming that our Universe is neutral. Also the product $g_H/g_{\mathrm{e}}/g_p$ in eq. (3.113) is 1, since the electron and proton have degeneracy 2, while the neutral Hydrogen has degeneracy 4 (one singlet state and one triplet state). Therefore and using eq. (3.114), the eq. (3.113) becomes in

$$\frac{X_{\mathrm{e}}^2}{1 - X_{\mathrm{e}}} = \frac{1}{n_{\mathrm{e}} + n_H} \left(\frac{m_{\mathrm{e}} T}{2\pi} \right)^{3/2} \mathrm{e}^{-B_H/T} \,, \tag{3.115}$$

which is an excellent approximation until the temperature drops to the order of B_H, because at that temperature the r.h.s. is of order 10^{15} computed neglecting the number of Helium atoms[8]. To keep the magnitude, the l.h.s. must have a very small denominator, which implies that X_{e} is very close to 1 *i.e.* all Hydrogen is ionized. Therefore, we need that the temperature drops far below B_H to obtain neutral Hydrogen, a process which is known as recombination. Nonetheless, as X_{e} falls, the recombination rate too, and the equilibrium becomes more difficult to maintain, which discredits eq. (3.113). Thus, if we want to follow the free electron fraction accurately, we need to solve the Boltzmann equation (eq. (3.91)). However, eq. (3.115) predicts very well the temperature of recombination $T_{rec} \approx 0.3\mathrm{eV}$, which after using eqs. (1.48) and (2.188) says that it happens at redshift $z_{rec} \approx 1320$ implying that recombination starts when we are already in MDE.

3.4.4 Photon Decoupling

The photons are most strongly coupled to the primordial plasma due to Compton scattering (eq. (2.190)). Which has an interaction rate given by $\Gamma_\gamma \simeq n_{\mathrm{e}}\sigma_T = \sigma_T n_b X_{\mathrm{e}}$, then as we know, the photons will decoupled when $\Gamma_\gamma \sim H$. Therefore, we need to track the free electron fraction using eq. (3.91), which gives a temperature of $T_{\mathrm{dec}} \approx 0.27\mathrm{eV}$ for the photon decoupling. We have to note that this value is not far from T_{rec}, which implies that a large degree of neutrality is needed in our Universe to become transparent to photon propagation. Using eq. (2.188), we also can know that the photon decoupling takes place at redshift $z_{\mathrm{dec}} \approx 1100$. These last photons released are known as the last scattering surface, which nowadays are watched as the Cosmic Microwave Background (CMB).

The photon decoupling causes a similar effect in the potentials that neutrinos, but the Universe is already in MDE, and then this difference in potentials becomes insignificant rapidly. Thus, higher multipoles than the dipole describe the photon distribution after decoupling even when decreasing in magnitude. To track them, we will start studying the equations for photon distribution in the tightly coupled limit before decoupling. Hence, we are going to use the equations found in section 2.9, specifically for photon distribution but without considering polarization, *i.e.* eq. (2.205) instead eq. (2.208). We are going to multiply eq. (2.205) by $\mathcal{P}_0(\mu)$ and $\mathcal{P}_1(\mu)$ and then, we will proceed to integrate them in the range $\mu \in [-1, 1]$ to obtain

$$\Theta_0' + k\Theta_1 + \phi' = 0 \,, \tag{3.116}$$

---

[8] The denominator of eq. (3.115) is then $n_{\mathrm{e}} + n_H = n_b$, where n_b is the total number density of baryons.

$$\Theta_1' - \frac{k}{3}\Theta_0 - \frac{k}{3}\psi = \tau_{op}'\left(\Theta_1 - \frac{i}{3}v_b\right) \,, \tag{3.117}$$

which are complemented by the equation for baryon velocity in eq. (2.244), rewritten as

$$v_b = -3i\Theta_1 + \frac{R_s}{\tau_{op}'}\left(v_b' + \mathcal{H}v_b + ik\psi\right) \,, \tag{3.118}$$

where the second term in the r.h.s. is much smaller than the first since the ratio R_s/τ_{op}' is small before recombination. Then, to lowest order we have $v_b \simeq -3i\Theta_1$, which we will replace into eq. (3.118) to obtain

$$v_b \simeq -3i\Theta_1 + \frac{R_s}{\tau_{op}'}\left(-3i\Theta_1' - 3i\mathcal{H}\Theta_1 + ik\psi\right) \,. \tag{3.119}$$

Now we are going to replace eq. (3.119) into eq. (3.117)

$$\Theta_1' + \mathcal{H}\frac{R_s}{1+R_s}\Theta_1 - \frac{k}{3\left((1+R_s\right)}\Theta_0 = \frac{k}{3}\psi \,, \tag{3.120}$$

then, we are going to take the derivative of eq. (3.116), where later we will replace the Θ_1' term using eq. (3.120) and the Θ_1 term using eq. (3.116), which gives

$$\Theta_0'' + \phi'' + \mathcal{H}\frac{R_s}{1+R_s}\left(\Theta_0' + \phi'\right) + \frac{k^2}{3}\psi + k^2c_s^2\Theta_0 = 0 \,, \tag{3.121}$$

where we defined the sound speed of the photon-baryon fluid as

$$c_s^2 \equiv \frac{1}{3\left(1+R_s\right)} \,. \tag{3.122}$$

We also can rewrite eq. (3.121) using that $R_s' = \mathcal{H}R_s$ due to eqs. (1.68) and (2.245), which finally gives

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + \frac{R_s'}{1+R_s}\frac{\mathrm{d}}{\mathrm{d}\tau} + k^2c_s^2\right]\left(\Theta_0 + \phi\right) = \frac{k^2}{3}\left(\frac{\phi}{1+R_s} - \psi\right) \,, \tag{3.123}$$

which is nothing more than the equation for a harmonic oscillator forced and damped. To solve eq. (3.123), we can first find the solutions of the homogeneous equation, and we can use the method of Green to find the full solution. The pressure term $c_s^2k^2$ is much larger than the damping term $R_s'/(1+R_s)$ at least for modes inside the horizon or in the tight-coupling regime ($R_s \ll 1$), we can neglect the second to find the homogeneous solutions. This is equivalent to say that pressure induces oscillations with a period much shorter than the damping induced by the expansion. Therefore, the homogeneous solution is a combination of sinusoidal functions

$$\left(\Theta_0 + \phi\right)^{(hom)}(k,\tau) = C_1 f_1(k,\tau) + C_2 f_2(k,\tau) \,, \tag{3.124}$$

where C_1, C_2 are integration constants and f_1, f_2 are

$$f_1 = \sin\left(kr_s\left(\tau\right)\right) \,, \tag{3.125}$$

$$f_2 = \cos\left(kr_s\left(\tau\right)\right) \,, \tag{3.126}$$

in which we defined the sound horizon as

$$r_s \equiv \int_0^{\tau} c_s\left(\tau'\right) \mathrm{d}\tau' \,, \tag{3.127}$$

that is the comoving distance traveled by a sound wave by time τ. Thus, the full solution in the tight-coupling regime is

$$\begin{aligned} \Theta_0\left(k,\tau\right) + \phi\left(k,\tau\right) = &\, C_1 f_1\left(k,\tau\right) + C_2 f_2\left(k,\tau\right) \\ &+ \frac{k^2}{3} \int_0^{\tau} \mathrm{d}\tau' \left[\phi(\tau') - \psi(\tau')\right] \frac{f_1(\tau')f_2(\tau) - f_1(\tau)f_2(\tau')}{f_1(\tau')f_2'(\tau') - f_1'(\tau')f_2(\tau')} \,, \end{aligned} \tag{3.128}$$

where we omitted the $k-$dependence in the terms inside the integral for simplicity. To set the constants C_1, C_2 we have to review the initial condition when both Θ_0 and ϕ are constants, then C_1 must vanishes and $C_2 = \Theta_0(0) + \phi(0)$. Finally, in the limit that we are working, the denominator of the integral reduces to $-kc_s(\tau') \to -k/\sqrt{3}$ and the numerator can be written just as $-\sin\left(k\left[r_s(\tau) - r_s(\tau')\right]\right)$, so

$$\begin{aligned} \Theta_0\left(k,\tau\right) + \phi\left(k,\tau\right) = &\left[\Theta_0(0) + \phi(0)\right] \cos\left(k r_s(\tau)\right) \\ &+ \frac{k}{\sqrt{3}} \int_0^{\tau} \mathrm{d}\tau' \left[\phi(\tau') - \psi(\tau')\right] \sin\left(k\left[r_s(\tau) - r_s(\tau')\right]\right) \,, \end{aligned} \tag{3.129}$$

which seems redundant since we are integrating over the potentials, but we must consider that we are in MDE, and the potentials are sourced mostly by dark matter, which makes them independent of photon distribution. The approximate result of eq. (3.129) also is important for other stuff detailed in [80], where one of them is that it gives an accurate expression for the frequency of oscillations called acoustic oscillations and therefore for the locations of the acoustic peaks. Those are found in the limit when the first term in eq. (3.129) dominates, then the peaks should appear at the extrema of $\cos\left(kr_s\right)$ given by

$$k_p = \frac{n\pi}{r_s} \qquad n = 1, 2, ... \qquad . \tag{3.130}$$

These oscillations are imprinted in the CMB temperature anisotropies even after photons decouple, which makes it possible to observe nowadays and reveal much information about the early Universe. An example is the comoving distance traveled by the sound waves until photon decoupling, which is characterized by the sound horizon, and also behaves as a standard yardstick.

3.4.5 Drag Epoch

After the photon decoupling the baryons and photons do not evolve as a single fluid anymore. Nonetheless, since there are more photons that baryons ($\rho_\gamma > \rho_b$), the photons decouple before that the baryons released from the Compton drag of photons. This epoch is known as drag epoch, which happened not far from photon decoupling, at approximate redshift $z_{drag} = 1020$. From then, all photons expand freely, and the baryons progressively fall into dark matter potentials wells; and causally, dark matter also is attracted to baryon overdensities.

We can study the behavior of the potentials using eq. (2.101), where we have to specify its sources: cold dark matter and radiation. Note that we are neglecting the baryons, which is a mistake that we will resolve later. For the dark matter we can use the eqs. (2.246) and (2.247) and for radiation we can use the eqs. (3.116) and (3.117). However, we ignore the term with τ'_{op} in eq. (3.117) for two reasons, the first is that for radiation we are considering both photons and neutrinos, where the last does not contain that term in its equations (see eq. (2.250)). The second reason is that we want to track the potentials that are sourced mostly by matter since we are in MDE. We are tracking the matter distribution, and what photons are doing after decoupling is irrelevant for our purpose, which allows us to neglect all the multipoles higher that the dipole even when we know that they describe, in fact, the photon distribution after photon decoupling. So, the equations to solve are

$$\Theta'_{r,0} + k\Theta_{r,1} + \phi' = 0\,, \tag{3.131}$$

$$\Theta'_{r,1} - \frac{k}{3}\Theta_{r,0} + \frac{k}{3}\phi = 0\,, \tag{3.132}$$

$$\delta'_{\mathrm{cdm}} + ikv_{\mathrm{cdm}} + 3\phi' = 0\,, \tag{3.133}$$

$$v'_{\mathrm{cdm}} + \mathcal{H}v_{\mathrm{cdm}} - ik\phi = 0\,, \tag{3.134}$$

noting that the subscript r here refers to radiation, both neutrinos and photons, since they obey the same equation in the limit of small baryon density and with the same initial conditions (section 3.3). We also set the potentials $\psi = -\phi$, which is valid in the limit that there is no quadrupole moment (eq. (2.254)). Then, we only need one equation for the potential (eqs. (2.101), (2.102) and (2.104)), where we choose the eq. (2.101)

$$k^2\phi + 3\mathcal{H}\left(\phi' + \mathcal{H}\phi\right) = 4\pi G a^2\left(\bar{\rho}_{\mathrm{cdm}}\delta_{\mathrm{cdm}} + 4\bar{\rho}_r\Theta_{r,0}\right)\,, \tag{3.135}$$

in which we already set $\psi = -\phi$ and where only the cold dark matter and the radiation are sourcing the potential according to eqs. (2.140), (2.142) and (2.248).

As we saw in section 2.7.4, the potential is frozen during MDE at all scales, which helps to solve eq. (3.135), but it still has a k-dependence. We are then tented to resolve the problem for small and large scales, which would leave us with two solutions that need to be patched for intermediate scales. Besides, the solutions in the sub-horizon limit need to be tracked from RDE, where they decrease as τ^{-2} (see eq. (2.136)). Also, its initial conditions were set by Inflation according to sections 3.1 and 3.3. Therefore, to solve those problems, it is common to create a function that matches both, the entire solution for the potential and that for large scales, which is called transfer function and it is defined as

$$T(k) \equiv \frac{\phi\left(k, a_T\right)}{\phi_{\text{Large-Scale}}\left(k, a_T\right)}\,, \tag{3.136}$$

where a_T is some arbitrary value for the scale factor during MDE where the potential is constant, but noting that it can not be a_{eq} from eq. (3.109) or a near value since matter does not govern entirely yet. Typically it is chosen to be $a_T \sim 0.03$. We also need to remember that the potential decreases in a factor 9/10 from RDE to MDE at large scales (eq. (3.111)), which helps us to connect the primordial potential set by inflation ϕ_P with $\phi_{\text{Large-Scale}}$ through

$$\phi_{\text{Large-Scale}}\left(k, a_T\right) = \frac{9}{10}\phi_P\left(k, a_i\right) \qquad \text{(At Large Scales)}\,. \tag{3.137}$$

Finally, we only need to connect the potential with that we observe nowadays, which we are going to do in the next section.

Now, we need to solve the baryon problem, which means that we need to consider them in the equations because they leave a fingerprint too in the matter distribution. We can start taking the adiabatic condition $\delta_b = 3\Theta_0$ from eq. (3.88) and insert it into eq. (3.123), which gives

$$\delta_b'' + \frac{R_s}{1+R_s}\mathcal{H}\delta_b' + k^2 c_s^2 = -k^2\psi - \frac{3R_s}{1+R_s}\mathcal{H}\phi' - 3\phi'' \,. \tag{3.138}$$

We also need to take account of baryons in the potential, which can be done rewriting eq. (3.135) as

$$k^2\phi + 3\mathcal{H}\left(\phi' - \mathcal{H}\psi\right) = 4\pi G a^2\left(\bar{\rho}_m\delta_m + 4\bar{\rho}_r\Theta_{r,0}\right)\,, \tag{3.139}$$

where now the subscript m refers to both baryons and cold dark matter. Note that we recovered the difference between the potentials ϕ and ψ, which implies that we need another equation. We could use eq. (2.253), but then we would need an equation for the dipole; instead, we are going to use eq. (2.103) with $\theta = ikv$, where v is the modulus of the velocity field to avoid confusions. Finally, to close the set of equations, we can add the eqs. (3.131) to (3.134) but recovering the potential ψ.

Like the fingerprint left in the CMB, the acoustic waves also left imprinted the baryon distribution, which signature is characterized by the sound horizon at the drag epoch

$$r_s\left(z_{\mathrm{d}rag}\right) = \int_0^{\tau_{\mathrm{d}rag}} c_s(\tau)\mathrm{d}\tau\,, \tag{3.140}$$

that is the comoving distance traveled by the sound waves by time $\tau_{\mathrm{d}rag}$. It is computed in the same way that eq. (3.127), but now at $z_{\mathrm{d}rag}$. Since z_{dec} differs from $z_{\mathrm{d}rag}$, then the sound horizons also are different.

3.5 Transition from MDE to DEDE

It is difficult to determine the precise beginning of the dark energy era for two reasons. First, we know that nowadays, our Universe has an accelerated expansion, which only can be possible given a source with an EoS $w < -1/3$. Then, many options surge to try to explain that, but still, no method to confirm them directly. Nonetheless, there is a favorite, in which the acceleration is driven by a cosmological constant in the Einstein equations, this is known as ΛCDM[9]. Others suggest that the acceleration is, in fact, a sign that the Einstein gravity is not accurate at certain scales, and we would need another theory to explain gravity at those. The second reason is that whatever causes the acceleration, it has an energy density whose value is in the same order that the whole matter nowadays. Meaning that DEDE is not totally ruled by dark energy, and instead, it shares the domination with the matter. Therefore, try to determinate an exact time when dark energy surpasses the matter density is infeasible. Nonetheless and considering the ΛCDM model, we can find when both energy densities are the same, which gives $a_{\mathrm{eq}} \sim 0.75$ at $z_{\mathrm{eq}} \sim 0.3$.

---

[9] The ΛCDM is the favorite model in the modern Cosmology, since the actual data favor it and then, it is most likely. However, it has some unresolved problems, like the fine-tuning problem and the *why now* problem.

After the drag epoch, the baryons start to fall in the cold dark matter potential wells, but also, the dark matter is attracted to the imprint left by baryon acoustic oscillations. Then, before MDE ends, both matter distributions approach each other. During this time, the evolution of matter perturbations is government mostly on small scales through eq. (2.131) with $w = c_s = 0$

$$\delta'' + \mathcal{H}\delta' - \frac{3}{2}\mathcal{H}^2\delta = 0\,, \tag{3.141}$$

where we will make a change of variables to N from eq. (2.124) and following eqs. (2.117), (2.125) and (2.126) to obtain

$$\frac{\mathrm{d}^2\delta}{\mathrm{d}N^2} + \frac{1}{2}\frac{\mathrm{d}\delta}{\mathrm{d}N} - \frac{3}{2}\delta = 0\,, \tag{3.142}$$

which solution is

$$\begin{aligned}\delta &= c_1\mathrm{e}^{-3N/2} + c_2\mathrm{e}^{N}\\ &= c_1 a^{-3/2} + c_2 a\,,\end{aligned} \tag{3.143}$$

where c_1, c_2 are integration constants. We can also recognize the growing mode $\delta_+ = c_1 a$, which motives us to define the growth function as

$$D(a) \equiv \frac{\delta_+(a)}{\delta_+(a_0)}\,, \tag{3.144}$$

with $a_0 = 1$ because eq. (1.10). Then $D(a) = a$ during MDE by eq. (3.143).

Once entered in DEDE, the potentials are still sourced only by matter at least in ΛCDM, where by definition, the cosmological constant has no perturbations, but since the background quantities change, it also makes that the potentials evolve differently. It is common to define a function that patches the potential now with that during MDE, in the same way that we did for the transfer function in eq. (3.136). This is going to be the same growth function $D(a)$ of eq. (3.144) for reasons that we will explain in section 4.2

$$\frac{D(a)}{a} = \frac{\phi(a)}{\phi\left(a_T\right)} \qquad (a > a_T)\ . \tag{3.145}$$

Note that we normalized the function by a, since in ΛCDM, the growth is linear $D(a) = a$ during MDE as we saw, which keeps the potential unaltered. Thus, we can connect the primordial potential from inflation with that we see nowadays using eqs. (3.136), (3.137) and (3.145)

$$\phi\left(k, a\right) = \frac{9}{10}\phi_P\left(k\right)T\left(k\right)\frac{D(a)}{a} \qquad (a > a_T)\ . \tag{3.146}$$

Chapter 4

From Theory To Observations

All the theory presented in chapters 1 to 3 is structured to connect it with the observations of our Universe. The reason is that historically, the latter were who droves the theory, but now we are focusing on a deductive method. For that, we first need to prepare the observables from the equations. Nonetheless, in a complex Universe like ours, that work is not easy, and we will request a lot of mathematical techniques like perturbation theory, non-linear evolution, statistics tools, and others.

4.1 Random Fields

We are going to study the statistics of random fields and its notation following mostly [200]. Therefore, a random field is a function $G\left(\vec{x}\right)$ whose values are randoms variables for any point $\vec{x}$, which does not belong exclusively to a $3D$ Euclidean space. The associated distribution function is given by

$$F_{1,2,...,n}\left(g_1,g_2,...,g_n\right)=F\left[G\left(\vec{x}_1\right)<g_1,G\left(\vec{x}_2\right)<g_2,...,G\left(\vec{x}_n\right)<g_n\right]\,, \tag{4.1}$$

where the G is the random field, which has value g among all the possible ones from the ensemble at a certain point $\vec{x}$. Then, the g_i refers to the different points $\vec{x}_i$ where G is evaluated. The probability that G takes some value g_1 at some point $\vec{x}_1$ is

$$p_1\left(g_1\right)\mathrm{d}g_1\,, \tag{4.2}$$

thus, p_1 is the probability distribution function

$$p_1\left(g_1\right)=\frac{\mathrm{d}F_1\left(g_1\right)}{\mathrm{d}g_1}\,, \tag{4.3}$$

which turns F_1 into the cumulative probability function with $F_1\left(-\infty\right)=0$ and $F_1\left(\infty\right)=1$.

The expectation value for a random field is defined through an ensemble average, which at some point $\vec{x}_1$ is

$$\left\langle G\left(\vec{x}_1\right)\right\rangle\equiv\int_\Omega g_1\,p_1\left(g_1\right)\mathrm{d}g_1\,, \tag{4.4}$$

where Ω denotes the ensemble.

In general, for two points $\vec{x}_1$ and $\vec{x}_2$, we have that $p_1(g_1) \neq p_2(g_2)$, which means that the probability distribution functions of the values that G may assume in $\vec{x}_1$ and $\vec{x}_2$ are different. When this does not happen, the probability of the realization is translationally invariant and we say that G is a statistically homogeneous random field, whose ensemble average (eq. (4.4)) is just

$$\langle G \rangle = \int_\Omega g\, p(g)\, \mathrm{d}g\,. \tag{4.5}$$

We also can ask for the probability that G takes the values g_1 and g_2 at $\vec{x}_1$ and $\vec{x}_2$ respectively, which is given by

$$p_{12}(g_1, g_2)\, \mathrm{d}g_1 \mathrm{d}g_2\,, \tag{4.6}$$

and can also be written as a derivative of the cumulative probability function $F_{1,2}$. Usually we have that $p_{12} \neq p_1(g_1)\, p_2(g_2)$, but for independent realizations it obeys $p_{12} = p_1(g_1)\, p_2(g_2)$, in which case the random process is called Poissonian. The two-dimensional probability distribution used in eq. (4.6) allow us to define the two-point correlation function as

$$\xi(\vec{x}_1, \vec{x}_2) \equiv \langle G(\vec{x}_1)\, G(\vec{x}_2) \rangle \equiv \int_\Omega g_1 g_2\, p_{12}(g_1, g_2)\, \mathrm{d}g_1 \mathrm{d}g_2\,, \tag{4.7}$$

which indicates the auto-correlation[1] of the random field G at two points. In the same way, we can extend the probability distribution for N points, where we can define the N-point correlation function as

$$\begin{aligned} \xi^{(N)}(\vec{x}_1, \vec{x}_2, ..., \vec{x}_N) &\equiv \langle G(\vec{x}_1)\, G(\vec{x}_2) ... G(\vec{x}_N) \rangle \\ &\equiv \int_\Omega g_1 g_2 ... g_N\, p_{12...N}(g_1, g_2, ..., g_N)\, \mathrm{d}^N g\,, \end{aligned} \tag{4.8}$$

where we used the same ensemble Ω for each point. We must take care of the order in the probability distribution function, because in general we have that $p_{12...N}(g_1, g_2, ..., g_N) \neq p_{21...N}(g_1, g_2, ..., g_N)$.

A random field is statistically isotropic if the probability of the realization is rotationally invariant, which implies that for any rotation matrix $\mathbf{R}$, the point $\vec{x}_{\mathbf{R}_1} = \mathbf{R}\vec{x}_1$ with probability distribution $P_{\mathbf{R}_1}$ satisfies

$$p_1(g_1) = P_{\mathbf{R}_1}(g_{\mathbf{R}_1})\ . \tag{4.9}$$

If the random field is statistically homogeneous, the two-point correlation function depends only in the difference between points

$$\xi(\vec{x}_1, \vec{x}_2) = \xi(\vec{x}_1 - \vec{x}_2)\ . \tag{4.10}$$

Besides, if the random field is both statistically homogeneous and statistically isotropic, the two-point correlation function depends only in the distance between points

$$\xi(\vec{x}_1, \vec{x}_2) = \xi(\vec{x}_1 - \vec{x}_2) = \xi(r_{1\,2})\ , \tag{4.11}$$

where $r_{1\,2} = |\vec{x}_1 - \vec{x}_2|$.

---

[1] For two different random fields we talk about cross-correlation.

Using the standard definition of variance from statistics $\sigma^2(X) = \langle X^2\rangle - \langle X\rangle^2$, we can define the ensemble variance of the random field for two points as

$$\begin{aligned}\sigma^2(\vec{x}_1,\vec{x}_2) &\equiv \langle G(\vec{x}_1)\,G(\vec{x}_2)\rangle - \langle G(\vec{x}_1)\rangle\langle G(\vec{x}_2)\rangle \\ &= \xi(\vec{x}_1,\vec{x}_2) - \langle G(\vec{x}_1)\rangle\langle G(\vec{x}_2)\rangle\,,\end{aligned} \tag{4.12}$$

which for a statistically homogeneous and isotropic random field, it is just

$$\sigma^2(r_{1\,2}) = \xi(r_{1\,2}) - \langle G\rangle^2\,, \tag{4.13}$$

remarking again that $\langle G\rangle^2$ is independent for the position (see eq. (4.5)). We also note that for a Poissonian random process we have that $\langle G(\vec{x}_1)\,G(\vec{x}_2)\rangle = \langle G(\vec{x}_1)\rangle\langle G(\vec{x}_2)\rangle$, then the variance is zero and

$$\xi(r_{1\,2}) = \langle G\rangle^2\,. \tag{4.14}$$

In some cases, we can compute just a spatial average of G since we only have one realization[2]. Therefore, we take account of the spatial average for some volume V of the random field G, which is

$$\overline{G} \equiv \frac{1}{V}\int_V \mathrm{d}^m\, x G(\vec{x})\ , \tag{4.15}$$

where m is the dimension of the space where $\vec{x}$ lives. To estimate the error of exchanging the ensemble average with the spatial one, we are going to define

$$X \equiv \overline{G} - \langle G\rangle\,, \tag{4.16}$$

which has a null expectation value

$$\langle X\rangle = \frac{1}{V}\int_V \mathrm{d}^m\, x\langle G(\vec{x})\rangle - \langle G\rangle = 0\,, \tag{4.17}$$

and a variance given by

$$\langle X^2\rangle = \frac{1}{V^2}\int_V \mathrm{d}^m x_1 \int_V \mathrm{d}^m\, x_2\langle G(\vec{x}_1)\,G(\vec{x}_2)\rangle - \langle G\rangle^2\,, \tag{4.18}$$

which for a statistically homogeneous random field can be written as

$$\langle X^2\rangle = \frac{1}{V^2}\int_V \mathrm{d}^m x_1 \int_V \mathrm{d}^m x_2\, \xi(\vec{x}_1 - \vec{x}_2) - \langle G\rangle^2\,, \tag{4.19}$$

where we know that it vanishes for a Poissonian process from eqs. (4.12) and (4.14). For a $3D$ Euclidean space and after a variable change, the integral from eq. (4.19) reduces to

$$\langle X^2\rangle = \frac{1}{V}\int_V \mathrm{d}^3 r\, \xi(\vec{r}) - \langle G\rangle^2\,, \tag{4.20}$$

which goes to zero in the limit $V \to \infty$ since the ergodic theorem. However, in most realistic cases, we have a finite volume, and then, a non zero variance[3]. Also assuming a statistically

---

[2] This is the case of observational Cosmology because we only have one Universe, ours.

[3] The variance $\langle X^2\rangle$ in Cosmology is known as cosmic variance.

isotropic random field and using spherical coordinates to describe the volume, we can write eq. (4.20) as

$$\langle X^2\rangle = \frac{3}{R^3}\int_0^R \mathrm{d}r\, r^2\xi(r) - \langle G\rangle^2\,. \tag{4.21}$$

It is also useful to study the random fields in Fourier space since, in almost all the cases, we can work in it as we did in the last chapters. We will work in a $3D$ Euclidean space from now, and also, we will assume that the Fourier transform of a random field is also a random field. Thus, for G the conventions to go and back from Fourier space are the same that before (see chapters 2 and 3)

$$G\left(\vec{x}\right) = \int \frac{\mathrm{d}^3k}{(2\pi)^3}\,\tilde{G}\big(\vec{k}\big)\mathrm{e}^{i\vec{k}\cdot\vec{x}}\,, \tag{4.22}$$

$$\tilde{G}\big(\vec{k}\big) = \int \mathrm{d}^3x\, G\left(\vec{x}\right)\mathrm{e}^{-i\vec{k}\cdot\vec{x}}\,, \tag{4.23}$$

which, for a finite volume as a box of volume L^3, take the form of Fourier series

$$G\left(\vec{x}\right) = \frac{1}{L^3}\sum_n G_n \mathrm{e}^{i\vec{k}_n\cdot\vec{x}}\,, \tag{4.24}$$

where the coefficients G_n are defined through

$$G_n = \int \mathrm{d}^3x\, G\left(\vec{x}\right)\mathrm{e}^{-i\vec{k}_n\cdot\vec{x}}\,, \tag{4.25}$$

and the wavenumbers $\vec{k}_n$ are quantized, since the periodic boundary conditions, as

$$\vec{k}_n = \frac{2\pi}{L}\vec{n}\,, \tag{4.26}$$

with $\vec{n}$ a generic vector whose components are integers. Note that if $L\to\infty$, we recover the usual convention from eq. (4.22).

If $G\left(\vec{x}\right)$ is a real field, then $\tilde{G}\big(-\vec{k}\,\big) = \tilde{G}^*\big(\vec{k}\,\big)$ with $\tilde{G}^*\big(\vec{k}\,\big)$ the complex conjugate of $\tilde{G}\big(\vec{k}\,\big)$, which is known as the reality condition. Now, we can compute the two-point correlation function for the Fourier transform of the random field $\tilde{G}$, using the reality condition, as

$$\big\langle \tilde{G}\big(\vec{k}\,\big)\tilde{G}^*\big(\vec{k}\,'\big)\big\rangle = \int \mathrm{d}^3x\int \mathrm{d}^3x'\,\langle G\left(\vec{x}\right)G\left(\vec{x}\,'\right)\rangle\,\mathrm{e}^{-i\vec{x}\cdot\vec{k}}\,\mathrm{e}^{i\vec{x}\,'\cdot\vec{k}\,'}\,, \tag{4.27}$$

which, after assuming homogeneity, becomes

$$\big\langle \tilde{G}\big(\vec{k}\,\big)\tilde{G}^*\big(\vec{k}\,'\big)\big\rangle = \int \mathrm{d}^3x\int \mathrm{d}^3x'\,\xi_G\left(\vec{x}\,'-\vec{x}\right)\mathrm{e}^{-i\vec{x}\cdot\vec{k}}\,\mathrm{e}^{i\vec{x}\,'\cdot\vec{k}\,'}\,, \tag{4.28}$$

where ξ_G is just the two-point correlation function of eq. (4.7) with the subscript G to specify the random field. Using the change of variables $\vec{z} = \vec{x}\,' - \vec{x}$, eq. (4.28) becomes

$$\begin{aligned}\big\langle \tilde{G}\big(\vec{k}\,\big)\tilde{G}^*\big(\vec{k}\,'\big)\big\rangle &= \int \mathrm{d}^3x\int \mathrm{e}^{-i\left(\vec{k}-\vec{k}\,'\right)\cdot\vec{x}}\,\mathrm{d}^3z\,\xi_G\left(\vec{z}\right)\mathrm{e}^{i\vec{k}\,'\cdot\vec{z}}\\ &= (2\pi)^3\,\delta\big(\vec{k}-\vec{k}\,'\big)P_G\big(\vec{k}\,\big)\,,\end{aligned} \tag{4.29}$$

where we used the integral representation of the Dirac delta, and also we defined the power spectrum as

$$P_G(\vec{k}) \equiv \int \mathrm{d}^3x\, \xi_G\left(\vec{x}\right) \mathrm{e}^{-i\vec{k}\cdot\vec{x}}\,, \tag{4.30}$$

which is the Fourier transform of the two-point correlation function. Assuming a statistical isotropic random field and using spherical coordinates, the eq. (4.30) can be developed as

$$\begin{aligned} P_G\left(k\right) &= 2\pi \int_0^\infty \mathrm{d}r\, r^2 \xi_G\left(r\right) \int_{-1}^1 \mathrm{d}u\, \mathrm{e}^{-ikru} \\ &= 4\pi \int_0^\infty \mathrm{d}r\, r^2 \xi_G\left(r\right) \frac{\sin\left(kr\right)}{kr}\,, \end{aligned} \tag{4.31}$$

where $u = \vec{k}\cdot\vec{x}$ is the cosine of the angle between those two vectors. We also can use $j_0(kr) = \sin\left(kr\right)/\left(kr\right)$ in eq. (4.31), where j_0 is the spherical Bessel function of order zero.

In most cases, observations are going to provide a direct measurement of the two-point correlation function. Then, we are going to invert the power spectrum, for that we will use eqs. (4.7) and (4.22) and the reality condition for a statistical homogeneous and isotropic random field

$$\begin{aligned} \xi_G\left(r\right) &= \left\langle G\left(\vec{x}\right) G\left(\vec{x}+\vec{r}\right)\right\rangle \\ &= \int \frac{\mathrm{d}^3k}{(2\pi)^3} \int \frac{\mathrm{d}^3k'}{(2\pi)^3} \left\langle \tilde{G}(\vec{k})\tilde{G}^*(\vec{k}\,')\mathrm{e}^{i\vec{k}\cdot\vec{x}-i\vec{k}\,'\cdot(\vec{x}+\vec{r})}\right\rangle. \end{aligned} \tag{4.32}$$

Note that we included the modes in the average to seek the differences between doing spatial or ensemble average. For the latter, the average is only carried on the fields, then eq. (4.32) becomes

$$\begin{aligned} \xi_G\left(r\right) &= \int \frac{\mathrm{d}^3k}{(2\pi)^3} \int \mathrm{d}^3k'\, P_G\left(k\right) \delta(\vec{k}-\vec{k}\,')\mathrm{e}^{i\vec{k}\cdot\vec{x}-i\vec{k}\,'\cdot(\vec{x}+\vec{r})} \\ &= \int \frac{\mathrm{d}^3k}{(2\pi)^3}\, P_G\left(k\right) \mathrm{e}^{-i\vec{k}\cdot\vec{r}}\,, \end{aligned} \tag{4.33}$$

where in the first line we used eq. (4.29), and in the second line we just used the property of the Dirac delta within an integral. Using spherical coordinates like in eq. (4.31), we can develop the angular integral of eq. (4.33)

$$\begin{aligned} \xi_G\left(r\right) &= \int \frac{\mathrm{d}k}{(2\pi)^2}\, k^2 P_G(k) \int_{-1}^1 \mathrm{d}u\, \mathrm{e}^{-ikru} \\ &= \int_0^\infty \frac{\mathrm{d}k}{2\pi^2}\, k^2 P_G(k) \frac{\sin\left(kr\right)}{kr} \\ &= \int_0^\infty \frac{\mathrm{d}k}{k}\, \Delta_G^2(k) \frac{\sin\left(kr\right)}{kr}\,, \end{aligned} \tag{4.34}$$

where in the last line, we defined the dimensionless power spectrum as

$$\Delta_G^2(k) \equiv \frac{k^3 P_G(k)}{2\pi^2}\,, \tag{4.35}$$

by customary, because the two-point correlation function ξ_G is dimensionless, which implies that the power spectrum P_G has dimension of volume.

Now, for the spatial average of eq. (4.32) we are going to use the Fourier series of eq. (4.24) and the spatial average definition of eq. (4.15). We also are going to write the spatial average using $\hat{\xi}_G$ instead of a bar

$$\hat{\xi}_G(r) = \frac{1}{V}\int_V \mathrm{d}^3x \sum_n \sum_m \frac{1}{V^2} G_n G_m^* \mathrm{e}^{i\vec{k}_n\cdot\vec{x} - i\vec{k}_m\cdot(\vec{x}+\vec{r})}\,, \tag{4.36}$$

where $G_n = \tilde{G}(\vec{k}_n)$ by comparing eqs. (4.23) and (4.25), and the $\vec{k}_n$ are defined in eq. (4.26). Note that the sums in eq. (4.36) are understood as sums over the three components of $\vec{n}$ and $\vec{m}$, then switching the integral with those sums gives the spatial integration $\int_V \mathrm{d}^3x \mathrm{e}^{i(\vec{k}_n - \vec{k}_m)} = \delta_{nm}V$, and the correlation function reduces to

$$\hat{\xi}_G(r) = \sum_n \frac{1}{V^2}|G_n|^2 \mathrm{e}^{-i\vec{k}_n\cdot\vec{r}}\,, \tag{4.37}$$

from which we infer the power spectrum

$$P_n \equiv P(\vec{k}_n) = \frac{|G_n|^2}{V}\,. \tag{4.38}$$

The signal to noise ratio for a Poissonian process is $1/\sqrt{N}$, where N is the number of independent realizations. Then, we can relate the cosmic variance with the power spectrum for a Poissonian process as

$$\frac{\sigma_P(k)}{P(k)} \simeq \frac{1}{\sqrt{N_k}}\,, \tag{4.39}$$

where N_k is the number of independent modes, which is $N_k = 4\pi k^2 \left(L/(2\pi)\right)^3 \mathrm{d}k$ between k and $k + \mathrm{d}k$, and $N_k = 4\pi k^2 \left(L/(2\pi)\right)^2$ inside a sphere with radius k. Therefore, the signal to noise ration in the first case is

$$\frac{\sigma_P(k)}{P(k)} \simeq \frac{1}{\sqrt{r_k}\,(kL)^{3/2}}\,, \tag{4.40}$$

with $r_k \equiv \mathrm{d}k/k$. While, for the second case

$$\frac{\sigma_P(k)}{P(k)} \simeq \frac{1}{kL}\,, \tag{4.41}$$

which tells us that the cosmic variance is negligible if the scale is much smaller than the dimension L of the survey, $kL \gg 1$.

4.1.1 Gaussian Random Fields

We can define a Gaussian random field in several ways. If we use the Fourier transform, a Gaussian random field has modes uncorrelated like in the second line of eq. (4.29), where the Dirac delta appears. So, Gaussianity implies statistical homogeneity. Furthermore, a Gaussian random field is characterized by the fact that all the correlators of odd order vanish, for example

$$\langle \tilde{G}(\vec{k}\,) \rangle = \langle \tilde{G}(\vec{k}_1)\tilde{G}(\vec{k}_2)\tilde{G}(\vec{k}_3) \rangle = 0\,. \tag{4.42}$$

Therefore, all the even order correlators can be written as a function of two-point correlators as the power spectrum. For example, for the 4-point correlator

$$\begin{aligned}\langle\tilde{G}(\vec{k}_1)\tilde{G}(\vec{k}_2)\tilde{G}(\vec{k}_3)\tilde{G}(\vec{k}_4)\rangle &= \langle\tilde{G}(\vec{k}_1)\tilde{G}(\vec{k}_2)\rangle\langle\tilde{G}(\vec{k}_3)\tilde{G}(\vec{k}_4)\rangle \\ &+ \langle\tilde{G}(\vec{k}_1)\tilde{G}(\vec{k}_3)\rangle\langle\tilde{G}(\vec{k}_2)\tilde{G}(\vec{k}_4)\rangle \\ &+ \langle\tilde{G}(\vec{k}_1)\tilde{G}(\vec{k}_4)\rangle\langle\tilde{G}(\vec{k}_2)\tilde{G}(\vec{k}_3)\rangle\,,\end{aligned} \tag{4.43}$$

Since all the Fourier modes are uncorrelated, their superposition is Gaussian distributed by the central limit theorem, which is why they are known as Gaussian random fields. Thus, the g values are distributed as a Gaussian, *i.e.*

$$p\,(g) = \frac{1}{\sqrt{2\pi\sigma_g^2}}\exp\left(-\frac{g^2}{2\sigma_g^2}\right)\,, \tag{4.44}$$

where we assumed a vanishing expectation value and a variance given by eq. (4.12). Note that in eq. (4.44) we used just g, because G is a statistical homogeneous random field.

4.2 Matter Power Spectrum

In section 3.1, we already computed the power spectrum for the comoving curvature perturbations, but at that time, we did not know the statistic meaning yet. We can now connect it with the potential and matter power spectra as in [80], using the equations found in chapters 2 and 3. We can recognize that $\mathcal{R}$ is a statistically homogeneous and isotropic random field since its power spectrum in eq. (3.77) only depends on the wavenumber modulus. Besides, if $\mathcal{R}$ is Gaussian, then its power spectrum contains all the statistical information, but most of the inflationary scenes predict primordial non-Gaussianity, which is encoded in higher-order correlation functions. We also want to know how are the statistics for other cosmological randoms fields, like the potential perturbations $\phi(k)$ and the density contrast $\delta(k)$. For that, we are going to start relating the potential perturbations after inflation with the comoving curvature using eq. (3.110) wor $w = 1/3$ since we are in RDE, this gives

$$\Phi = \frac{2}{3}\mathcal{R}\,, \tag{4.45}$$

which also remains in Fourier space. Then, the relation between the primordial power spectrum for Φ and the comoving curvature perturbation is

$$\begin{aligned}P_\Phi(k) &= \frac{4}{9}P_{\mathcal{R}}(k) \\ &= \frac{8\pi^2}{9k^3}A_s\left(\frac{k}{H_0}\right)^{n_s-1}\,,\end{aligned} \tag{4.46}$$

where we used eq. (3.77) in the second line. We can rewrite A_s in eq. (4.46) using the amplitude of density perturbations at horizon crossing δ_H, which obeys $\delta_H^2 = (4/25)\,A_s$ according to [153], thus the power spectrum for the potential after inflation is

$$P_\Phi(k) = \frac{50\pi^2}{9k^3}\delta_H^2\left(\frac{k}{H_0}\right)^{n_s-1}\,. \tag{4.47}$$

Nonetheless, we know that the potential evolves and it is affected by early Universe processes, summarized by eq. (3.146). Then, the final power spectrum for the potential *i.e.* the one that we can measure at late Universe is

$$\begin{aligned} P_\Phi(k,a) &\equiv |\phi|^2 \\ &= \frac{81}{100} T^2(k) \frac{D^2(a)}{a^2} P_{\Phi_P}(k) \\ &= \frac{9\pi^2\delta_H^2}{2a^2k^3} T^2(k) D^2(a) \left(\frac{k}{H_0}\right)^{n_s-1} . \end{aligned} \tag{4.48}$$

For modes that cross the horizon at late times $k \gg \mathcal{H}$, which is equivalent to the sub-horizon limit, we can directly relate the density contrast with the potential at late times using eq. (2.128). Where $\bar{\rho}\delta = \bar{\rho}_m\delta_m$ for MDE and $DEDE$ since the radiation perturbations are negligible and dark energy, if it is Λ, has not perturbations by definition[4]. Therefore, we have that

$$\begin{aligned} \delta_m &= \frac{k^2}{4\pi G\bar{\rho}_m a^2}\phi \\ &= \frac{2k^2 a}{3H_0^2\Omega_{m,0}}\phi , \end{aligned} \tag{4.49}$$

where we used eq. (1.68) to compute $\bar{\rho}_m = \rho_{m,0}a^{-3}$ and eq. (1.35) to transform $\rho_{m,0}$ into $\Omega_{m,0}$. Note that ϕ evolves as $\delta(k,a)/a$, which is the true reason why we defined the growth function of eq. (3.145) in that way. Then, the matter power spectrum at late times can be computed using the potential power spectrum in eq. (4.48), as

$$\begin{aligned} P_m(k,a) &\equiv |\delta_m|^2 \\ &= \frac{4a^2k^4}{9\Omega_{m,0}^2H_0^4} P_\Phi(k,a) \\ &= \frac{2\pi^2\delta_H^2}{\Omega_{m,0}^2H_0^3} T^2(k) D^2(a) \left(\frac{k}{H_0}\right)^{n_s} . \end{aligned} \tag{4.50}$$

Some authors [80] redefine δ_H^2 to include the $\Omega_{m,0}^2$ factor in the denominator of eq. (4.50), which makes disappear the normalized matter energy density parameter in the matter power spectrum.

Due to all the constants carried in the final spectrum, it is common to define some normalization for $P_m(k)$. For that, we will use the variance of the mass density field smoothed on some comoving scale R corresponding to a mass $M = (4\pi/3)\,\bar{\rho}R^3$, where $\bar{\rho}$ is the background mass density. This variance can be computed using the power spectrum through

$$\begin{aligned} \sigma_R^2(a) &= \int \frac{\mathrm{d}^3k}{(2\pi)^3} P_m(k,a) W(kR) \\ &= D^2(a) \int \frac{\mathrm{d}^3k}{(2\pi)^3} W(kR) \frac{2\pi^2\delta_H^2}{\Omega_{m,0}^2H_0^3} T^2(k) \left(\frac{k}{H_0}\right)^{n_s} , \end{aligned} \tag{4.51}$$

---

[4] Indeed, for other types of dark energy, the perturbations are tiny and negligible compared to those of matter.

where we used eq. (4.50), and $W(kR)$ is the window function corresponding to the smoothing of the density field. In most of the cases, the mass functions are fitted assuming a spherical top-hat smoothing, which corresponds to

$$W(x) = 3\frac{j_1(x)}{x}\,, \tag{4.52}$$

where $j_1(x)$ denotes the spherical Bessel function of order 1. Note that in the second line of eq. (4.51), we put out the growth function from the integral. Thus, since $D(a=1)=1$, we can write the variance of the mass density at redshift $z=0$ or $a=1$ as

$$\sigma^2_{R,0} \equiv \sigma^2_R(a=1) = \int \frac{\mathrm{d}^3k}{(2\pi)^3}\, W(kR)\, \frac{2\pi^2\delta_H^2}{\Omega^2_{m,0}H_0^3} T^2(k) \left(\frac{k}{H_0}\right)^{n_s}\,, \tag{4.53}$$

and then we can write the variance of the mass density at any time just as

$$\sigma_R(a) = D(a)\sigma_{R,0}\,, \tag{4.54}$$

where we dropped out the squares. Since the original idea is to normalize the matter power spectrum, we must agree with some scale R, which by convention is defined at $R = 8h^{-1}Mpc$ because at that scale, the linear perturbation theory breaks. Finally, the normalized matter power spectrum can be written as

$$P_m(k) = \frac{P_m(k,a)}{\sigma^2_8(a)}\,, \tag{4.55}$$

where $P_m(k,a)$ comes from eq. (4.50). Note that even the both functions in the fraction of eq. (4.55) depend on time, the quotient does not.

4.3 Correlation Function Estimators

In the section 4.1, we saw that the cosmic variance could not be avoided since we have no access to different ensembles to make the correct average. Then, we only can reduce it, considering a large volume to get a better approximation between volume and ensemble averages. Nevertheless, we also have no random field; instead we have a collection of points corresponding to galaxies with their coordinates. Therefore, we must reconstruct the random field from the galaxies, but this introduces a new error: the cosmic bias. These galaxies can be modeled as spatial point processes, where we will need to use all the framework of statistical analysis to found the so-called estimators, for which we will follow the reviews in [203, 139]. First, we are going to introduce a first-order characteristic, such as the averaged number density $\bar{n}$ or also called intensity, which is

$$\bar{n} \equiv \frac{N}{V}\,, \tag{4.56}$$

where N is the total number of objects inside the volume V. Then, we are going to introduce second-order characteristics, such as the correlation function $\xi(r)$ and the pair correlation function $g(r)$, which satisfy

$$g(r) = 1 + \xi(r)\,, \tag{4.57}$$

where $\xi(r)$ is defined in eq. (4.7) and for a statistically homogeneous and isotropic field in eq. (4.11). A more practical definition for these measures is considering a infinitesimal ball B of volume $\mathrm{d}V$ where the probability of having a point inside B is $\bar{n}\mathrm{d}V$. Now, if there are two balls B_1 and B_2, of volumes $\mathrm{d}V_1$ and $\mathrm{d}V_2$, and intercenter distance r, then the probability to have a point in each ball can be expressed as

$$\mathrm{d}P = \bar{n}^2 g(r)\mathrm{d}V_1\mathrm{d}V_2 = \bar{n}^2\left[1+\xi(r)\right]\mathrm{d}V_1\mathrm{d}V_2\,. \tag{4.58}$$

It is clear that in the case of complete randomness of the point distribution[5], the factor of proportionality $g(r) = 1 + \xi(r) = 1$. Thus, the correlation function $\xi(r)$ is the excess of probability, over a random point distribution, to find two galaxies separated by a distance r.

Now, we need a way to measure those characteristics considering that we only have a catalog with galaxies positions. We will introduce some popular estimators in the literature, which are not perfect because they have a certain grade of error due to cosmic bias or cosmic variance. We are also going to see that, in most cases, it is better to have estimators with low cosmic variance than cosmic bias. We start considering a catalog with the points $\{\vec{x}_i\}_{i=1}^{N_g}$, which, for the moment, are statistically homogeneous distributed. Also, N_g is the number of galaxies, which are over a volume of V_c. Therefore, an estimator for the averaged number of galaxies $\hat{\bar{n}}$ is just

$$\hat{\bar{n}} = \hat{n}_g = \frac{N_g}{V_c}\,, \tag{4.59}$$

which is an unbiased estimator of the number density. For the correlation function, we can find a biased but straightforward estimator

$$\hat{g}_0(r) = 1 + \hat{\xi}_0(r) = \frac{1}{N_g}\sum_{i=1}^{N_g}\frac{n_i^{\Delta}(r)}{4\pi r^2\Delta\hat{n}_g}\,, \tag{4.60}$$

where Δ is known as the bin width, which is finite since we are working with a finite number of points, we can only have access to the correlation information in the interval $[r, r+\Delta]$ and not to a certain r. To define n_i^{Δ}, we first need to introduce the number of points in a sphere with radius r around a point $\vec{x}_i$

$$N_i(r) = \sum_{j=1;j\neq i}^{N}\mathbb{1}_{[0,r]}\left(|\vec{x}_i - \vec{x}_j|\right)\,, \tag{4.61}$$

where N is the total number of points $i.e$ $N = N_g$, and $\mathbb{1}_A(x)$ is defined as

$$\mathbb{1}_A = \begin{cases} 1 & x\in A \\ 0 & x\notin A \end{cases}\,, \tag{4.62}$$

which denotes the indicator function of the set A. Thus, n_i^{Δ} is the number of points in the shell with a radius between $[r, r+\Delta]$ around a point $\vec{x}_i$

$$\begin{aligned} n_i^{\Delta}(r) &= \sum_{j=1;j\neq i}^{N}\mathbb{1}_{[r,r+\Delta]}\left(|\vec{x}_i - \vec{x}_j|\right) \\ &= N_i\left(r+\Delta\right) - N_i\left(r\right)\,. \end{aligned} \tag{4.63}$$

---

[5] Complete randomness refers to that the measures in B_1 and B_2 are uncorrelated, which is the case for points that are Poisson distributed.

The estimator in eq. (4.60) considers equally all the points that are inside the sample geometry $\mathcal{D}$, but for the points $\vec{x}_i$ near the boundary of $\mathcal{D}$ and for large radii r, the number $N_i(r)$ is underestimated and $\hat{\xi}_0$ is biased towards smaller values. However, there are more estimators defined in the same way as eq. (4.60) that try to fix the problem of boundary points, which can be found in [139].

We can have a property similar to unbiasedness if the expectation of an estimator converges towards the real mean value of $\xi(r)$ for $\Delta \to 0$; this is called approximately unbiased. Also, we can have ratio-unbiased estimators, which are defined as the quotient of two unbiased quantities; this is the case for some of the estimates in [139], which are ratio–unbiased in the limit $\Delta \to 0$.

In the cosmological literature, the correlation function estimators are often constructed using together another set of random points $\{\vec{y}_i\}_{i=1}^{N_r}$, which are Poisson distributed, but they are inside the same sample geometry $\mathcal{D}$ that the original catalog with N_g galaxies. This election is because we want to use the basic idea introduced before, that the correlation function is the excess of probability over a random point distribution. For the real catalog, we are going to count the number of *data-data* pairs as

$$DD(r) = \frac{1}{2}\sum_{i=1}^{N_g} n_i^{\Delta}(r)\,, \tag{4.64}$$

where $n_i^{\Delta}(r)$ is defined in eq. (4.63). Note that the *data-data* pairs $DD(r)$ are also settled for a bin $[r, r+\Delta]$ like $\hat{\xi}_0$. Therefore, the sum in $DD(r)$ counts all the distances between galaxies that belong to the above bin, but since pairs are counted twice we introduced the factor 1/2. We can do the same for the random catalog, introducing the *random-random* pairs

$$RR(r) = \frac{1}{2}\sum_{i=1}^{N_r} rr_i^{\Delta}(r)\,, \tag{4.65}$$

with $rr_i^{\Delta}(r)$ is defined equivalently as $n_i^{\Delta}(r)$ in eq. (4.63), but now for the random catalog

$$rr_i^{\Delta}(r) = \sum_{j=1;j\neq i}^{N_r} \mathbb{1}_{[r,r+\Delta]}\left(|\vec{y}_i - \vec{y}_j|\right)\,, \tag{4.66}$$

which also is defined for a bin. The *random-random* pairs $RR(r)$ count the distances between the random points that belong to the bin $[r, r+\Delta]$. Since It also counts twice, we introduced again the factor 1/2 in eq. (4.65). Finally, we can consider the pairs between the data and the random catalogs, for that, we are going to introduce the *data-random* pairs

$$DR(r) = \sum_{i=1}^{N_g} \mathrm{d}r_i^{\Delta}(r)\,, \tag{4.67}$$

with

$$\mathrm{d}r_i^{\Delta}(r) = \sum_{j=1}^{N_r} \mathbb{1}_{[r,r+\Delta]}\left(|\vec{x}_i - \vec{y}_j|\right)\,, \tag{4.68}$$

where both are also defined for a bin. In general, the *data-random* pairs are not equal to the *random-data* pairs, which are defined in a similar way to the first but interchanging the points and sums. However, for statistically homogeneous data and random catalogs, both are the same and we can take advantage of it. Note that this time we did not introduce the 1/2 factor, since the *data-random* pairs are not counted twice, this is because $(|\vec{x}_i - \vec{y}_j|) \neq (|\vec{y}_i - \vec{x}_j|)$ in eq. (4.68), where we also did not impose the restriction $i \neq j$ for the same reason.

Before introducing the estimators based on pairs counted for the correlation function, we have to note that if the data and random catalogs have a different number of element *i.e.* $N_g \neq N_r$, we first need to normalize the pairs; otherwise, the signal can be canceled. For that, we are going to normalize for the total number of pairs the counts in eqs. (4.64), (4.65) and (4.67) as

$$DD(r) \to \overline{DD}(r) = \frac{DD(r)}{\frac{1}{2}N_g\,(N_g - 1)}\,, \tag{4.69}$$

$$RR(r) \to \overline{RR}(r) = \frac{RR(r)}{\frac{1}{2}N_r\,(N_r - 1)}\,, \tag{4.70}$$

$$DR(r) \to \overline{DR}(r) = \frac{DR(r)}{N_g N_r}\,, \tag{4.71}$$

because $N(N-1)/2$ are all the distances between N objects. We have to note that the only difference is for the DR pairs since they are the distances between two different catalogs. Many authors use $N(N-1)/2 \simeq N^2/2$ because generally, the catalogs have a large number of points to ensure a good statistic.

With eqs. (4.69) and (4.70), we can introduce the first estimator for the correlation function that was proposed firstly by Peebles & Hauser (1974) [196]

$$\hat{\xi}_{PH}(r) = \hat{g}_{PH}(r) - 1 = \frac{\overline{DD}(r)}{\overline{RR}(r)} - 1\,, \tag{4.72}$$

following the idea of *probability excess.* This estimator is ratio-unbiased for $\Delta \to 0$.

With eqs. (4.69) and (4.71), we can present the estimator proposed by Davis & Peebles (1983) [69]

$$\hat{\xi}_{DP}(r) = \hat{g}_{DP}(r) - 1 = \frac{\overline{DD}(r)}{\overline{DR}(r)} - 1\,. \tag{4.73}$$

This estimator is biased [147], but whose bias is negligible on small scales. Nonetheless, on large scales, the bias becomes essential, and $\xi(r)$ may be overestimated by $\hat{\xi}_{DP}(r)$ as it is demonstrated in [139].

Before introducing the final two estimators, we are going to use eqs. (4.70) and (4.71) to define

$$\hat{b}(r) = \frac{\overline{DR}(r)}{\overline{RR}(r)}\,, \tag{4.74}$$

which is not an estimator for the correlation function, but it will help us to quantify the bias.

With eqs. (4.69) to (4.71), we can introduce the estimator proposed by Hamilton [117]

$$\hat{\xi}_H(r) = \hat{g}_H(r) - 1 = \frac{\overline{DD}(r)\overline{RR}(r)}{\left[\overline{DR}(r)\right]^2} - 1\,. \tag{4.75}$$

Then, $\hat{g}_H(r) = \hat{g}_{DP}(r)/\hat{b}(r)$ according to eqs. (4.73) and (4.74). The Hamilton estimator is unbiased only in the unlikely case where the biases from $1/\hat{b}(r)$ and $\hat{g}_{DP}(r)$ cancel [139].

Finally, with eqs. (4.69) to (4.71), we introduce the estimator proposed by Landy & Szalay (1993) [147]

$$\begin{aligned} \hat{\xi}_{LS}(r) = \hat{g}_{LS}(r) - 1 &= \frac{\overline{DD}(r) - 2\overline{DR} + \overline{RR}(r)}{\overline{RR}(r)} \\ &= \frac{\overline{DD}(r)}{\overline{RR}(r)} - 2\frac{\overline{DR}(r)}{\overline{RR}(r)} + 1 \\ &= \hat{\xi}_{PH} - 2\hat{b} + 2\,, \end{aligned} \tag{4.76}$$

where we also used the eqs. (4.72) and (4.74). Since the Peebles & Hauser estimator is ratio-unbiased for $\Delta \to 0$, $\hat{\xi}_{LS}$ is approximately unbiased only if the expectation value of $\hat{b}(r)$ is 1. The Landy & Szalay estimator was designed to minimize both the cosmic variance and cosmic bias *i.e.* it is theoretically optimal with respect to both errors at least in the weak correlation limit $|\xi|, |\bar{\xi}(L)| << 1$, where L is the dimension of the survey as we saw in eqs. (4.39) to (4.41). Thus, it is quite insensitive to edge effects[6] and discreteness[7]. However, it is still affected by finite-volume effects[8], proportional to $\bar{\xi}(L)$, indeed the latter cannot be reduced without prior assumptions about clustering at scales larger than those probed by the survey [36].

In many cases, we do not face a catalog statistically homogeneous distributed. In fact, the points near the catalog boundary are more likely to be surrounded by areas with a low density number. This issue is accentuated in the direction of observation for far galaxies because we are restricted to the instrument's observational resolution. To fix that problem, it is crucial to know the selection function Φ that encodes the catalog geometry and how the galaxies were selected. Thus, Φ depends on the position but also in astrophysical parameters. Therefore, it is common to weigh with w_i every galaxy, which is going to try to homogenize the distribution and improve the problems above. In the literature, the most common weights are those that minimize the variance in the two-point estimator. This is the case of the Hamilton pair weighting [117]

$$w_{1\,2} = \frac{1}{\left(1 + \bar{n}\Phi J\right)^2}\,, \tag{4.77}$$

where $\bar{n}$ is the averaged number density, Φ the selection function to that pair and $J \simeq \int_0^r \xi \mathrm{d}V$ is the volume integral of the correlation function until the pair separation r. However, it is

---

[6] They are related to the geometry of the catalog. In general, estimators give less weight to galaxies near the edge than those far away from the boundaries.

[7] They are related to the fact that we assume that the observed galaxy distribution is a discrete representation of an underlying smooth field. The discreteness errors are proportional to $1/N_g$ at some power, and they become negligible for large enough N_g.

[8] They are because we can have access to only a finite number of structures of a given size in surveys, in particular, the mean density itself is not always well determined. These effects are roughly proportional to the average of the two-point correlation function over the survey $\bar{\xi}(L)$.

preferable to adopt separable weights because it is more efficient to compute the correlation estimator. We can approximate eq. (4.77) to obtain an individual weighting

$$w_i = \frac{1}{1 + \bar{n}\Phi_i J}\,, \tag{4.78}$$

where Φ_i is the selection function at point i. Thus, we can obtain again the pair weighting as $w_{1\,2} = w_1 w_2$. Note that J lost the argument of pair separation in eq. (4.78) and then, we need to estimate it for some characteristic scale r before even know the correlation estimator. This looks cyclic, but we are often concerned about studying the correlation for specific scales because what we want is to get the best signal to noise for that scale since the Hamilton weighting minimizes the variance.

Another weighting that minimizes the variance is the proposed by Feldman, Kaiser & Peacock [96], which is analogous to the Hamilton weighting but for the power spectrum instead of the correlation function[9], nonetheless, it is mostly used also in two-point correlation function analysis. The Feldman, Kaiser & Peacock weighting is given by

$$w_i = \frac{1}{1 + \bar{n}P(k)}\,, \tag{4.79}$$

which depends on the power spectrum at some scale k. Again, this could look cyclic, but we must focus on the scales where we want the best signal to noise for our analysis.

Note that $\bar{n}$ is not anymore constant since our initial dilemma before introduce the weights. However, we can consider that it depends only on the distance from us or r because in that direction we have the large variation of concentration.

In a more detailed analysis, some other weights count for galaxies' astrophysical properties, like luminosity and mass. Therefore, the total weight for each galaxy can be modeled as the multiplication of individual weights

$$w_i = \prod_j w_j^i\,. \tag{4.80}$$

To modify the pairs counts for the correlation estimators and include the weights, we only need to multiply them inside the sum. For example, the *data-data* pairs now are

$$DD(r) = \frac{1}{2}\sum_{i=1}^{N_g} w_i n_i^{\Delta}(r)\,, \tag{4.81}$$

with

$$n_i^{\Delta}(r) = \sum_{j=1;j\neq i}^{N} w_j \mathbb{1}_{[r,r+\Delta]}\left(|\vec{x}_i - \vec{x}_j|\right)\,, \tag{4.82}$$

where we just redefined eqs. (4.63) and (4.64). We can do the same for DR and RR from eqs. (4.70) and (4.71), but the random catalog must have the same selection function that data and also the same form of the number distribution in the line of sight direction.

---

[9] The power spectrum and the correlation function are Fourier counterparts, but its estimators in general, not.

4.4 Velocity Field And Redshift Space Distortions

So far, we have considered that the distribution of points is also static *i.e.* they do not move, but we know that galaxies are moving because the expansion of the Universe or by proper movements (eq. (1.1)). Obviously, to our scales, their movements do not change the distribution that we directly observe, but there is still a velocity field associated that can be measured and also reveal much cosmological information. To study them, we are going to base most of our discussions in [118, 80, 23, 76]. Since we want to recover the velocity field, we are going to use the continuity equation in Fourier space from eq. (2.106) for matter *i.e.* $w = c_s = 0$ and in the Newtonian limit $\phi' \to 0$[10]

$$\delta' + \theta = 0\,, \tag{4.83}$$

where we can use $\theta = \delta_i v^i \to \theta = i k_i v^i$ (see eq. (2.64)). Besides, we know that for irrotational fluids the velocity has the same direction of Fourier modes $v^i = v k^i$, then $\theta = ikv$ with v the velocity modulus. Replacing this into eq. (4.83) gives

$$v = \frac{i}{k}\delta'\,. \tag{4.84}$$

Since we are interested in the growing modes of δ, because they remain until late times, we can write $\delta = \delta_0 D(a)$ and then $\delta' = \delta D'(a)/D(a)$. This turns eq. (4.84) into

$$v = \frac{i}{k}\frac{D'(a)}{D(a)}\delta\,. \tag{4.85}$$

It is common to define the logarithm growth rate of matter perturbations as

$$\begin{aligned} f &\equiv \frac{\mathrm{d}\ln D(a)}{\mathrm{d}\ln a} \\ &= \frac{a}{D(a)}\frac{\mathrm{d}D(a)}{\mathrm{d}a}\,, \\ &= \frac{1}{\mathcal{H}}\frac{D'(a)}{D(a)}\,, \end{aligned} \tag{4.86}$$

then we can replace the last line into eq. (4.85) to obtain

$$\begin{aligned} v &= \frac{i\mathcal{H}f\delta}{k} \\ &= \frac{iaHf\delta}{k}\,, \end{aligned} \tag{4.87}$$

which gives us a direct relationship between the density field and the velocity field. Nonetheless, we have to remember two things; the first is that the velocity is a vector in the $\vec{k}$ direction

$$\vec{v} = iaHf\delta\frac{\vec{k}}{k^2}\,. \tag{4.88}$$

and the second is that the relation between the velocity and the density holds only in linear theory. Therefore, we can compute two-point correlations between galaxies velocities, which depend on the matter power spectrum and geometry in linear regimes.

---

[10] This is equivalent to the sub-horizon limit.

One of the problems to compute correlations for galaxy velocities is that we can not measure them directly. In fact, we obtain the velocities from the redshift, but we know that the Hubble flow of eq. (1.1) contaminates it. On the other hand, we determinate the position of a galaxy also by its redshift, but its proper movement contaminates the measure. Therefore, what we really can observe is the redshift space, but not the real one, which makes appear distortions in the direction of the line of sight called redshift space distortions (RSD). Thus, the coherent motions of galaxies towards overdense regions, in the linear regime, induce a squeezing effect on the correlation function observed. Besides, at small scales, the collapsing and virialized regions of galaxy clusters give rise to the so-called fingers-of-god (FOG). To these effects, we have to add that the proper movement of our galaxy, which resides in an overdense region of the Universe, increases the distortions. Hopefully, we have knowledge of most of these processes, which allows us to relate what we observe in redshift space with the real one. For that, we first need to connect the coordinates in both spaces, which can be done assuming that we are at a random point in our Universe in rest with respect to the CMB[11]. Then, we start defining the line-of-sight component of the galaxy proper velocity $\vec{v}_p$, as

$$v_{\parallel} \equiv \vec{v}_p \cdot \frac{\vec{r}}{r} \,, \tag{4.89}$$

where $r = |\vec{r}|$ and $\vec{r}$ is the source position in our frame, which defines the line-of-sight direction. Deriving again eq. (1.1), but now using the conformal time, we find $\vec{v} = \mathcal{H}\Delta\vec{r} + \vec{v}_p$ for real space. In redshift space, we can not distinguish the proper velocity, then we have just $\vec{v}_s = \mathcal{H}\Delta\vec{s}$, where the subscript s indicates redshift coordinates as the position $\vec{s}$ in that space. Nonetheless, both velocities $\vec{v}$ and $\vec{v}_s$ must be the same in the line-of-sight direction, since we observe them directly through the redshift. Then, we can relate both coordinates systems using just the line-of-sight components of the positions $\vec{r}$ and $\vec{s}$, because in that direction the distortions are produced, which gives

$$\begin{aligned} \vec{s} &= \vec{r} + \frac{\vec{v}_p}{\mathcal{H}} \\ &= \vec{r} + \frac{v_{\parallel}}{aH}\frac{\vec{r}}{r} \,, \end{aligned} \tag{4.90}$$

noting that in the second line we used eqs. (1.4) and (4.89). The redshift-space density field $\delta_s(\vec{s})$ can be obtained from the real-space one by requiring mass conservation *i.e.* $[1 + \delta_s(\vec{s})]\,\mathrm{d}^3 s = [1 + \delta(\vec{r})]\,\mathrm{d}^3 r$, which gives

$$\delta_s(\vec{s}) = [1 + \delta(\vec{r})] \left| \frac{\mathrm{d}^3 s}{\mathrm{d}^3 r} \right|^{-1} - 1 \,, \tag{4.91}$$

where the term in $|.|$ is the Jacobian $|J|$ given by the derivative of $\vec{s}$ in eq. (4.90) with respect to $\vec{r}$. Working in the plane-parallel approximation, the Jacobian can be written simply as

$$|J| = \left| \frac{\mathrm{d}^3 s}{\mathrm{d}^3 r} \right| \simeq 1 - f\partial_{\parallel} u \,, \tag{4.92}$$

where $\partial_{\parallel}$ indicates the derivative with respect to the line-of-sight direction in real space. Also, u is defined as

$$u \equiv -\frac{v_{\parallel}}{faH} \,, \tag{4.93}$$

---

[11] This is not true but helps us to start relating the redshift space with the real one. Obviously, we have to take account of these effects at some point.

and f the growth rate of eq. (4.86). Inserting eq. (4.92) into eq. (4.91), gives

$$\delta_s(\vec{s}) = \left(\delta(\vec{r}) + f\partial_{\parallel} u\right)\left(1 - f\partial_{\parallel} u\right)^{-1} . \tag{4.94}$$

For irrotational velocity fields we can write $u = \partial_{\parallel}\left(\nabla^2\right)^{-1}\theta$, where θ is the velocity divergence defined in eq. (2.64) and $\left(\nabla^2\right)^{-1}$ is the inverse of the Laplacian operator. Then, eq. (4.94) can be recast

$$\delta_s(\vec{s}) = \left(\delta(\vec{r}) + f\partial_{\parallel}^2\left(\nabla^2\right)^{-1}\theta\right)\left(1 - f\partial_{\parallel}^2\left(\nabla^2\right)^{-1}\theta\right)^{-1} . \tag{4.95}$$

We also know that in Fourier space, the direction of the irrotational velocity fields matches the direction of Fourier modes. Thus, we can write the operator $\partial_{\parallel}^2\left(\nabla^2\right)^{-1}$ in Fourier space as $\left(k_{\parallel}/k\right)^2$, which motivates us to define de cosine of the angle between the Fourier modes and the line-of-sight direction as

$$\begin{aligned} \mu &\equiv \frac{\vec{k}\cdot\vec{r}}{kr} \\ &= \frac{k_{\parallel}}{k} , \end{aligned} \tag{4.96}$$

noting that this definition rescues the spirit of the one made in eq. (2.204) but for another direction. Therefore, the calculation becomes clearest in Fourier space, but we must take care because now we have two Fourier spaces: one for the redshift and another for the real space. Expanding eq. (4.95) in redshift Fourier space, gives modes of the form

$$\begin{aligned} \delta_s\big(\vec{k}\big) &= \int \mathrm{d}^3 s\, \mathrm{e}^{-i\vec{k}\cdot\vec{s}}\delta_s(\vec{s}) \\ &= \int \mathrm{d}^3 r\, \mathrm{e}^{-i\vec{k}\cdot\vec{r}}\, \mathrm{e}^{ikf\mu u}\left[\delta(\vec{r}) + \mu^2 f\theta(\vec{r})\right] , \end{aligned} \tag{4.97}$$

where in the second line, we turned the redshift dependence into the real one using eq. (4.90), which results in an extra exponential factor. Besides, the change of variables in the integral makes appears a Jacobian, which cancels with that of eq. (4.95) in the plane-parallel approximation.

Now, we can compute the matter power spectrum in redshift space using section 4.4 and following eq. (4.29)

$$\begin{aligned} \left\langle\delta_s\big(\vec{k}\big)\delta_s^*\big(\vec{k}\,'\big)\right\rangle &= \int \mathrm{d}^3 s \int \mathrm{d}^3 s'\, \mathrm{e}^{-i\vec{k}\cdot\vec{s}}\, \mathrm{e}^{i\vec{k}\,'\cdot\vec{s}\,'} \quad \left\langle\delta_s(\vec{s})\,\delta^{s'}(\vec{s}\,')\right\rangle \\ &= \int \mathrm{d}^3 r \int \mathrm{d}^3 r'\, \mathrm{e}^{-i\vec{k}\cdot\vec{r}}\, \mathrm{e}^{i\vec{k}\,'\cdot\vec{r}\,'} \quad \Big\langle \mathrm{e}^{ifk\mu u}\, \mathrm{e}^{-ifk'\mu' u'}\left[\delta(\vec{r}) + \mu^2 f\theta(\vec{r})\right] \\ &\qquad \left[\delta(\vec{r}\,') + (\mu')^2 f\theta(\vec{r}\,')\right]\Big\rangle \\ &= (2\pi)^3\,\delta\big(\vec{k} - \vec{k}\,'\big)P_s(k,\mu) , \end{aligned} \tag{4.98}$$

where $P_s(k,\mu)$ is the anisotropic redshift-space power spectrum, which is defined, following eq. (4.30), as

$$P(k,\mu) = \int \mathrm{d}^3 x\, \mathrm{e}^{-i\vec{k}\cdot\vec{x}}\left\langle \mathrm{e}^{-ifk\mu\Delta u}\left[\delta(\vec{r}) + \mu^2 f\theta(\vec{r})\right]\left[\delta(\vec{r}\,') + \mu^2 f\theta(\vec{r}\,')\right]\right\rangle , \tag{4.99}$$

where $\Delta u = u\left(\vec{r}'\right) - u'\left(\vec{r}\right)$. The power spectrum in eq. (4.99) is almost exact because the only approximation made was the plane-parallel limit of eq. (4.92), which is valid for samples with pairs covering angles lower than 10° [165]. Besides, eq. (4.99) captures all the different regimes of distortions; for example, its square brackets describe the squeezing effect of Kaiser effect [136], while the exponential prefactor is responsible for the FOG. We can use the ansatz of [224], which assumes that the exponential prefactor and the term involving the density and velocity fields can be separated in the ensemble average. Under this, eq. (4.99) becomes in

$$P_s\left(k,\mu\right) = \mathrm{e}^{-(fk\mu\sigma_v)^2}\left[P_{\delta\delta}(k) + 2\mu^2 f P_{\delta\theta} + \mu^4 f^2 P_{\theta\theta}(k)\right]\,, \tag{4.100}$$

where $P_{\delta\delta}, P_{\delta\theta}, P_{\theta\theta}$ are the non-linear mass density-density, density-velocity divergence, and velocity divergence-velocity divergence power spectrum respectively, which are defined in the same way that eq. (4.30) but taking care of the random fields involved. The ensemble average over the exponential prefactor makes appear the pairwise velocity dispersion, defined in [76] as

$$\sigma_v^2 = \frac{1}{6\pi^2}\int P_{\theta\theta}(k)\mathrm{d}k\,. \tag{4.101}$$

Note that in linear regime (eq. (4.83)), the different power spectrum are the same $P_{\delta\delta} = P_{\delta\theta} = P_{\theta\theta} = P_m(k)$, where P_m is the matter power spectrum of eq. (4.50). Also, for $k\sigma_z << 1$ we can recover the original Kaiser formula in [136]

$$P_s\left(k,\mu\right) = \left[1 + 2\mu^2 f + \mu^4 f^2\right] P_m\left(k\right)\,. \tag{4.102}$$

The redshift-space anisotropic two-point correlation function can be obtained by Fourier-transforming the anisotropic power spectrum of eq. (4.100) or eq. (4.102), through eq. (4.33). This gives us the correlation function in the real redshift space, also called configuration space, which is

$$\begin{aligned}\xi\left(r_\perp, r_\parallel\right) &= \int \frac{\mathrm{d}^3k}{(2\pi)^3}\,\mathrm{e}^{i\vec{k}\cdot\vec{s}} P_s\left(k,\mu\right) \\ &= \sum_l \xi_l^s(s)\mathcal{P}_l\left(\nu\right)\,,\end{aligned} \tag{4.103}$$

where ν is the cosine of the angle between the line-of-sight direction and the position $\vec{s}$ in configuration space, and $r_\perp, r_\parallel$ are the distance perpendicular and parallel to the line-of-sight direction, respectively, which satisfy

$$\nu = \frac{r_\parallel}{s}\,, \tag{4.104}$$

$$r_\perp = \sqrt{s^2 - r_\parallel^2}\,. \tag{4.105}$$

Also, in the second line of eq. (4.103), we made a multipole expansion of the correlation function, where the $\mathcal{P}_l$ are the already known Legendre polynomials and the ξ_l^s are the multipoles, defined as

$$\xi_l^s(s) = i^l \int \frac{\mathrm{d}^3k}{(2\pi)^3}\,P_l^s(k) j_l(ks), \tag{4.106}$$

where j_l denotes the spherical Bessel functions of order l and the P_l^s are the multipoles of the anisotropic power spectrum, defined as

$$P_l^s(k) = \frac{2l+1}{2}\int_{-1}^{1}\mathrm{d}\mu\,P_s\left(k,\mu\right)\mathcal{P}_l\left(\mu\right)\,. \tag{4.107}$$

Although the study of redshift space distortions helps us to better understand the distribution of galaxies observed, we also must consider that those galaxies are just tracers of the matter density field, and then, what we observe is biased with respect to the theoretical results. Nonetheless, we can relate both densities fields by a bias factor, defined as

$$b \equiv \frac{\delta_g}{\delta} \,, \tag{4.108}$$

where δ_g is the galaxy density contrast. In the linear regime, we can assume that b is constant, which means that it does not depend on k or r, neither time. Nevertheless, it depends on the type of galaxy we are tracing and considering, and also, it can adopt estranger functional forms. It is important to remark that even though the density contrast is biased, the velocity field is not, which implies that galaxies violate the linear continuity equation (see eq. (4.83)). However, do not be aware because this violation could be caused by some hydrodynamical effects like galaxies appearing and disappearing, which breaks continuity. Therefore, the power spectrum from galaxies P_g is related to the matter power spectrum P_m as

$$P_g = b^2 P_m \,. \tag{4.109}$$

Nonetheless, we must take care of the anisotropies induced by RSD because the bias is only applied for the components that carry δ, for example, the linear anisotropic power spectrum in eq. (4.102) becomes in

$$\begin{aligned} P_{s,g}\left(k,\mu\right) &= \left[b^2 + 2\mu^2 f b + \mu^4 f^2\right] P_m\left(k\right) \\ &= b^2 \left[1 + \beta\mu^2\right]^2 P_m\left(k\right) \,, \end{aligned} \tag{4.110}$$

where we defined the distortion parameter β as

$$\beta \equiv \frac{f}{b} \,, \tag{4.111}$$

which governments the strength of the anisotropy. We can obtain a direct measure of β taking the ratio of the quadrupole to the monopole defined in eq. (4.107), which for the anisotropic power spectrum of eq. (4.110) is

$$\frac{P^s_2}{P^s_0} = \frac{\frac{4}{3}\beta + \frac{4}{7}\beta^2}{1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2} \,. \tag{4.112}$$

4.5 Geometrical Distortions

As we mentioned, the clustering of galaxies can be studied by correlation estimators, which only need point positions. A problem that we ignored is that we have access only to the angular position of a galaxy in the sky and its redshift, where the latter introduces several errors, as we saw in section 4.4. Although we can analyze angular correlations, the $3D$ correlators contain more information, like anisotropies in the line-of-sight direction. Therefore, in order to construct them we require the galaxy positions in comoving-Cartesian coordinates, but the translation from angular and redshift separations to comoving separations depends on $D_A(z)$ (eq. (1.64)) and $H(z)$ (eq. (1.72)), respectively. Thus, we need to assume some cosmological model and the values of its parameters, but the assumption of an incorrect cosmology leads to an apparent anisotropy of clustering statistics, which is similar to that caused by RSD. The challenge to measure this effect is that redshift distortions are generally larger than cosmological or geometrical distortions, so we also need appropriate treatment of the first to obtain reliable measures of the latter. We are going to follow mostly the procedure of [228, 23, 226, 122].

Imagine that we observe an object or event at mean redshift z but with sizes $\delta\theta$ and δz. Then, its comoving sizes in the line-of-sight ($r_\parallel$) and transverse ($r_\perp$) directions are related to δz and $\delta\theta$, respectively as

$$r_\parallel = \frac{c\delta z}{H(z)}\,, \tag{4.113}$$

$$r_\perp = (1+z)\,\mathrm{d}_A(z)\delta\theta\,, \tag{4.114}$$

where in the first equation, we used eq. (1.53) for a small δz, which is why we recover the c units. While, for the second equation, we used the definition of eq. (1.61), but since we are seeking the comoving size, it is corrected by a scale factor (eq. (1.62)). For the transverse direction, the combination $r_\perp/\mathrm{d}_A$ does not depend on the cosmology because $(1+z)\delta\theta$ is what we measure. Therefore, if we assume some cosmology to compute the distance to our object, its real size is

$$r_\perp = \frac{\mathrm{d}_A}{\mathrm{d}_{A_r}} r_{\perp_r}\,, \tag{4.115}$$

where the subscript r refers to the reference cosmology. Similarly, form eq. (4.113) the product $r_\parallel H$ remains constant in the line-of-sight direction under a change of cosmology; thus we have the next relation

$$r_\parallel = \frac{H_r}{H} r_{\parallel_r}\,. \tag{4.116}$$

Note that the measures $r_\parallel$ and $r_\perp$ change in a different way, which implies that if the two-point correlation function or the power spectrum are isotropic in some cosmology, they will become anisotropic for any other cosmology which is known as the Alcock-Paczynski effect [18].

To study geometrical distortions, we need an object or event, which means that we are going to use a standard ruler for our purposes, like in chapter 1. For that, we remember the two length scales encoded in the matter power spectrum, which are the comoving Hubble horizon size at the matter-radiation equality and the comoving sound horizon size at the drag epoch at which baryons were released from photons, where the last one is also encoded in the so-called BAO signature of the two-point correlation function. As observers and using eqs. (4.113) and (4.114), we can measure four distance ratios described in [228], we will focus on those of the BAO peak

$$\delta\theta_s = \frac{r_s\left(z_{\mathrm{d}rag}\right)}{(1+z)\,\mathrm{d}_A(z)}\,, \tag{4.117}$$

$$\delta z_s = \frac{r_s\left(z_{\mathrm{d}rag}\right)H(z)}{c}\,, \tag{4.118}$$

noting that we used the same distance measure in the line-of-sight and transverse directions, since we assumed that such process is isotropic[12]. Also, for eqs. (4.117) and (4.118) we will

---

[12] Then, the anisotropies that we observed nowadays were developed later by the redshift and geometrical distortions, and the non-linear evolution.

not have the same ratio as eqs. (4.115) and (4.116) to change of cosmology because now we must consider that the sound horizon r_s (eq. (3.127)) and $z_{\mathrm{d}rag}$ also depend on cosmology. Then for the BAO signature, we have to use the next relations to relate the real and reference cosmology

$$r_\perp = \frac{\mathrm{d}_A r_{s_r}\left(z_{\mathrm{d}rag_r}\right)}{\mathrm{d}_{A_r} r_s\left(z_{\mathrm{d}rag}\right)} r_{\perp_r} = \alpha_\perp r_{\perp_r}\,, \tag{4.119}$$

$$r_\parallel = \frac{H_r r_{s_r}\left(z_{\mathrm{d}rag_r}\right)}{H r_s\left(z_{\mathrm{d}rag}\right)} r_{\parallel_r} = \alpha_\parallel r_{\parallel_r}\,, \tag{4.120}$$

where we defined the ratios $\alpha_\perp$ and $\alpha_\parallel$ as

$$\alpha_\perp \equiv \frac{\mathrm{d}_A r_{s_r}\left(z_{\mathrm{d}rag_r}\right)}{\mathrm{d}_{A_r} r_s\left(z_{\mathrm{d}rag}\right)}\,, \tag{4.121}$$

$$\alpha_\parallel \equiv \frac{H_r r_{s_r}\left(z_{\mathrm{d}rag_r}\right)}{H r_s\left(z_{\mathrm{d}rag}\right)}\,. \tag{4.122}$$

We can also define some related quantities, as the shift parameter α

$$\alpha = \alpha_\parallel^{1/3}\alpha_\perp^{2/3}\,, \tag{4.123}$$

and the warping parameter ϵ

$$1+\epsilon = \left(\frac{\alpha_\parallel}{\alpha_\perp}\right)^{1/3}\,. \tag{4.124}$$

This notation separates the isotropic deformations or dilation in α from anisotropic deformations or warping in ϵ [185]. In other words [211], α is the BAO measurement expected from spherically averaged clustering measurements and ϵ the significance of the BAO feature introduced into the quadrupole by the assumption of a wrong cosmology. Therefore, in surveys that have not the power of constraining anisotropies, we can relate both cosmologies using just the isotropic deformation as

$$D_V(z) = \alpha \frac{r_s}{r_{s_r}} D_{V_r}(z)\,, \tag{4.125}$$

where the sound horizons are evaluated at their respective drag redshifts and D_V is the volume-averaged distance, defined as

$$D_V(z) \equiv \left[cz\left(1+z\right)^2 \frac{\mathrm{d}_A^2(z)}{H(z)}\right]^{1/3}\,. \tag{4.126}$$

Finally, since we have complete knowledge of the reference or fiducial cosmology because we use it to compute the galaxy distances, we also can constrain the true cosmology by the equations found in this chapter using the parameters $\alpha, \epsilon, \alpha_\parallel, \alpha_\perp$.

Chapter 5

Algorithms Insights

In chapter 4, we studied some estimators for the two-point correlation function, which just are based on counting pairs to a certain distance. However, at the time to compute them, we have limited resources; for example, if we have N galaxies, counting the DD pairs over M bins takes the order of $\mathcal{O}(MN^2)$, which increases a lot for $N-$point correlation functions. This issue is an obvious problem for big surveys of galaxies because it requires hardware and time-consuming. Also, we need to produce the catalog of random points with the same geometry of the original data, but with another distribution in space, but not in redshift. Then, we need to learn the redshift distribution from data to generate samples from the randoms.

Furthermore, we have developed all the cosmological theory behind. Now we need to contrast our estimators with physics formulas to recover essential parameters like the matter density nowadays or for field reconstruction like the potential. This problem also requires a lot of time-consuming if we want to present competent values. Luckily, we have developed several techniques in data mining and machine learning that help us in this task, which we are going to review in this chapter following mostly [164, 258, 195, 133, 22].

5.1 Calculations Of Pairwise Distances

To compute the two-point correlation function using some of the estimators in eqs. (4.72), (4.73), (4.75) and (4.76), we need to count all the distances between points which takes $DD \sim \mathcal{O}(N_1^2)$, $DR \sim \mathcal{O}(N_1 N_2)$, $RR \sim \mathcal{O}(N_2^2)$, where N_1 and N_2 are the numbers of points in the data and random catalogs, respectively. Then, we need to group those distance into the M bins, which increases the orders mentioned above by a factor M, and in the end, we obtain a very inefficient algorithm. A more practical approach uses dual-trees, an efficient class of divide-and-conquer algorithms based on the nearest-neighbor problem.

The nearest-neighbor problem consists in finding the nearest point[1] to $\vec{x}_i \in \mathbb{R}^m$[2] between

---

[1] For distance; we are going to refer to any metric measure, where the standard Euclidean distance is the most common choice, especially for our purposes.

[2] We can work these algorithms in any metric space. However, in our case, we will work in the real space

the points of a set $\mathcal{X}$. The naive approach computes the $\binom{N}{2}$ distances, with N the number of points in $\mathcal{X}$, but this takes $\mathcal{O}(N^2)$ in time. However, a smart observation is to consider first a subset of points called $\mathcal{R}$, obviously, $\mathcal{R} \subseteq \mathcal{X}$, which can be closed into a "bounding box" or a rectangular $m-$volume containing all the points in $\mathcal{R}$. The advantage of this box is that we can quickly compute two bounds: the closest distance that any point in $\mathcal{R}$ can be to $\vec{x}_i$, which we will call $\mathrm{d}_l\left(\vec{x}_i, \mathcal{R}\right)$, and the farthest possible pairwise distance, called $\mathrm{d}_u\left(\vec{x}_i, \mathcal{R}\right)$. Furthermore, if we compute the distance between any test point in $\mathcal{R}$ and $\vec{x}_i$, we can say that it is the smallest distance we have seen so, called $\hat{\mathrm{d}}$. Now, $\hat{\mathrm{d}}$ provides us an upper bound on the distance between $\vec{x}_i$ and the true nearest neighbor. Therefore, if $\hat{\mathrm{d}} < \mathrm{d}_l\left(\vec{x}_i, \mathcal{R}\right)$, we know automatically that none of the points in $\mathcal{R}$ can be the true nearest neighbor of $\vec{x}_i$, which could save us much time. In order to use this observation, we will also employ a divide-and-conquer approach to make a fast algorithm, because if we iterate through the pairs one-by-one, we would have a similar order of time that the naive method. This can be done by a space-partitioning tree data structure, where we can recursively split the computation into smaller subproblems.

We can use a $K-$dimensional tree (KD tree) for the aforementioned tree, which generalizes the two-dimensional or Quad-trees and the 3-dimensional or Oct-trees to an arbitrary number of dimensions. The KD tree is a binary tree structure which recursively partitions the parameter space along the data axes, dividing it into nested orthotropic regions into which data points are filed. Also, its construction is very fast because partitioning is performed only along the data axes, and no distances need to be computed. Once constructed, the nearest-neighbor of a query point can be determined with only $\mathcal{O}(\log(N))$ distance computations[3].

Another option is the Ball trees, which were developed to address the inefficiencies of KD Trees in higher dimensions. Instead of partition the data along Cartesian axes like KD tress, the Ball trees partition data in a series of nesting hyper-spheres. The price to pay is that the construction of a Ball tree is costlier than of the KD tree. However, the benefit, between others, is obtaining a data structure that can be very efficient on highly structured data, even in very high dimensions. A Ball tree recursively divides the data into nodes defined by a point center $\vec{c}_i$ and a radius r, such that each point in the node lies within the hyper-sphere defined by r and $\vec{c}_i$. Besides, the number of candidate points for a neighbor search is reduced using the triangle inequality $|x+y| \leq |x|+|y|$, which makes that a single distance calculation between a test point and the point center is sufficient to determine a lower and upper bound on the distance to all points within the node. Finally, another advantage of Ball trees is that it is easy to compute the number of pairwise distances in a bin. For example, for all the distances between R and $R + \mathrm{d}R$, we only need to apply two queries (one for R and one for $R + \mathrm{d}R$). Then, we rest both to obtain the wanted, which makes this data structure excellent for two-point correlation functions.

Finally, to improve the trees mentioned above, we can construct two trees, one on references and one on queries (hence dual-tree algorithm). Then, the computations for the queries are going to take just $\mathcal{O}(M \log(N))$, where M is the number of bins, but the trees constructions will take $\mathcal{O}(MN \log(N))$.

---

with $m = 3$ dimensions.

[3] For low-dimensional ($m < 20$) neighbors searches, the KD tree is fast, but it becomes inefficient as m grows very large: this is one manifestation of the so-called “curse of dimensionality”

5.2 Kernel Density Estimation

One typical problem, for a set of points in a $m-$dimensional space, is to find the underlying probability distribution function (PDF) -or just the density distribution- of those. It is intuitive to construct its histogram for one-dimensional points, which can be generalized to more dimensions. One trouble, among others, with histograms, is that they depend on the bin width, which for a large size, makes that the probability distribution function looks uniform. Also, if we want to take new samples from the probability distribution function, we need to fit some analytic or parametric curve to the PDF. In contrast, the kernel density estimation (KDE) is a method for estimating the probability distribution function of the data, nonparametrically *i.e.* with virtually no assumptions on the functional form of the PDF.

For a given set of N measurements $\vec{x}_i$, the kernel density estimator at an arbitrary position $\vec{x}$ is defined as

$$\hat{f}\left(\vec{x}\right) = \frac{1}{Nh^m}\sum_{i=1}^{N} K\left(\frac{\mathrm{d}\left(\vec{x},\vec{x}_i\right)}{h}\right) , \tag{5.1}$$

where m is the dimension number of the parameter space, and h is known as the bandwidth, which defines the size of the kernel $K(x)$. Then, the local density is estimated as a weighted mean of all points, where the weights are specified through the kernel and typically decrease with distance $\mathrm{d}\left(\vec{x},\vec{x}_i\right)$, which mostly is the standard Euclidean distance. Also, the kernel function can be any function, that is positive at all points ($K(x) \geq 0$), normalized to unity ($\int K(x)\mathrm{d}x = 1$), with a vanishing mean ($\int xK(x)\mathrm{d}x = 0$) and with a variance greater than zero ($\int x^2K(x)\mathrm{d}x > 0$). A kernel often used is the Gaussian kernel

$$K(x) = \frac{1}{\sqrt{(2\pi)^m}}\mathrm{e}^{-x^2/2} , \tag{5.2}$$

but also, there are many kernels like the top-hat, epanechnikov, exponential, linear, between others. Since the estimator in eq. (5.1) requires the computation of all distances, the naive approach takes $\mathcal{O}\left(N^2\right)$. However, in order to improve the scaling, approximations can be made if the data is space-partitioning in a tree data structure -like a KD or a Ball trees- as we saw in section 5.1, because the contribution for points with $\mathrm{d}\left(\vec{x},\vec{x}_i\right) \gg h$ can be ignored, but this could introduce errors in the estimator.

The bandwidth h in eqs. (5.1) and (5.2) acts as a smoothing parameter, controlling the tradeoff between bias and variance in the result. A large h leads to a very smooth density distribution with high bias; while, a small bandwidth leads to an unsmooth density distribution with high variance. For those reasons, it is essential to choose an optimal h, which can be found using, for example, a cross-validation technique.

Another useful application of KDE is learning a non-parametric generative model from a dataset to draw new samples efficiently. Therefore, the new data consist of linear combinations of the input data, with weights probabilistically drawn given the KDE model. Furthermore, we will use this application to learn the PDF from the data redshift distribution, to then apply the estimator to the random samples.

5.3 Bayesian Statistical Inference

The study of statistics can be summarized into two frameworks: the classical or frequentist statistics and the Bayesian inference. The tenets of the first are that probabilities refer to relative frequencies of events that depend on parameters that are fixed and unknown constants. Since they are not fluctuating, the probability statements about parameters are meaningless. In contrast, we have that the probability describes the degree of subjective belief for Bayesian inference, not the limiting frequency. Also, the probability can be made about things other than data, including model parameters and models. Therefore, inferences about parameters can be made by producing their probability distribution, which quantifies the uncertainty of our knowledge about that parameter.

The Bayesian method has become popular over the last few decades because of the maturing development of its philosophical and technical foundations and the ability to perform the required computations. Besides, assuming that all of the supplied information is correct, the Bayesian method yields optimal results. The frequentist versus Bayesian controversy sets in when we apply the theorem of Bayes to the likelihood function $f\left(D|M\right)$

$$f\left(M|D\right)=\frac{f\left(D|M\right)f\left(M\right)}{f\left(D\right)}\,, \tag{5.3}$$

where M, D stand for model and data, respectively. To be more precise, we are going to make explicit acknowledge the presence of prior information I and the fact that models are described by parameters $\vec{\theta}$ that we want to estimate from data

$$f\big(M,\vec{\theta}\,\big|D,I\,\big)=\frac{f\big(D\,\big|M,\vec{\theta},I\,\big)f\big(M,\vec{\theta}\,\big|I\,\big)}{f\big(D\,\big|I\,\big)}\,. \tag{5.4}$$

This quantifies the rule for "combining an initial belief with new data to arrive at an improved belief" and says that the "improved belief" is proportional to the product of the "initial belief" and the probability that the "initial belief" generates the observed data. Furthermore, $f\big(M,\vec{\theta}\,\big|D,I\,\big)$ is called the posterior PDF for the model M with parameters $\vec{\theta}$, given a data D and other prior information I. The term $f\big(D\,\big|M,\vec{\theta},I\,\big)$ is the likelihood of the data given some model M and the fixed parameters $\vec{\theta}$ that describe it. The term $f\big(M,\vec{\theta}\,\big|I\,\big)$ is the *priori* joint probability -also called just prior- for the model M and its parameters $\vec{\theta}$ in the absence of any data. Finally, the term $f\big(D\,\big|I\,\big)$ is the probability of data, or the prior predictive probability for D, which is not usually explicitly computed since it is also the proper normalization for the posterior PDF.

5.4 Metropolis-Hastings

Sometimes the functions that enter in the likelihood or prior are challenging to handle analytically, which forces us to obtain the Bayesian inference using computational resources. However, the direct exploration of the posterior for a model with a large number of parameters is impractical and often impossible, even using a computer. For example, if we study the parameter space using a lattice of N points for each parameter, the posterior would take $\mathcal{O}(N^m)$, where m is the number of parameters. Therefore, we need some mechanism to avoid those computations, and here some algorithms based on generating Markov chains appear. A Markov chain is a sequence of random variables in which a given value nontrivially depends only on its preceding value, which means that given the present value θ, past θ_{i-1} and future θ_{i+1} values are independent. While more samples we obtain from the chain, we make a better integration of the parameter space as in a Monte Carlo integration. This is why there are several algorithms based on Markov Chain Monte Carlo (MCMC), and one of the most famous is the Metropolis-Hastings algorithm.

In order for a Markov chain to reach a stationary distribution proportional to $p\big(\theta\big)$, the probability of arriving at a point θ_{i+1} must be proportional to $p\big(\theta i+1\big)$,

$$p\big(\theta_{i+1}\big) = \int T\big(\theta_{i+1}|\theta_i\big)p\big(\theta_i\big)\mathrm{d}\theta_i\,, \tag{5.5}$$

where $T(\theta_{i+1}|\theta_i)$ is the transition probability, also called the jump kernel or transition kernel, which satisfies the detailed balance condition

$$T\big(\theta_{i+1}|\theta_i\big)p\big(\theta_i\big) = T\big(\theta_i|\theta_{i+1}\big)p\big(\theta_{i+1}\big)\,. \tag{5.6}$$

Many of the MCMC algorithms differ in their choice of the transition kernel. The Metropolis-Hastings algorithm adopts the kernel

$$T\big(\theta_{i+1}|\theta_i\big) = p_{acc}\big(\theta_i,\theta_{i+1}\big)K\big(\theta_{i+1}|\theta_i\big)\,, \tag{5.7}$$

where $K\big(\theta_{i+1}|\theta_i\big)$ is an arbitrary density distribution, usually a Gaussian PDF, and p_{acc} is the accepted probability which we can compute using the above equations as

$$p_{acc}\big(\theta_i,\theta_{i+1}\big) = \frac{K\big(\theta_i|\theta_{i+1}\big)p\big(\theta_{i+1}\big)}{K\big(\theta_{i+1}|\theta_i\big)p\big(\theta_i\big)}\,. \tag{5.8}$$

If $p_{acc} \geq 1$, the proposed point θ_{i+1} is always accepted. But when θ_{i+1} is rejected, θ_i is added to the chain instead.

Nonetheless, we need to take care because when the posterior is multimodal, the simple Metropolis-Hastings algorithm can become stuck in a local mode and not find the global best minima within a reasonable running time. There are many better algorithms, such as Gibbs sampling, parallel tempering, various genetic algorithms, and nested sampling.

5.5 Neural Networks

To understand neural networks, we first need to study a single neuron. This has a number I of inputs x_i and one output, which we will call y. Associated with each input we have a weight w_i with $i = 1, ..., I$, but there could be an additional parameter w_0 of the neuron called a bias, which we may view as being the weight associated with an input X_0 that is set to 1. With the inputs x_i we can compute the activation of the neuron, defined as

$$a = \sum_i^I w_i x_i\,, \tag{5.9}$$

where the sum is from $i = 0$ if there is a bias and from $i = 1$ otherwise. Then, the output y is set as a function $f(a)$ of the activation, but there are several possible activation functions, which can be linear as

$$f(a) = a\,, \tag{5.10}$$

sigmoid as the logistic function

$$f(a) = \frac{1}{1+\mathrm{e}^{-a}} \quad y \in (0,1)\,, \tag{5.11}$$

or the hyperbolic tangent function

$$f(a) = \tanh(a) = \frac{\mathrm{e}^{a} - \mathrm{e}^{-a}}{\mathrm{e}^{a} + \mathrm{e}^{-a}} \quad y \in (-1,1)\,, \tag{5.12}$$

between other deterministic activation functions, but also, there could be some that are stochastic activation functions.

The single neuron can be seen as an extrapolation or regression problem in which we fit some curve to the data to determine the weights and predict another output from unknown inputs. To find these weights, it is common to minimize the difference between the known outputs and the predicted ones through a loss function, which is a problem that we will board after to introduce general neural networks.

To obtain a higher degree of complexity that is required in some problems, we can connect multiple neurons, making the output of one neuron be the input to another and creating a neural network. These can be divided into two classes based on their connectivity; the first are feedforward networks in which all the connections are directed such that the network forms a directed acyclic graph. And the second ones are those that are not feedforward networks, also known as feedback networks.

The multilayer perceptron is an example of a feedforward network with input neurons, hidden neurons, and output neurons, which are the layers. These multilayer networks are also known as backpropagation networks in supervised learning. Besides, the hidden neurons may be arranged in more layers; however, the most common multilayer perceptrons have a single hidden layer and are known as two-layer networks because we do not include the inputs in the counting. This kind of network defines a nonlinear parameterized mapping from the inputs x_i to the outputs y_j. Both are connected by a continuous function that also depends on the parameters or weights w_k and the net's architecture *i.e.* the functional form of the mapping, denoted by $\mathcal{A}$.

Imagine a network with only one hidden layer in a regression problem, with the following architecture $\mathcal{A}$: N inputs x_i that are connected with D nodes in the hidden layer, finally connected with the M outputs y_j. The activation function for the first connection is going to be $f(a)$ for all the nodes; while, for the second connection, the activation function will be $h(a)$. Then, the mapping of the network has the next form for the hidden layer

$$a_k = \sum_{i=0}^{N} w_{ki}^{(1)} x_i\,; h_k = f\left(a_k\right)\,, \tag{5.13}$$

and for the output layer

$$a_j = \sum_{k=0}^{D} w_{jk}^{(2)} h_k \, ; y_j = h\left(a_j\right) \, . \tag{5.14}$$

5.6 Density-Estimation Likelihood-Free Inference

The density-estimation likelihood-free inference (DELFI) [48, 87, 188, 159, 20, 22] and the approximate bayesian computation (ABC) [156, 52, 262], between others, belong to the family of likelihood-free inference (LFI). They perform Bayesian inference under very complex models, using only forward simulations. We will focus on studying DELFI, which aims to train a flexible density estimator for the target of the posterior PDF from a set of simulated data-parameters pairs. Also, it can yield high-fidelity posterior inference from orders-of-magnitude fewer simulations than traditional ABC-based methods.

DELFI turns inference into a density estimation tack on set of simulated parameter-data pairs $\left\{\vec{\theta}, \vec{t}\,\right\}$. For this, there are principally three ways:

1. Fit a model to the joint density $p\big(\vec{\theta}, \vec{t}\,\big)$, then obtain the posterior by evaluating the joint density at the observed data $\vec{t}_O$, as $p\big(\vec{\theta}\,\big|\vec{t}\,\big) \propto p\big(\vec{\theta}, \vec{t} = \vec{t}_O\big)$ [21].
2. Fit a model to the conditional density $p\big(\vec{\theta}\,\big|\vec{t}\,\big)$, then obtain the posterior by evaluating at the observed data $\vec{t}_O$ [188, 158].
3. Fit a model to the conditional density $p\big(\vec{t}\,\big|\vec{\theta}\big)$, obtain the likelihood by evaluating at the observed data $\vec{t}_O$, and multiply by the prior to obtain the posterior $p\big(\vec{\theta}\,\big|\vec{t}\,\big) \propto p\big(\vec{t}\,\big|\vec{\theta}\big)p\big(\vec{\theta}\big)$ [159, 190].

In our case, and given the advantages summarized in [22], we will use the third way. With this choice made, DELFI can be broadly summarized as follows:

- Run simulations at different parameter values $\vec{\theta}$ to obtain simulated parameters-data pairs $\big\{\vec{\theta}, \vec{t}\big\}$.
- Fit a parametric conditional density estimator $p\big(\vec{t}\,\big|\vec{\theta}\,; \vec{w}\big)$ to the simulations $\big\{\vec{\theta}, \vec{t}\big\}$, where $\vec{w}$ are the weights.
- Evaluate the estimated conditional density at the observed data $\vec{t}_O$ to obtain the (learned) likelihood function $p\big(\vec{t}_O\big|\vec{\theta}\,; \vec{w}\big)$.

Besides, some questions must be addressed in order to perform an efficient algorithm for DELFI; they can be found in [22]. Within this matter and as a summarize, we use neural density estimators (NDEs) as a flexible and efficient conditional density estimation framework for DELFI [188, 190, 159]. For efficient acquisition of simulations, we use active learning, allowing the NDEs to call the simulator to run new simulations on-the-fly based on the current likelihood-surface approximation. Finally, it is important to review data compression schemes for accelerating DELFI.

The NDEs provide a flexible framework to obtain parametric models for conditional probability densities $p\big(\vec{t}\,\big|\vec{\theta}\,; \vec{w}\big)$, parameterized by neural networks with weights $\vec{w}$, which can be trained on a set of simulated data-parameters pairs $\left\{\vec{\theta}, \vec{t}\right\}$. We will use masked autoregressive flows (MAFs) [189] as NDEs, but there are more classes as, for example, the mixture density networks (MDNs) [39]. We use the chain rule to separate the probability density as a product of one-dimensional conditionals

$$p\big(\vec{t}, \vec{\theta}\big) = \prod_{i=1}^{\dim \vec{t}} p\big(t_i \big| \vec{t}_{1:i-1}, \vec{\theta}\big). \tag{5.15}$$

Then, the neural autoregressive density estimators construct parametric densities for this set of one-dimensional conditionals, where the parameters of each of the conditionals are parameterized as a neural network [254]. As an example, we could model each conditional $p\big(t_i \big| \vec{t}_{1:i-1}, \vec{\theta}\big)$ as a Gaussian, where its mean and variance are free functions of $\big(\vec{t}_{1:i-1}, \vec{\theta}\big)$, parameterized by a neural network. This is the case for masked autoencoders for density estimation (MADEs) [104], because the means and variances of each conditional density are parameterized by the neural network, where its layers weights are masked in such a way that the output nodes for $p\big(t_i \big| \vec{t}_{1:i-1}, \vec{\theta}\big)$ only depends on $\big(\vec{t}_{1:i-1}, \vec{\theta}\big)$ *i.e.* the autoregressive property is preserved. Besides, the hidden layers of the MADE have some non-linear activation functions (see eqs. (5.11) and (5.12)), while the output nodes associated with the conditional means have linear activation (see eq. (5.10)), and the output nodes associates with the variance have exponential activation to ensure positivity.

By learning the means and variances of the autoregressive conditionals, a MADE can be thought of as learning the transform of the random variable $\vec{t}$ back to the unit normal

$$\vec{t}\big|\vec{\theta} \to \vec{u}\big(\vec{t}, \vec{\theta}; \vec{w}\big) \sim \mathcal{N}\left(\mathbb{0}, \mathbb{1}\right) , \tag{5.16}$$

$$t_i\big|\vec{\theta} \to u_i = \frac{t_i - \mu_i\big(\vec{t}_{1:i-1}, \vec{\theta}; \vec{w}\big)}{\sigma_i\big(\vec{t}_{1:i-1}, \vec{\theta}; \vec{w}\big)} , \tag{5.17}$$

where $\vec{w}$ are the masked weights of the neural network. Therefore, the parametric density estimator for a MADE is given by

$$\begin{aligned} p\big(\vec{t}\big|\vec{\theta}; \vec{w}\big) &= \prod_i p\big(t_i \big| \vec{t}_{1:i-1}, \vec{\theta}; \vec{w}\big) \\ &= \mathcal{N}\left(\vec{u}\big(\vec{t}, \vec{\theta}; \vec{w}\big) \Big| \mathbb{0}, \mathbb{1}\right) \left| \frac{\partial \vec{u}\big(\vec{t}, \vec{\theta}; \vec{w}\big)}{\partial \vec{t}} \right| \\ &= \mathcal{N}\left(\vec{u}\big(\vec{t}, \vec{\theta}; \vec{w}\big) \Big| \mathbb{0}, \mathbb{1}\right) \prod_{i=1}^{\dim \vec{t}} \sigma_i\big(\vec{t}, \vec{\theta}; \vec{w}\big) . \end{aligned} \tag{5.18}$$

Single MADEs have some limitations, like their sensitivity to factorization order in eq. (5.15) and the assumption of simple conditionals may be overly restrictive. Nonetheless, masked autoregressive flows (MAF) [189] address these limitations by constructing a stack of MADEs. Then, the output of each MADE is taken as input for the next, with random reordering of the chain rule factorization between each MADE. Then, MAFs define the following density estimator

$$\begin{aligned} p\big(\vec{t}\,\big|\vec{\theta}\,;\vec{w}\big) &= \prod_i p\big(t_i\big|\vec{t}_{1:i-1},\vec{\theta}\,;\vec{w}\big) \\ &= \mathcal{N}\left(\vec{u}\big(\vec{t},\vec{\theta}\,;\vec{w}\big)\Big|\mathbb{0},\mathbb{1}\right)\prod_{n=1}^{N_{\mathrm{MADEs}}}\prod_{i=1}^{\dim\vec{t}}\sigma_i^n\big(\vec{t},\vec{\theta}\,;\vec{w}\big)\,, \end{aligned} \tag{5.19}$$

where $\vec{u}$ is the output from the final MADE.

In order to fit a neural density estimator to a set of simulated samples $\{\vec{\theta},\vec{t}\,\}$, we need to find the weights of the neural network that minimize the Kullback-Leibler divergence

$$D_{KL}\left(p^*|p\right) = \int p^*\big(\vec{t}\,\big|\vec{\theta}\big)\ln\left(\frac{p\big(\vec{t}\,\big|\vec{\theta}\,;\vec{w}\big)}{p^*\big(\vec{t}\,\big|\vec{\theta}\big)}\right)\mathrm{d}\vec{t}, \tag{5.20}$$

between the parametric density estimator $p\big(\vec{t}\,\big|\vec{\theta}\,;\vec{w}\big)$ and the target $p^*\big(\vec{t}\,\big|\vec{\theta}\big)$. Since we have not access to the target density, we use only the samples $\{\vec{\theta},\vec{t}\,\}$ to take the -negative logarithmic- loss function, as

$$-\ln U\left(\vec{w}\,\Big|\{\vec{\theta},\vec{t}\,\}\right) = -\sum_{i=1}^{N_{\mathrm{samples}}}\ln p\big(\vec{t}_i\,\big|\vec{\theta}_i\,;\vec{w}\big)\,, \tag{5.21}$$

which is a Monte Carlo estimation of the KL-divergence, and it is also equivalent to the negative logarithmic-likelihood of the simulated data $\{\vec{\theta},\vec{t}\,\}$ under the conditional density estimator $p\big(\vec{t}\,\big|\vec{\theta}\,;\vec{w}\big)$. For MAF conditional estimators, the loss is given by

$$-\ln U\left(\vec{w}\,\Big|\{\vec{\theta},\vec{t}\,\}\right) = -\sum_i \ln\left[\mathcal{N}\left(\vec{u}\big(\vec{t}_i,\vec{\theta}_i\,;\vec{w}\big)\Big|\mathbb{0},\mathbb{1}\right) + \sum_{n=1}^{N_{\mathrm{MADEs}}}\sum_{m=1}^{\dim\vec{t}}\ln\sigma_m^n\big(\vec{t}_i,\vec{\theta}_i\,;\vec{w}\big)\right]\,. \tag{5.22}$$

Therefore, the neural density estimators are trained in the usual way by minimizing the negative log-loss with respect to the network weights or inferring a posterior density over the weights given the training data (and some network weight prior). The over-fitting can be mitigated using the standard regularization methods used for neural networks, such as early-stopping or dropout.

In some cases, perform the forward simulations is expensive. Then, the goal is to achieve the highest fidelity posterior inference with the fewest simulations possible. Preferentially, we would like to run simulations in the most exciting regions of the parameter space, which is not known *a priori*. Active learning allows the neural density estimators to call the simulator independently during the training, automatically deciding on-the-fly where the best parameters to run new simulations are, based on their current state of knowledge of the target posterior. Some of the active learning approaches for adaptive acquisition of simulations for DELFI are the sequential neural likelihood [190] and the Bayesian optimization style acquisition rules [159].

Chapter 6

The Correlation Function For The NYU-VAGC

Laying the foundations of Cosmological theory, in chapters 1 to 3 we reviewed from the basics, the principal's quantities that make up the ΛCDM scenario. In chapter 4, we saw the principal observable that we will seek in this thesis and how to obtain it through the observation of galaxies distribution. We finalized, in chapter 5, by studying some of the smart and new techniques in the last decades that help us to construct that observable and infer cosmological information from it. In this chapter, we will apply those techniques to construct the correlation function from the New York University Value-Added Galaxy Catalog (NYU-VAGC) [7, 45, 186], which is a cross-matched collection of galaxy catalogs with spectroscopic information.

6.1 NYU Value-Added Galaxy Catalog

The NYU-VAGC combine galaxies information from several catalogs, which include the Sloan Digital Sky Survey (SDSS [268]), the Faint Images of the Radio Sky at Twenty-centimeters survey (FIRST [33]), the Two-Micron All-Sky Survey (2MASS [230]), the Two-degree Field Galaxy Redshift Survey (2dFGRS [61]), the Infrared Astronomical Satellite Point Source Catalog Redshift (IRAS PSCz [220]) and the Third Reference Catalog of Galaxies (RC3 [78]). Together, they create a robust spectroscopic survey at low redshift ($0.003 < z < 0.05$) with $559,391$ galaxies shown in figs. 6.1 and 6.2. Besides, the NYU-VAGC provides a random mask that covers the same area that the galaxy catalog, which corresponds to j. This mask is formed with $1,905,377$ points and their sky coordinates, but without redshift information. It is also shown in fig. 6.1.

6.2 Kernel Density Estimation For Random Mask

We proceed to estimate the redshift distribution from galaxies to apply it to the random points. We use the kernel density estimation technique, explained in section 5.2, with a Gaussian kernel from eq. (5.2). For the estimation of the hyper-parameter h, we used $10-$fold

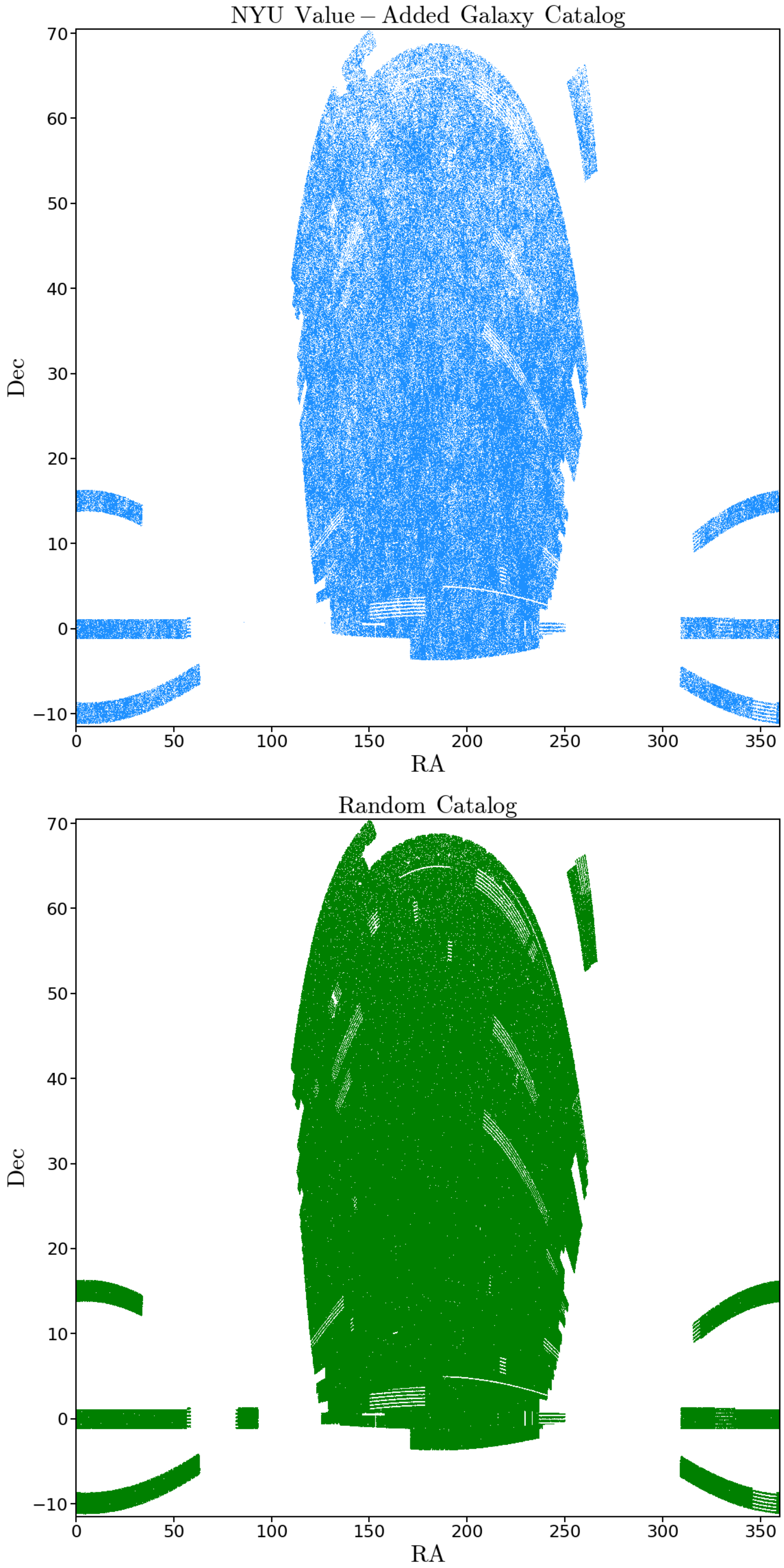


Figure 6.1: *Top.* The NYU-VAGC galaxies showing its sky covering. *Bottom.* The mask of random points Poisson distributed over the same geometry and area than NYU-VAGC.

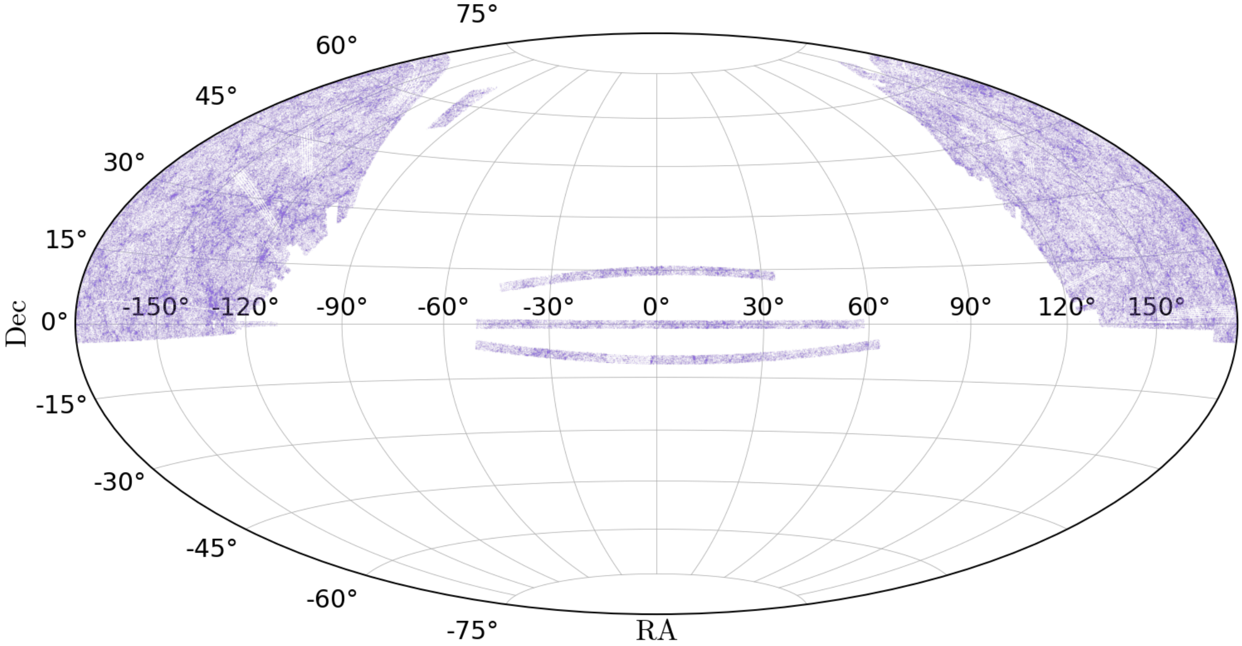


Figure 6.2: The NYU-VAGC sky covering using the Hammer projection.

cross-validation *i.e.* we cut randomly our dataset into 10 subsets. Thus, we take 9 subsets that we call train set to obtain the redshift probability distribution function to finally compare their samples in the remaining subset or test set. We do this procedure for several values of h, and also iterate and change the test set over the 10 partitions of our dataset. We used a grid of $h-$values that contain different magnitude orders, which is essential in this kind of analysis. It is given by

$$h_{grid} = [0.00001, 0.00005, 0.00008, 0.001, 0.003, 0.005, 0.01, 0.05] \,. \tag{6.1}$$

The best bandwidth, which minimizes the sampler and real redshift over the ten iterations, is given by $h = 0.001$. Using it, we create the samples for the randoms points from with the seed 51294. These are shown in fig. 6.3 with the NYU-VAGC redshift distribution. In the same image, we can observe the KDE probability distribution function together with the both samples, which also shows their agreement.

6.3 From Redshift To Comoving Distances

Once the redshift distribution for randoms is obtained, we can compute the comoving distance between the targets from catalogs using eq. (1.53). For that, we need to assume some fiducial cosmology, which, in our case, is Planck 2018 cosmology [59]. We use the cosmological parameters obtained by the constraining of the $TT, TE, EE + \text{low}E + \text{lensing}$ signals, where we use the mean values given at the 68% confidence intervals. Besides, we measure

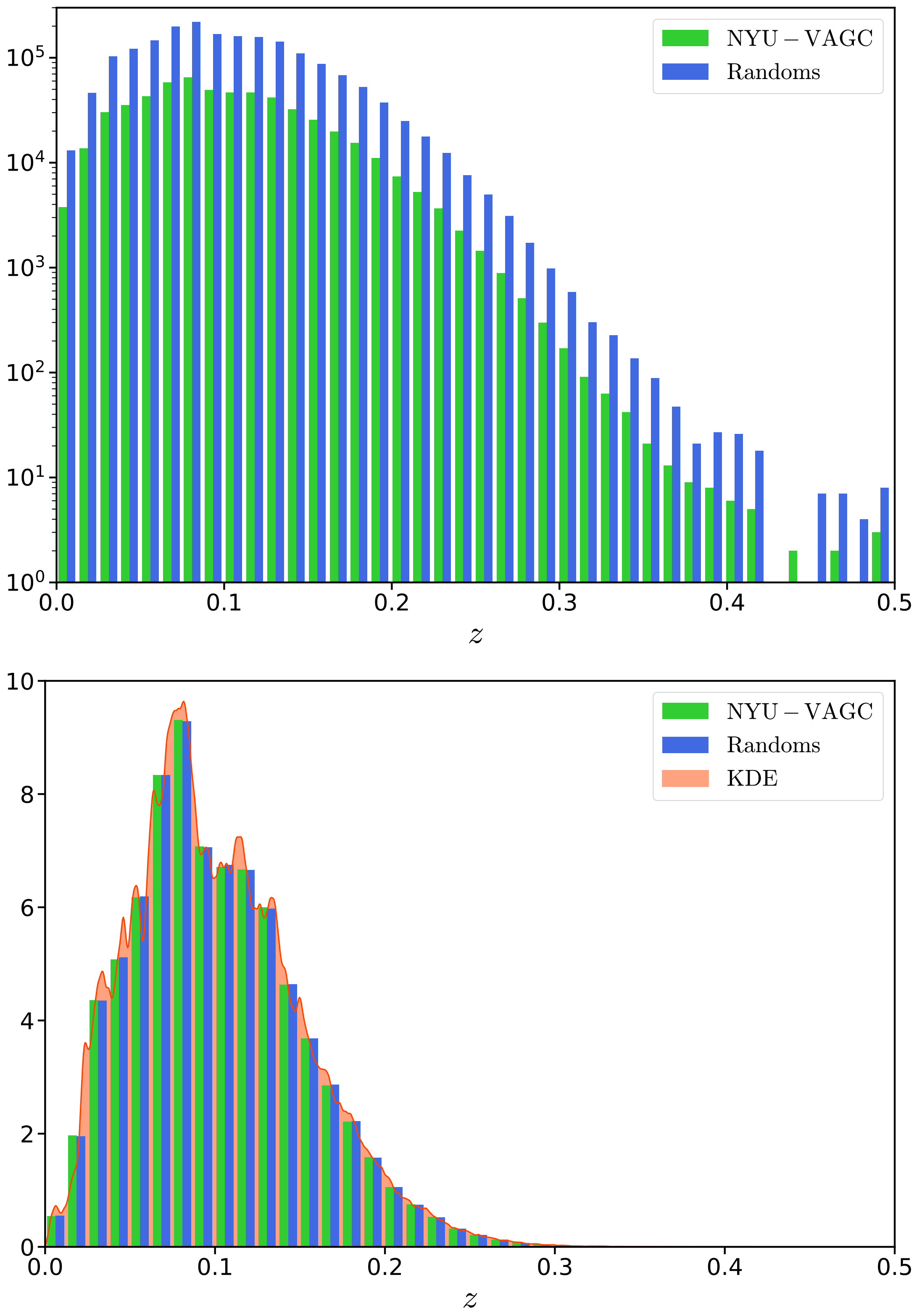


Figure 6.3: *Top.* The histogram with a logarithm $y-$axis scale shows the number of galaxies as a function of redshift for the NYU-VAGC (green) and the samples obtained from the KDE technique for the random points (blue). *Bottom.* The normalized histogram with a linear $y-$axis scale with the NYU-VAGC galaxy redshifts (green), the redshift samples obtained from KDE for randoms (blue), and the KDE probability distribution function (orange).

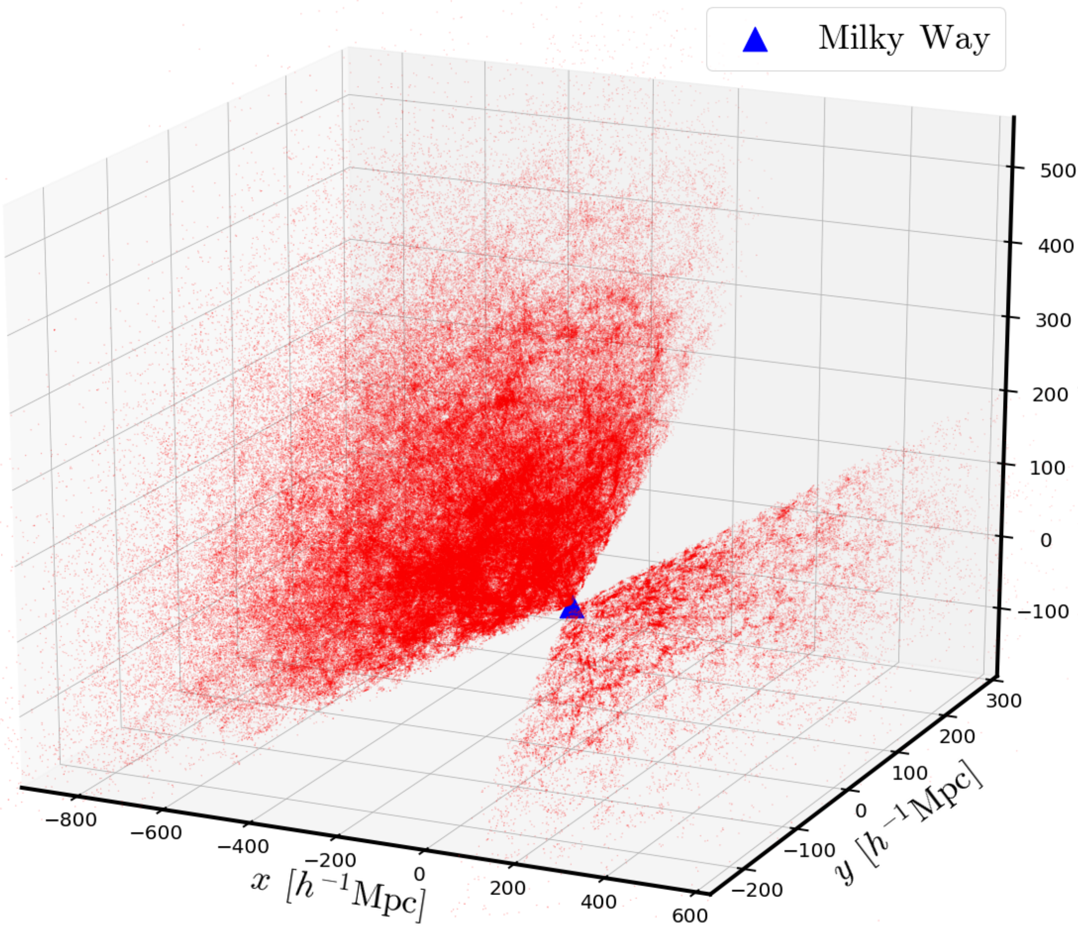


Figure 6.4: The distribution of galaxies in redshift-real space from the NYU-VAGC, assuming Planck 2018 [59] as fiducial cosmology. The blue triangle indicates our position in the origin of the coordinate reference system of eq. (6.3).

those distances in h^{-1}Mpc, with h the dimensionless reduced Hubble parameter defined in section 1.1. Therefore, assuming a flat ΛCDM scenario, we only need to fix the parameter $\Omega_{m,0}$, which is

$$\Omega_{m,0} = 0.31\,, \tag{6.2}$$

because H_0 is encoded in h which is free, Ω_r is negligible at low redshift, and $\Omega_{\Lambda,0} = 1 - \Omega_{m,0}$ by eq. (1.34). Finally, we can construct a 3D map for the NYU-VAGC and the random catalog, where the first is shown in fig. 6.4. For that, we take the next coordinate transformation

$$\begin{aligned} x &= \chi \sin(\delta) \cos(\alpha) \\ y &= \chi \sin(\delta) \sin(\alpha) \\ z &= \chi \cos(\delta)\;, \end{aligned} \tag{6.3}$$

where χ is the comoving distance, δ the declination, and α the right ascension.

6.4 Weight Scheme

For the weight scheme discussed in section 4.3, we are going to use the FKP weights from eq. (4.79). Then, we need the number density and the power spectrum at some interesting scale. For the first, we split the NYU-VAGC into 224 redshift bins, where each one contains approximately 2500 galaxies. In this way, we can obtain a better measurement of the number density, in which we need to compute the volume occupied by each bin. Nonetheless, this partition generates bins that are not linear, neither logarithmic.

To obtain the bin volume, we are going to use spherical coordinates. Then, the volume differential is $\mathrm{d}V = \chi^2 \mathrm{d}\chi\Omega$, where Ω is the solid angle differential. For a bin between z_i and z_{i+1}, the volume is

$$V_i = \Omega_{survey} \frac{\chi^3 \left(z_{i+1}\right) - \chi^3 \left(z_i\right)}{3} \,, \tag{6.4}$$

with $\Omega_{survey} = 2.44$ as the solid angle occupied by the survey, which we assume is equal for all bins because the partition is in redshift space but not in spherical angles. The comoving bin limits are obtained, as before, using the fiducial cosmology. Thus, we can obtain the number density as

$$\bar{n}_i = \frac{n_i}{V_i} \,, \tag{6.5}$$

where $n_i = 2500$ for all bins except the last one, which contains $n_{224} = 1891$ galaxies.

Now, we need to assign a weight for each galaxy, but first, we have to know what is the bin to each galaxy. This is not difficult, since we know the galaxy redshift and the bin limits in redshift. For the randoms points, we use the same number density function that galaxies because they have the same redshift distribution.

Finally, we need to set the power spectrum at some interesting scale. We use three values $P(k) = 10000,\ 15000,\ 20000$ [Mpc^3], obtaining three different weights. Note that we did not use $h^{-3}\mathrm{Mpc}^3$ as a unit because we do not want to deal with the h measurement in the weights. Under the fiducial cosmology [59] at $z = 0$, the above power spectrums correspond to the scales $k = 0.1032,\ 0.0786,\ 0.0642$ [Mpc^{-1}], respectively. We will see the results for the correlation function estimator in the next section, but it is important to remark that all the scales aforementioned are close to BAO.

6.5 The NYU-VAGC Two-Point Correlation Function

With the comoving distances and weights at both catalogs, we have all the ingredients to construct our two-point correlation function estimator, as we saw in section 4.3. We need to compute the DD and RR pairwise distances in each catalog and the DR cross-pairwise distance. Then, we group them by bins to obtain the estimator. We are going to use a ball tree algorithm, explained in section 5.1, already implemented in `Python` by Jarvis, Bernstein, & Jain (2004) [135] through a library called `TreeCorr`.

Once computed the pairwise distances, with the normalization explained in eqs. (4.69) to (4.71), and the weight scheme from eq. (4.81), we need to group them in bins. We choose two binning schemes: linear and logarithmic. For the first we take 20 bins between $r = 20$

$[h^{-1}\text{Mpc}]$ to $r = 160$ $[h^{-1}\text{Mpc}]$, which makes the bin size $\Delta r = 7.5$ $[h^{-1}\text{Mpc}]$. For the second we take 45 bins between $r = 15$ $[h^{-1}\text{Mpc}]$ to $r = 180$ $[h^{-1}\text{Mpc}]$, which makes the logarithm bin size $\Delta \ln(r) = 0.024$ $[\ln(h^{-1}\text{Mpc})]$.

The binning is not a trivial election, since, for most correlations functions, which tend to be approximately power laws, the logarithmic binning is the most appropriate. Nevertheless, for uses cases where the scales of interest span only a relatively small range of distances, it may be convenient to use linear binning. This last choice is the preferred one for BAO when we study it on large scales. Besides, there is a technical problem, due to the implementation and search in ball trees, the logarithmic binning is faster than the linear. In `TreeCorr`, the speed is also parameterized by the `brute` and the `bin_slop` arguments. The first is a Boolean that, if it is false, the algorithm stops at non-leaf cells whenever the error in the separation is compatible with the given `bin_slop`. But if it is true, we use the brute force algorithm, which goes to the leaves. The second can be defined as how much slop allows in the placement of pairs in the bins. Therefore, using the brute force algorithm takes more time since the ball tree structure is no longer useful. Similarly, using a `bin_slop` equal or close to 0 takes more time since we must perform a search tree in all the branches. For our case, we use `bin_slop`$= 0.1$ for the logarithmic binning and `bin_slop`$= 0.1$ for the linear.

Finally, we construct the two-point correlation function using the Peebles & Hauser (eq. (4.72)) and the Landy & Szalay (eq. (4.76)) estimators. Since the second also requires cross-pairwise distances, it takes more computation time. The results are shown in fig. 6.5.

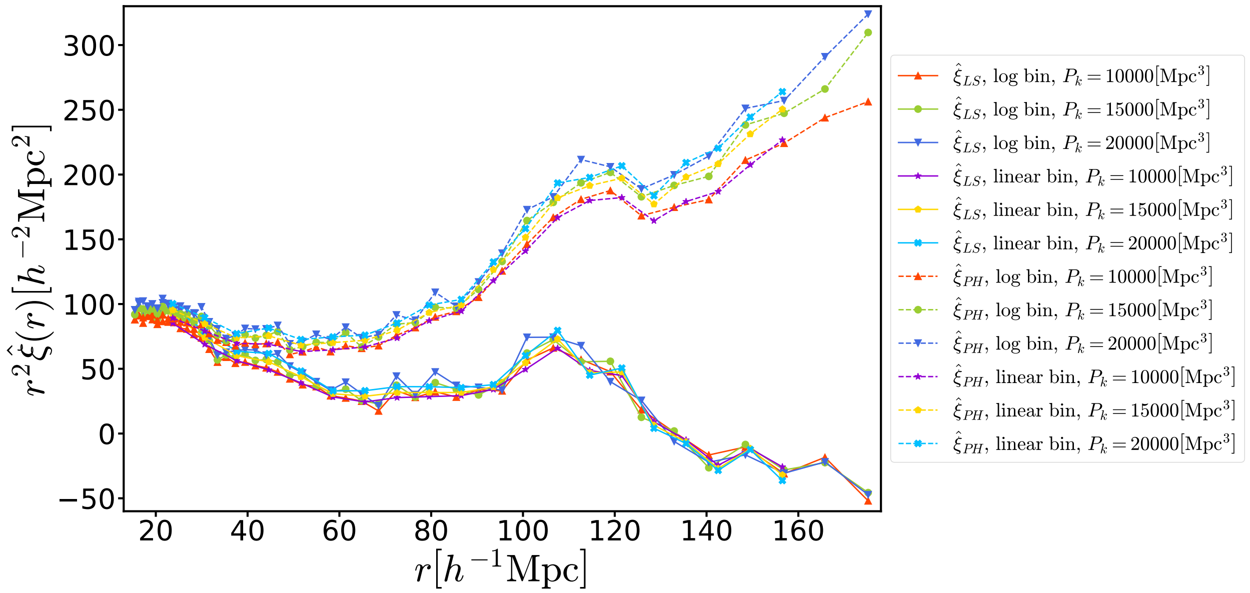


Figure 6.5: The two-point correlation signal for the NYU-VAGC using three different power spectrums in the weighting scheme, as we argued in section 6.4. We plot the $r^2\hat{\xi}$ signal to better distinguish the expected BAO peak near 110 $[h^{-1}\text{Mpc}]$. Finally, we show the signal computed using the Landy & Szalay estimator $\hat{\xi}_{LS}$ from eq. (4.76) (solid lines) and the Peebles & Hauser estimator $\hat{\xi}_{PH}$ from eq. (4.72) (dashed lines). For both cases, we display the linear and logarithmic binning used and detailed in section 6.5.

6.6 The Covariance Matrix For The NYU-VAGC

Our two-point correlation estimator also requires an uncertainty measurement to make Bayesian inference in the next section. There are several ways to obtain it, but that does not mean that all are useful. For example, we can consider that all the galaxies were taken correctly in the survey, and then, their positions and redshifts have no uncertainty; also, there are no other possible systematic errors like the induced by astrophysical processes. In that case, the only source of uncertainty is the shot noise, which for the two-point estimator is the Poisson error due to the counting of galaxies. It has the next form [94]

$$\hat{\sigma}^2\left[\hat{\xi}(r)\right] = \frac{\left(1+\hat{\xi}(r)\right)^2}{DD(r)}, \tag{6.6}$$

which depends on the two-point correlation estimator and the galaxies' pairwise distances for a certain bin. It is shown in fig. 6.6 for the signal computed in the last section. Therefore, with this method, the covariance matrix will be diagonal, and there is no way to estimate the off-diagonal terms. However, this method is based on unrealistic assumptions, and it underestimates the variance.

To take a better estimator of the covariance matrix, we perform a jackknife method. In this case, we split the sky coverage of the survey into $N_{patch} = 40$ patches shown in fig. 6.7. We will also estimate the two-point correlation function using the Landy & Szalay estimator, the logarithmic binning detailed in the last section, and a weighting scheme based on the

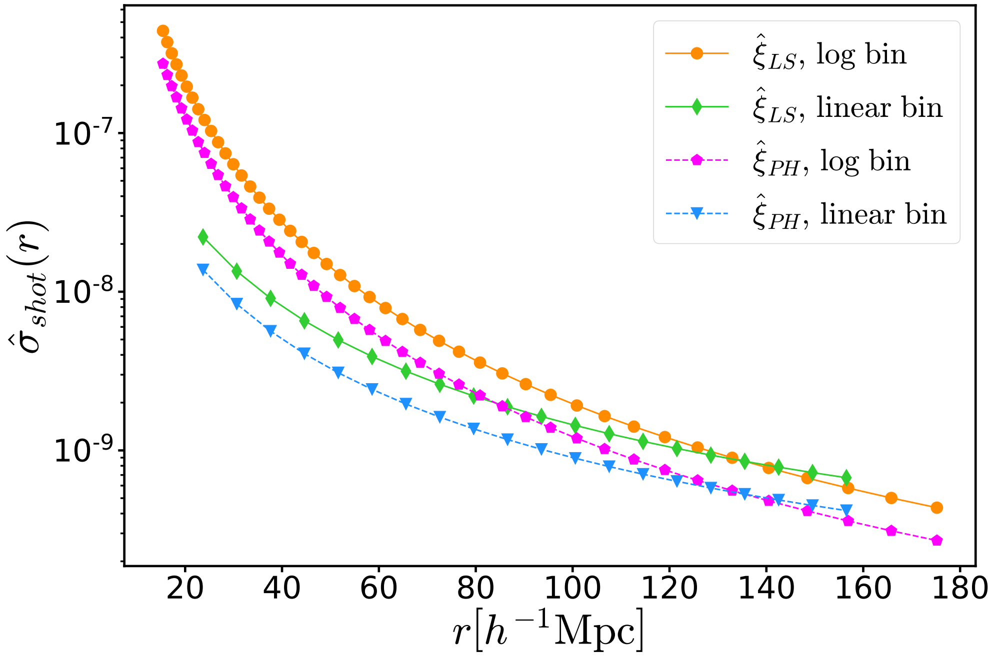


Figure 6.6: The shot noise for the NYU-VAGC computed by eq. (6.6), using only one weight scheme with $P(k) = 15000$ $\left[\text{Mpc}^3\right]$ since there is not a big difference between the three power spectrum used. We show the error for the Landy & Szalay estimator $\hat{\xi}_{LS}$ from eq. (4.76) (solid lines) and for the Peebles & Hauser estimator $\hat{\xi}_{PH}$ from eq. (4.72) (dashed lines). In both cases, we display the linear and logarithmic binning used and detailed in section 6.5.

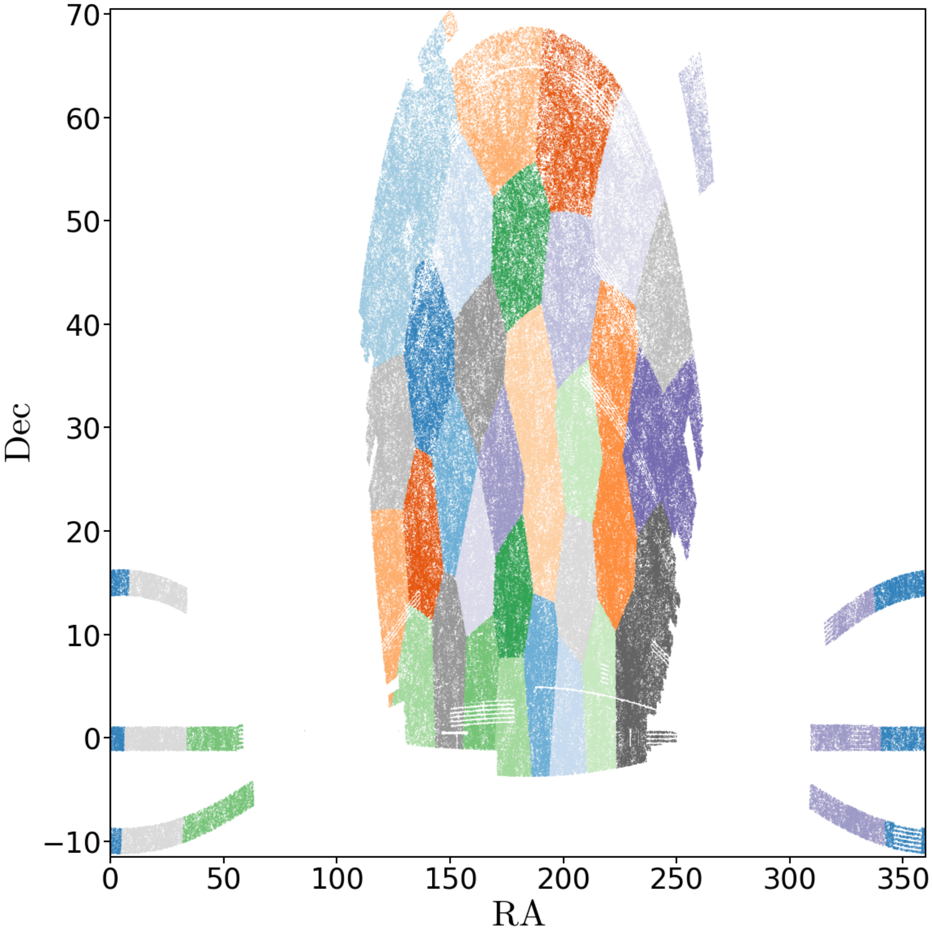


Figure 6.7: The $N_{patch} = 40$ sky patches of the NYU-VAG that were used to compute the variance of two-point correlation signal through the jackknife technique. They were obtained using the $k-$means algorithm from `scikit-learn`.

power spectrum $P(k) = 15000$ [Mpc^3]. These signals are shown in fig. 6.8. Then, we estimate the covariance matrix based on the scatter in the measurement by excluding one patch at a time. It is given by

$$C_{ij} = \frac{N_{patch} - 1}{N_{patch}} \sum_{a=1}^{N_{patch}} \left(\hat{\xi}_a\left(r_i\right) - \bar{\xi}\left(r_i\right)\right)\left(\hat{\xi}_a\left(r_j\right) - \bar{\xi}\left(r_j\right)\right) , \tag{6.7}$$

where $\bar{\xi}$ is the mean two-point correlation function between all the patches.

Finally, we also propagate the shot noise into the covariance matrix (see fig. 6.9) obtained from jackknife, which does not induce significant variance. Nonetheless, it is essential to remark that the size of the sky patches and their number can considerably affect the determination of the uncertainty on scales of interest. However, they are restricted by the survey volume and our incapability to compute ensemble averages, as we argued in section 4.1. We can try to remedy that by using simulations that require initial conditions and modeling to proportionate mock catalogs, from which we can compute ensemble averages. This method is used mostly because it gives a better error estimation, but it requires a considerable number of mocks, which is expensive, and good modeling on different scales that are likely affected by non-linear effects. Besides, those simulations are affected by cosmological and astrophysical parameters in processes that are not clear yet, turning it into an active research field.

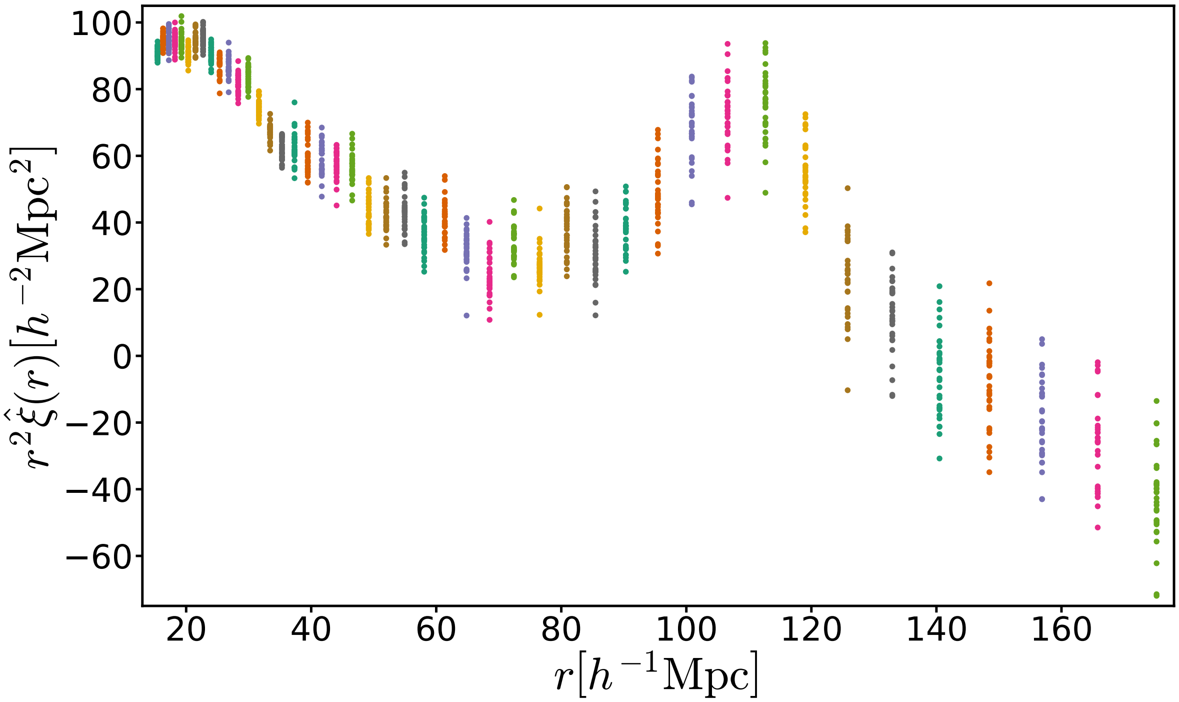


Figure 6.8: The variability in the two-point correlation signal computed through jackknife technique for $N_{patch} = 40$ sky patches of the NYU-VAGC coverage (see fig. 6.7) and detailed in section 6.6. We plot the $r^2\hat{\xi}$ signal to better distinguish the expected BAO peak near 110 $[h^{-1}\mathrm{Mpc}]$.

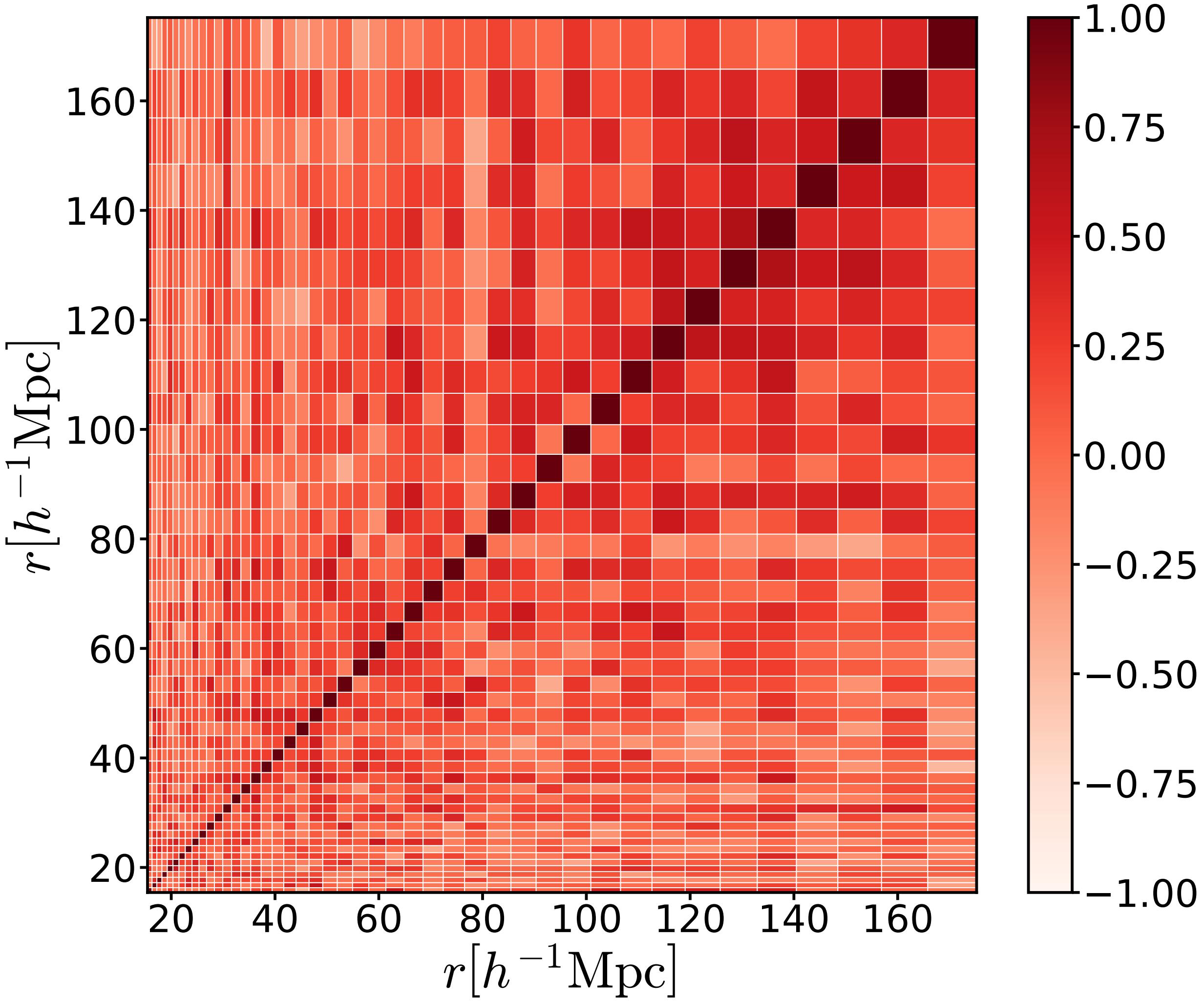


Figure 6.9: The normalized covariance matrix obtained by jackknife technique (see section 6.6) for $N_{patch} = 40$ sky patches of the NYU-VAGC coverage. The normalization is computed as $C_{ij}^{(norm)} = C_{ij}/\sqrt{C_{ii}C_{jj}}$, where C_{ij} is obtained from eq. (6.7). Note that its form is because we are using a logarithmic binning.

Chapter 7

Cosmological Inference

Once constructed the two-point correlation function and its covariance matrix, we can proceed to make a statistical analysis using the Bayesian inference framework of section 5.3. From it, we can recover cosmological information encoded in the two-point correlation function using the power spectrum from chapter 4, and considering in the model the distortions studied in the same chapter. These are important for our catalog because it is likely to suffer redshift space distortion since the proper velocities are comparable with the expansion rates at low redshift. Besides, to compute the comoving distances, we assumed a fiducial cosmology, which can differ from the encoded in the catalog, introducing geometrical distortions.

In this chapter, we will review the classic and new techniques explained in chapter 5 to constrain cosmological information. Since we already have the correlation function and its covariance matrix from chapter 6, we only need to specify the model. We are going to review some of the most used in the literature. Finally, we are going to compare the inference obtained from the different algorithms.

7.1 Models For $\xi(r)$

It is important to remark that the inference over cosmological parameters is restricted to a model that we need to specify. It is important to take into account the physical model and always consider the possible distortions in the measurement. We start using the power spectrum from eq. (4.50), which gives us information on all scales in the late Universe. However, that power spectrum is not always valid since we constructed it using linear theory, and we know that the model fails on small scales because of the non-linearities. Besides, we will see that those non-linear effects have a small contribution to large scales. From eq. (4.50), we can recover the correlation function by eq. (4.34) without considering anisotropies. This integration is performed numerically in logarithmic space to avoid the loss of information from BAO wiggles. Thus, we do the next change of variables $u = \log(kr) \Rightarrow \mathrm{d}u = \mathrm{d}k/k$, which transforms the integral into

$$\xi(r) = \int_{-\infty}^{\infty} \frac{\mathrm{d}u}{2\pi^2}\, \mathrm{e}^{2u} \frac{P(u)}{r^3} \sin\left(\mathrm{e}^{u}\right) . \tag{7.1}$$

This change preserves the cancellation of units between k and r required to obtain a dimensionless two-point correlation function. Then, those variables must have the same units. This issue can be looked trivial, but it is necessary to work with the explicit variable h in the units because it encodes cosmological information that can enter a systematic error. Since the power spectrum decays to zero for $k \to 0$ and $k \to \infty$, we do not need to perform the integral for all the domain, which is difficult because we are doing a numerical integration. Thus, we need to determinate some limits of integration, which can be found studying where $k^2 P(k)/r \sim 1$, because it gives us the scales k that do not contribute considerably to the integral beyond the peak for any r. Once obtained these limits k_- and k_+, we can perform the numerical integration in logarithm space of eq. (7.1) as

$$\xi(r) = \int_{\log(k_- r)}^{\log(k_+ r)} \frac{\mathrm{d}u}{2\pi^2}\, \mathrm{e}^{2u} \frac{P(u)}{r^3} \sin\left(\mathrm{e}^u\right) \,, \tag{7.2}$$

noting that we included the r value in the limits because of the change of variables. Also, the units of volume of the power spectrum are canceled by those in r^3, which is why they must have the same measurement units.

7.1.1 Purely Cosmological Model

Using eq. (7.2), we can start to make inference using the data constructed in chapter 6. Since the model for $\xi(r)$ depends on one power spectrum, which is produced by the cosmological parameters $\Omega_b, \Omega_{\mathrm{cdm}}, A_s, n_s, h, \tau_{op}, N_{eff}$ and the eq. (4.50), we can move on that parameter space. However, this introduces two big problems, which the first is that this has not much sense because we constructed the estimator using a fiducial cosmology [59], where we set the Ω_m parameter. Therefore, the inference that we could make in the parameter-space above is biased even if we model the geometrical distortions well. The second problem is that the movement in that parameter space requires, at each step, the computation of one power spectrum, which requires the solving of several coupled perturbed equations studied in chapters 2 and 3. This problem makes it extremely expensive even for algorithms based on Markov Chain Monte Carlo (MCMC).

7.1.2 Geometrical Model With Shift Parameter

In section 4.5, we studied how to take account of the differences between the fiducial and real cosmology. We saw that the principal effects are the distortions induced, which can be solved by the shift parameter α of eq. (4.123), and the warping ϵ. Under this model, we can compare the two-point correlation function constructed from data using the estimator with the theoretical one, which is produced using the power spectrum at the fiducial cosmology [59], as long as we consider the dilation by introducing the shift parameter in the model.

To face the problems of the last section, the biasing and the expensive computations, we can compare the estimator from chapter 6 with a geometrical model, which is constructed using the two-point correlation function through the power spectrum (see eq. (7.2)) at the fiducial cosmology [59]. By the moment, this would give us a framework from which we can not extract cosmological information, and also it preserves a biased inference. Nevertheless, since we know how to take account of the geometrical distortions influenced by cosmology, we can recover some of the cosmological parameters. Then, these kinds of models are called geometrical because we use the fact that a wrong assumption on the cosmological parameters affects the geometrical distribution of galaxies, which is observable in their two-point correlation function. To take account of the differences between the fiducial and real cosmology, we introduced the shift parameter α (see eq. (10.18)) in the model, assuming $\epsilon = 0$ at first instance *i.e.* without considering anisotropies. Then, the model is

$$\xi_{\rm model}(r, z = z_{\text{🌀}}) = B\xi_{th}\left(\alpha r, z, \vec{p}_r\right) \, , \tag{7.3}$$

where ξ_{th} is the two-point correlation function obtained using the fiducial or reference cosmology [59] parameters $\vec{p}_r$ at some redshift z, which in this case corresponds to the mean galaxy redshift of the catalog $z_{\text{🌀}} = 0.101$; while, B takes account of possible differences in the amplitude of the signal. These will be clear through the use of other models.

7.1.3 Geometrical Model With Nuisance Parameters

We also can take account of possible systematic errors by using nuisance parameters. These enter the model as measurements of the different functions that can compete against the theoretical model of the two-point correlation function. Since we expect that the signal decreases as a power-law r^{-2}, we will insert nuisance parameters until that order. Then, the model is the sum of eq. (7.3) and the power-law functions

$$\xi_{\rm model}(r, z = z_{\text{🌀}}) = B\xi_{th}\left(\alpha r, z, \vec{p}_r\right) + a_0 + \frac{a_1}{r} + \frac{a_2}{r^2} \, , \tag{7.4}$$

where the a_i are the nuisance parameters.

7.1.4 Geometrical Model Considering Non-Linearities

The non-linear structure formation can affect acoustic oscillations' signature because the motions that lead to the growth of correlations also change their shape. The mode-coupling effect can also explain this in the power spectrum, which generates additional oscillations out of phase with those in the linear spectrum. Then, when it is Fourier transformed, these out of phase oscillations induce a percent-level shift in the acoustic peak of the two-point correlation function [66, 244]. To take account of the non-linear suppression of the BAO signal, we can introduce a damping term in the matter power spectrum

$$P_{\rm damped}(k, z) = P_m(k, z)\mathrm{e}^{-(k/k_*)^2} \, , \tag{7.5}$$

where P_m is the spectrum from eq. (4.50) and k_* is the damping scale, which can also be written as

$$P_{\rm damped}(k, z) = P_m(k, z)\mathrm{e}^{-(\sigma_v k)^2} \, , \tag{7.6}$$

with $\sigma_v = 1/k_*$ as the pairwise velocity dispersion, which was also defined in eq. (4.101) because it encodes the same effect. It also depends on redshift, and it is important to consider the mean galaxy redshift of the catalog if we want to set/constrain it. Then, the damped two-point correlation function is obtained through eq. (4.34), which is denoted as

$$\xi_{\rm damped}(r, z) = \xi_L(r, z) \otimes G(r) \, , \tag{7.7}$$

where $\otimes$ term denotes convolution, $\xi_L(r,z)$ is the two-point correlation function obtained from the linear matter power spectrum and $G(r)$ is the Fourier transform of $\tilde{G}(k)$ or damped function, which was implicitly defined before as

$$\tilde{G}(k) = \mathrm{e}^{-(\sigma_v k)^2} = \mathrm{e}^{-(k/k_*)^2} \,. \tag{7.8}$$

Then, for the convolution, we just multiply the power spectrum and the damped function in Fourier space to obtain the damped two-point correlation function through eq. (4.101).

Also, to take account of non-linearities to second-order caused by the mode-coupling of different Fourier modes, we can include an additional term to the model

$$\xi_{NL}(r,z) = \xi_L(r,z) \otimes G(r) + A_{MC}\xi^{(1)}(r,z)\frac{\partial \xi_L(r,z)}{\partial r} \,, \tag{7.9}$$

where the first term damps the baryonic acoustic feature in the linear two-point correlation function ξ_L, as we saw, A_{MC} is the $k-$mode coupling (MC) parameter, and $\xi^{(1)}$ is

$$\xi^{(1)}(r,z) = \hat{r}\cdot\nabla^{-1}\xi_L(r,z) = \int_0^\infty \frac{k}{2\pi^2}P_m(k,z)j_1(kr)\mathrm{d}k \,, \tag{7.10}$$

where $P_m(k,z)$ comes from eq. (4.34) and $j_1(kr)$ is the spherical Bessel function of first order.

In most of the surveys, the parameter k_* or σ_v is constrained, even when we can determinate it by eq. (4.101). On the other hand, the parameter A_{MC} has a low power of constraining because at high redshift, the mode coupling effect is small, and at low redshift, it is negligible compared to RSD. Therefore, in most of the analyses done, it is set to $A_{MC} = 0$ or $A_{MC} = 1$.

The inclusion of the damped scale affects the small-scale clustering amplitude, but we know that it should affect only the BAO amplitude. This can be solved adding a term that includes a wiggle-free power spectrum $P_{\mathrm{dw}}(k,z)$ [244], which possesses the same shape as $P_m(k,z)$ but with the baryon oscillation component deleted, through

$$\begin{aligned} P_{\mathrm{damped}}(k,z) &= \tilde{G}(k)P_m(k,z) + \left[1-\tilde{G}(k)\right]P_{\mathrm{dw}}(k,z) \\ &= \left[P_m(k,z) - P_{\mathrm{dw}}(k,z)\right]\exp\left(-\frac{1}{2}k^2\Sigma_{NL}^2\right) + P_{\mathrm{dw}}(k,z) \,, \end{aligned} \tag{7.11}$$

where in the second line we reordered the terms introducing the same damped function but now with the variable Σ_{NL}, which is just

$$\Sigma_{NL}^2 = 2\sigma_v^2 = \frac{2}{k_*^2} \,, \tag{7.12}$$

because it is most used in this kind of analysis.

Finally, the model considering the shift, nuisance parameters, and non-linearities is

$$\xi_{\mathrm{model}}(r, z = z_{\text{🌀}}) = B\xi_{NL}\left(\alpha r, z, \vec{p}_r | \Sigma_{NL}, A_{MC}\right) + a_0 + \frac{a_1}{r} + \frac{a_2}{r^2} \,, \tag{7.13}$$

with ξ_{NL} from eq. (7.9) or the obtained after Fourier transforming one of the damped power spectrum from eqs. (7.5), (7.6) and (7.11).

7.1.5 Geometrical Model With Redshift Space Distortions

In the last models, we grouped all the effects that could change the amplitude of the signal in the B parameter. Since we already know that the RSD have an impact on the amplitude from section 4.4, we are going to model them using the multipoles. We will use the monopole because it requires the same integration that the isotropic two-point correlation function. Inserting the Kaiser formula from eq. (4.110), because we are going to consider also the bias, into eq. (4.107) for $l = 0$, we obtain

$$P_0^s(k, z) = \left(b^2 + \frac{2}{3} f b + \frac{1}{5} f^2\right) P_m(k, z) \,. \tag{7.14}$$

Before computing the monopole, we will normalize the power spectrum, as we did in eq. (4.55).

$$P_0^s(k, z) = \left((\sigma_8 b)^2 + \frac{2}{3}(\sigma_8 f)(\sigma_8 b) + \frac{1}{5}(\sigma_8 f)^2\right) P_m(k) \,, \tag{7.15}$$

noting that we introduced the variance of the mass density at $8h^{-1}Mpc$ together the b and f parameters. Besides, the z dependence is lost in the power spectrum. Therefore, the monopole is

$$\xi_0^s(r) = \left((\sigma_8 b)^2 + \frac{2}{3}(\sigma_8 f)(\sigma_8 b) + \frac{1}{5}(\sigma_8 f)^2\right) \Xi(r) \,, \tag{7.16}$$

where $\Xi(r)$ is the normalized two-point correlation function *i.e.* the obtained after the integration of the $k-$ dependence only power spectrum.

We can combine this model with the previous, obtaining

$$\begin{aligned} \xi_{\text{model}}(r, z = z_{🌀}) = &\left((\sigma_8 b)^2 + \frac{2}{3}(\sigma_8 f)(\sigma_8 b) + \frac{1}{5}(\sigma_8 f)^2\right) \Xi\left(\alpha r, \vec{p}_r | \Sigma_{NL}, A_{MC}\right) \\ &+ a_0 + \frac{a_1}{r} + \frac{a_2}{r^2} \,, \end{aligned} \tag{7.17}$$

where we can normalize the damped power spectrum or the mode-coupling effect in the correlation function from eqs. (7.5), (7.6), (7.9) and (7.11) by the fiducial σ_8^2 at the mean galaxy redshift of the catalog to then obtain the normalized two-point correlation function.

It is common to constrain $f\sigma_8$ and $b\sigma_8$ in this kind of analysis instead σ_8, f, and b directly. Because that would generate a correlation between the three parameters. Nonetheless, the combined parameters are independent between each other, making them a robust measurement because a dependence of the bias is dangerous in galaxy catalogs that are, for example, considering different types of galaxies or tracers.

It is also important to remark that the measurements obtained for $f\sigma_8$ and $b\sigma_8$ depend on the fiducial cosmology chosen to construct the correlation function. We need to take account of the bias generated on these measurements due to geometrical distortions before using them in other analyses. The principal change is generated in the volume, which generates an anisotropy in the power spectrum described by the Alcock-Paczynski effect in [18] as we saw in section 4.5. This issue can be corrected by considering how the two-point correlation function transforms under a change of cosmology, which is treated in [16]. From which we start relating both power spectrum amplitudes

$$\sigma_8^2 b^2 \left[1+\beta\mu^2\right]^2 = \frac{H_r(z)\mathrm{d}_A^2(z)2}{H(z)\mathrm{d}_{A_r}^2(z)}\sigma_{8_r}^2 b_r^2 \left[1+\beta_r\mu_r^2\right]^2 \mathcal{C}\,, \tag{7.18}$$

where r indicates that the variables are obtained using the fiducial cosmology, and $\mathcal{C}$ is a constant that relates the two power spectrum. Following the procedure of [16] to eliminate the line-of-sight direction variable μ and the bias b through approximations, we obtain the next relation

$$f\sigma_8 = \left(f\sigma_8\right)_r C \left(\frac{H(z)\mathrm{d}_{A_r}^2(z)}{H_r(z)\mathrm{d}_A^2(z)}\right)^{\frac{3}{2}} \left(\frac{\sigma_8}{\sigma_{8_r}}\right)^2 , \tag{7.19}$$

with C explicitly given as

$$C = \int_{k_1}^{k_2} \mathrm{d}k \sqrt{\frac{P_r(k)}{P(k')}}\,, \tag{7.20}$$

where k_1 and k_2 are the limits of the scales that we are interested in for the power spectrum data or the correlation function. The scale k' denotes the modes in the underlying cosmology and it obeys $k' = k/\alpha$, with α the shift parameter from eq. (4.123). However, C is close to 1 in most of the surveys, which allows us to ignore it.

Another way to take account of that effect is by merely rescaling the $f\sigma_8$ measurements with $H(z)\mathrm{d}_A(z)$ ratios from the cosmology used to that of the fiducial one as in [161]. Then, the measurements need to be corrected by

$$f\sigma_8 = \left(f\sigma_8\right)_r \frac{H_r(z)\mathrm{d}_{A_r}(z)}{H(z)\mathrm{d}_A(z)}\,. \tag{7.21}$$

Although, it is most common to correct the model because, in that case, we have not to change the possible covariance matrix that relates the measurements when we make Bayesian inference. Therefore, for the model, the correction is

$$\left(f\sigma_8\right)_{mc} = \left(f\sigma_8\right)_m \frac{H(z)\mathrm{d}_A(z)}{H_r(z)\mathrm{d}_{A_r}(z)}\,, \tag{7.22}$$

where mc indicates model-corrected, and m is the model value obtained through the evaluation at some point in the parameter space. It also determines the values of $H(z)$ and $\mathrm{d}_A(z)$. It is important to note that the value $H(z)\mathrm{d}_A(z)$ is independent of H_0, which is another reason to use this method.

7.1.6 Empirical Model

We also can construct an empirical model for the two-point correlation function, since we know *a priori* that it evolves as a power-law with a peak at BAO scales that can be modeled as a Gaussian function. We consider this kind of model from [245, 255] with the form

$$\xi_{\mathrm{model}}(r) = B + \left(\frac{r}{r_0}\right)^{-\gamma} + \frac{N}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(r-r_m\right)^2}{2\sigma^2}\right), \tag{7.23}$$

where the parameters r_0 and γ model the shape of the correlation at small scales and its power-law behavior, B takes into account a possible negative correlation at large scales, and r_m, σ and N are the parameters of the Gaussian function used to model the BAO feature. It is important to remark, that according to [255], the correct BAO peak position is shifted to smaller scales with respect to the Gaussian median value r_m.

The model in eq. (7.23) can be used to detect the BAO peak position accurately, but we can not extract directly any additional cosmological information.

7.2 Results I

To find the optimal parameter values, we use a Monte Carlo Markov Chain (MCMC) based on a Metropolis-Hastings algorithm (see section 5.3), which is developed in the `Python` library `emcee` [100]. We assume a likelihood function with the form

$$L = \frac{\exp\left(-\frac{1}{2}\chi^2(\vec{\theta})\right)}{\sqrt{(2\pi)^k \det\left(\hat{C}\right)}} \propto \exp\left(-\frac{\chi^2(\vec{\theta})}{2}\right), \tag{7.24}$$

where $\hat{C}$ is the covariance matrix estimator, k is its dimension, and also, we will use the second proportionality because the covariance matrix does not depend on the parameters $\vec{\theta}$.The χ^2 statistic is computed as

$$\chi^2(\vec{\theta}) = \left[\vec{\xi}_{\text{model}}(\vec{\theta}) - \vec{\xi}_{\text{obs}}\right]^T \hat{\Psi} \left[\vec{\xi}_{\text{model}}(\vec{\theta}) - \vec{\xi}_{\text{obs}}\right] . \tag{7.25}$$

In the equation above, $\vec{\xi}_{\text{model}}$ is the model that depends on the parameters $\vec{\theta}$; while, $\vec{\xi}_{\text{obs}}$ is the observed one. Both are grouped in a vector at each position r. Finally, $\hat{\Psi}$ is the precision matrix, which is the inverse of the covariance matrix. Nonetheless, it is important to remark that taking the inverse of the data covariance matrix, which is an estimator, induces an error that must be propagated. This issue is why the estimator for the precision matrix is not necessarily equal to the inverse estimator of the data covariance matrix. For our purposes, we will assume that $\hat{\Psi} = \hat{C}^{-1}$, but considering that this can induce a bias in our parameters to constrain. A better handle of this effect can be found in chapters 9 and 10.

Parameter	Flat Priors	Parameter Constraints
B	$[-5, 5]$	-0.0012 ± 0.0004
γ	$[0, 10]$	$2.9^{+0.257}_{-0.249}$
r_0 $[h^{-1}\text{Mpc}]$	$[0, 200]$	$12.904^{+1.495}_{-1.548}$
r_m $[h^{-1}\text{Mpc}]$	$[0, 200]$	$107.535^{+1.194}_{-1.434}$
N $[h^{-1}\text{Mpc}]$	$[0, 100]$	0.175 ± 0.02
σ $[h^{-1}\text{Mpc}]$	$[0, 100]$	$13.016^{+1.407}_{-1.188}$

Table 7.1: The flat prior intervals and parameter constraints derived from the Metropolis-Hastings algorithm to the empirical model described in section 7.1.6.

In order to focus our analysis in the BAO peak, we will take the two-point correlation estimator from section 6.5 between 46.6 $[h^{-1}\text{Mpc}]$ and 157 $[h^{-1}\text{Mpc}]$. We also will take the covariance matrix from section 6.6 between the same bins. Then, we are going to use several of the model described in the last section.

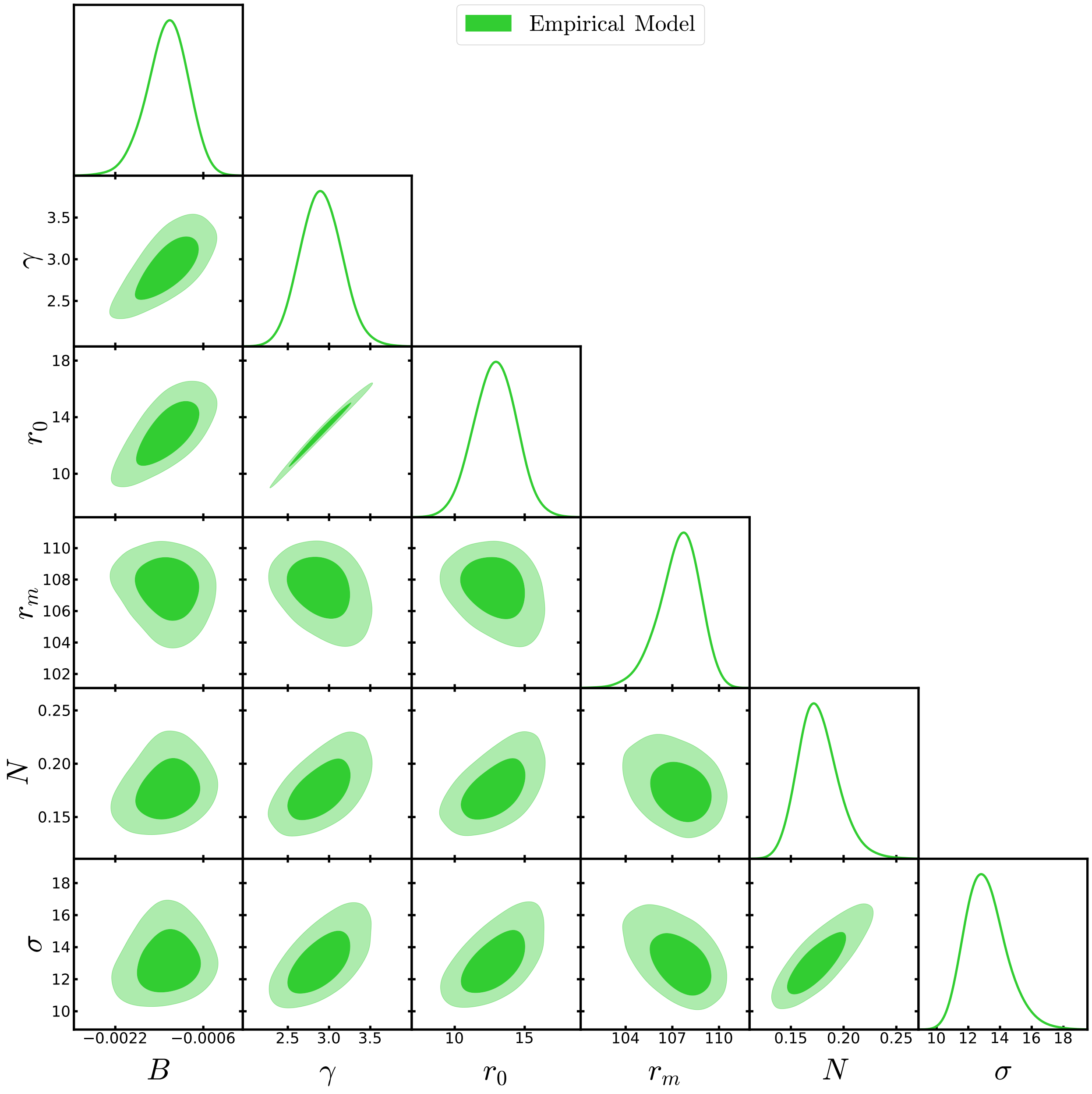


Figure 7.1: The contour plot for the empirical model, which shows the posterior probability distribution function, obtained from the Metropolis-Hastings algorithm, for all the parameters involved.

Model 0

We start using the empirical model from section 7.1.6, which we call model-0, from it we constrain the parameters

$$\vec{\theta}_{\text{model}-0} = \{B, \gamma, r_0, r_m, N, \sigma\} \ . \tag{7.26}$$

We use flat priors in the analysis, which are detailed in table 7.1. The results are shown as mean values and their 68% confidence intervals in table 7.1 and fig. 7.1.

Model 1

Now, we are going to use the geometrical model from section 7.1.2, which we call model-1, from it we constrain the parameters

$$\vec{\theta}_{\rm model-1} = \{B, \alpha\} \ . \tag{7.27}$$

We use flat priors in the analysis, which are detailed in table 7.2. The results are shown as mean values and their 68% confidence intervals in table 7.2 and fig. 7.2.

Model 2

We add the nuisance parameters to the last model, which is detailed in section 7.1.3. We name this model as model-2, from it we constrain the parameters

$$\vec{\theta}_{\rm model-2} = \{B, \alpha, a_0, a_1, a_2\} \ . \tag{7.28}$$

We use flat priors in the analysis, which are detailed in table 7.2. The results are shown as mean values and their 68% confidence intervals in table 7.2 and figs. 7.2 and 7.4.

Model 3

Now, we consider non-linearities (see section 7.1.4) using the damped power spectrum from eq. (7.11). For this model, we also require the wiggle-free power spectrum, which can be computed using the formula from [83]. In our case, we use a Savgol filter[1] to delete the baryon oscillation component. We name this model as model-3, from it we constrain the parameters

$$\vec{\theta}_{\rm model-3} = \{B, \alpha, a_0, a_1, a_2, \Sigma_{NL}\} \ . \tag{7.29}$$

We use flat priors in the analysis, which are detailed in table 7.2. The results are shown as mean values and their 68% confidence intervals in table 7.2 and fig. 7.2.

Model 4

We add the effect of redshift space distortions using the monopole, as we detailed in section 7.1.5, to the model in the last section. We call this model as model-4, from it we constrain the parameters

$$\vec{\theta}_{\rm model-4} = \{b\sigma_8, f\sigma_8, \alpha, a_0, a_1, a_2, \Sigma_{NL}\} \ . \tag{7.30}$$

We use flat priors over these parameters in the analysis, which are detailed in table 7.2. The results are shown as mean values and their 68% confidence intervals in table 7.2 and fig. 7.3.

Since the parameter $f\sigma_8$ seems not constrained, we also consider in this analysis scales between 26.8 [h^{-1}Mpc] and 165.8 [h^{-1}Mpc], because the RSD effect should be most notorious

---

[1] `https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.savgol_filter.html`

Parameter	Flat Priors	Model-1	Model-2	Model-3	Model-4	Model-4B	Model-4C	Model-5	Model-6	Model-7
$b\sigma_8$	$[0, 5]$	-	-	-	$1.58^{+0.15}_{-0.14}$	$1.85^{+0.09}_{-0.11}$	1.22 ± 0.05	-	-	-
$f\sigma_8$	$[0, 1]$	-	-	-	$0.51^{+0.32}_{-0.34}$	$0.65^{+0.19}_{-0.23}$	$0.47^{+0.14}_{-0.15}$	-	-	-
B	$[0, 10]$	2.46 ± 0.09	$3.73^{+0.25}_{-0.26}$	$4.91^{+0.64}_{-0.53}$	-	-	-	2.45 ± 0.10	2.45 ± 0.10	$4.20^{+0.51}_{-0.44}$
α	$[0, 2]$	0.94 ± 0.01	0.94 ± 0.01	0.95 ± 0.01	0.95 ± 0.01	1.15 ± 0.01	-	0.94 ± 0.01	0.94 ± 0.01	0.94 ± 0.01
Σ_{NL} [Mpc]	$[0, 100]$	-	-	12.8 ± 3.0	13.0 ± 3.0	$17.0^{+1.7}_{-3.2}$	-	-	-	-
k_* $\left[\text{Mpc}^{-1}\right]$	$[0, 1]$	-	-	-	-	-	-	$0.42^{+0.37}_{-0.19}$	$0.42^{+0.35}_{-0.19}$	$0.17^{+0.33}_{-0.05}$
A_{MC}	$[0, 10]$	-	-	-	-	-	-	-	$4.7^{+3.6}_{-3.2}$	$5.2^{+3.2}_{-3.6}$
a_0	$[-10, 10]$	-	-0.010 ± 0.001	-0.010 ± 0.001	-0.010 ± 0.001	$-0.0133^{+0.0003}_{-0.0005}$	$-3.99 \times 10^3 \pm 2.50 \times 10^{-05}$	-	-	-0.010 ± 0.001
a_1 $[h^{-1}\text{Mpc}]$	$[-10, 10]$	-	2.56 ± 0.27	$2.88^{+0.33}_{-0.32}$	$2.93^{+0.26}_{-0.25}$	$2.82^{+0.17}_{-0.07}$	0.88 ± 0.01	-	-	$2.78^{+0.40}_{-0.32}$
a_2 $\left[h^{-2}\text{Mpc}^2\right]$	$[-500, 100]$	-	-136.1 ± 16	$-176.2^{+22.9}_{-26.5}$	$-180.7^{+21.2}_{-23.0}$	$-128.3^{+4.1}_{-14.0}$	$-38.5^{+0.5}_{-0.8}$	-	-	$-160.0^{+24.3}_{-31.9}$

Table 7.2: The flat prior intervals and parameter constraints derived from Metropolis-Hastings algorithm to the models $1, 2, 3, 4, 4B, 4C, 5, 6$ and 7 (see section 7.2).

at smaller scales. We call this model as model-$4B$, which has the same parameters from eq. (7.30).

We also add another model because, from fig. 7.3, we can see that adding smaller scales shifts the α parameter. Therefore, we use another model with the same parameters from eq. (7.30), but fixing α and Σ_{NL} to those obtained with the model at large scales *i.e.* model-4. Besides, we consider scales between 15.4 $[h^{-1}\text{Mpc}]$ and 165.8 $[h^{-1}\text{Mpc}]$ and we call this model as model-$4C$. Finally, the results for both models $4B$ and $4C$ are also shown in table 7.2 and fig. 7.3.

Model 5

Now, we are going to test the other models for non-linearities (see section 7.1.4). We construct this model using as a base the geometrical model from section 7.1.2, with the damped power spectrum from eq. (7.5). Finally, we call this model as model-5, from it we constrain the parameters

$$\vec{\theta}_{\text{model}-5} = \{B, \alpha, k_*\} \ . \tag{7.31}$$

We use flat priors in the analysis, which are detailed in table 7.2. The results are shown as mean values and their 68% confidence intervals in table 7.2 and fig. 7.4.

Model 6

We add, to the model of the last section, the effect of mode-coupling using the eq. (7.9). We use the damped power spectrum from eq. (7.5) and we call this model as model-6, from which we are going to constrain the parameters

$$\vec{\theta}_{\text{model}-6} = \{B, \alpha, k_*, A_{MC}\} \ . \tag{7.32}$$

We use flat priors in the analysis, which are detailed in table 7.2. The results are shown as mean values and their 68% confidence intervals in table 7.2 and fig. 7.4.

Model 7

Finally, we add the nuisance parameters to the model of the last section. We call this model as model-7, from it we constrain the parameters

$$\vec{\theta}_{\text{model}-7} = \{B, \alpha, k_*, A_{MC}, a_0, a_1, a_2\} \ . \tag{7.33}$$

We use flat priors in the analysis, which are detailed in table 7.2. The results are shown as mean values and their 68% confidence intervals in table 7.2 and fig. 7.4.

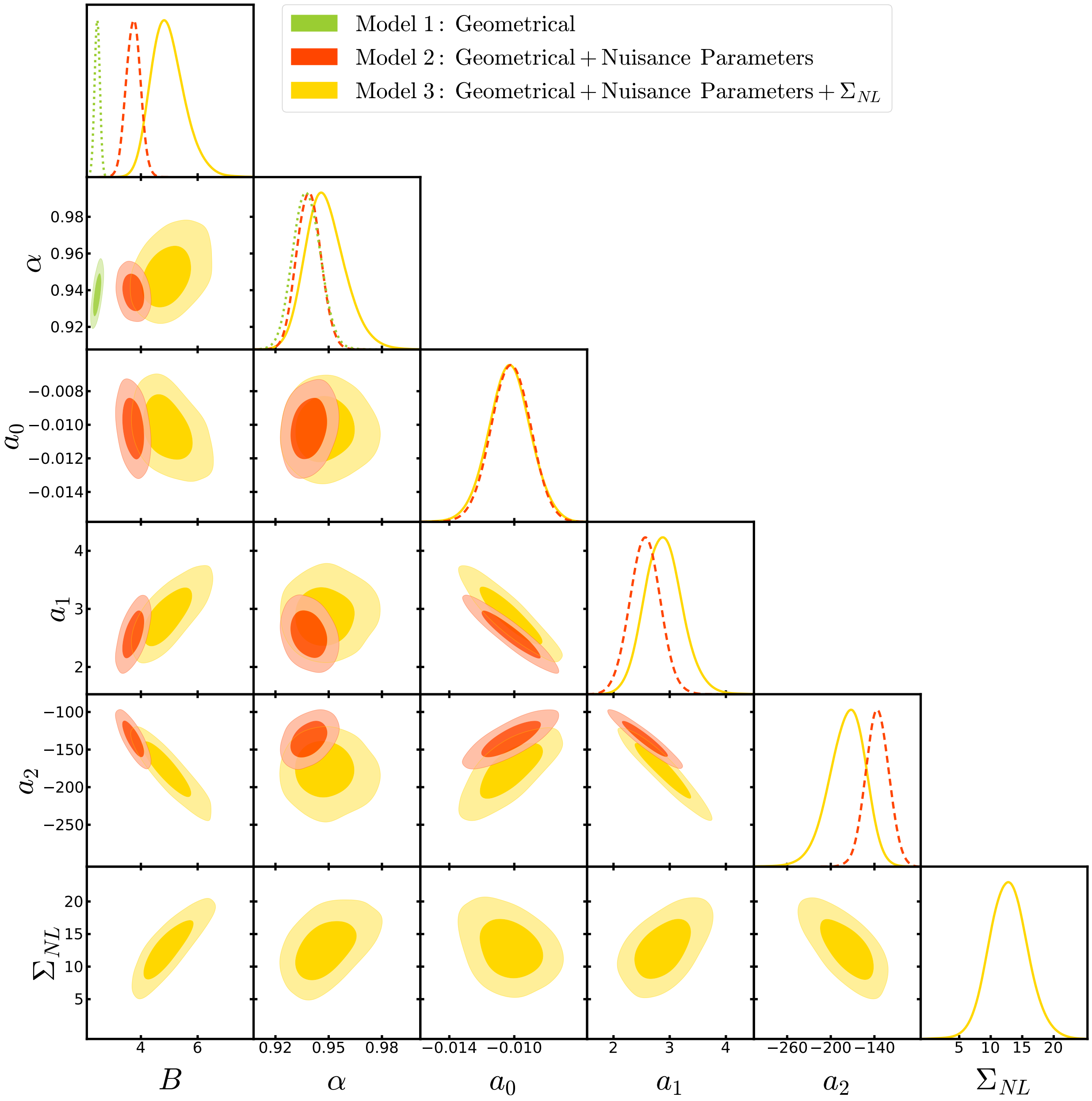


Figure 7.2: The contour plot for the models-1, 2, 3, which shows the posterior probability distribution function, obtained from the Metropolis-Hastings algorithm, for all the parameters involved. The numerical results are presented in table 7.2.

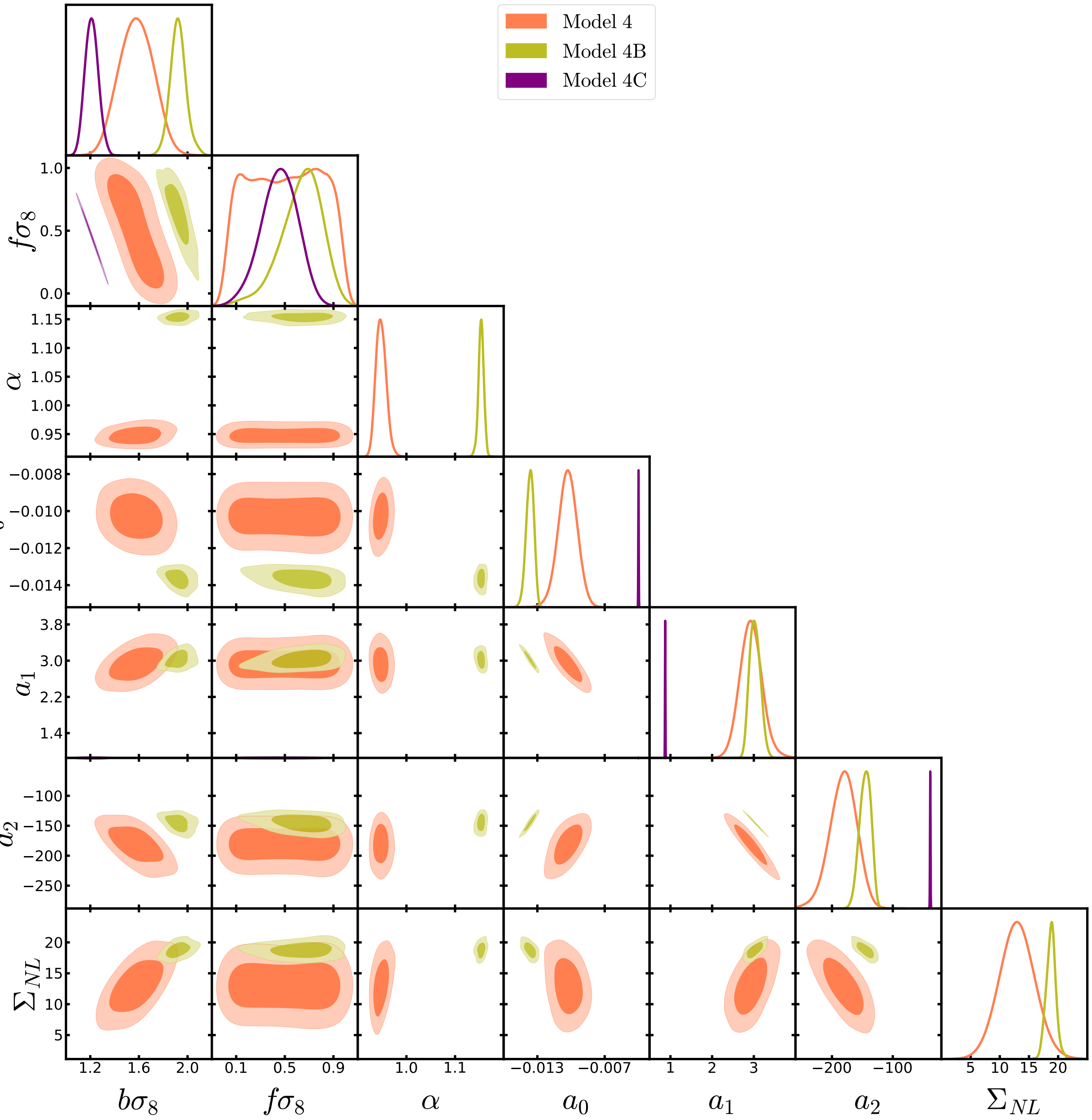


Figure 7.3: The contour plot for the models-4, $4B, 4C$, which shows the posterior probability distribution function, obtained from the Metropolis-Hastings algorithm, for all the parameters involved. The numerical results are presented in table 7.2.

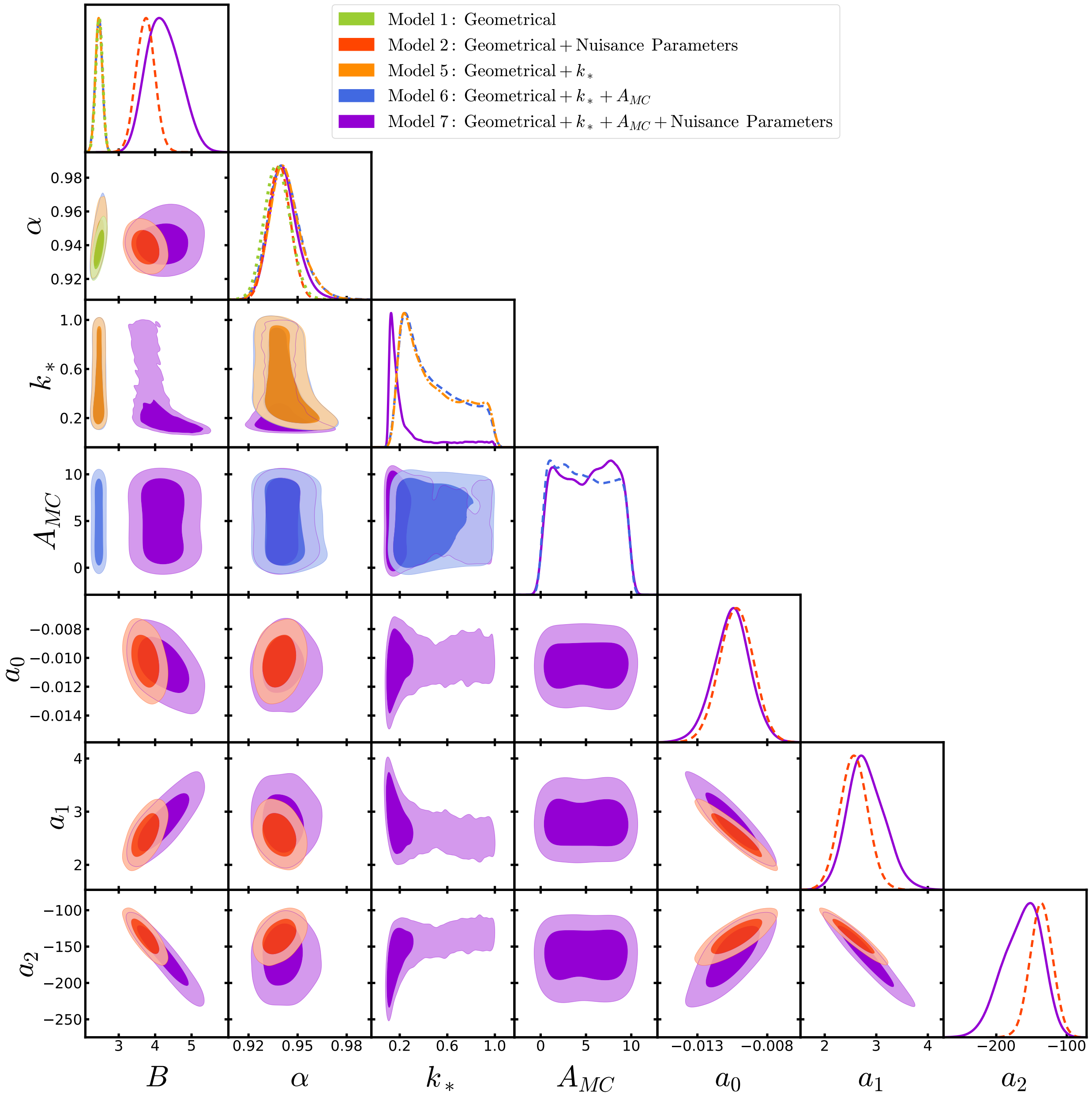


Figure 7.4: The contour plot for the models-1, 2, 5, 6, 7, which shows the posterior probability distribution function, obtained from the Metropolis-Hastings algorithm, for all the parameters involved. The numerical results are presented in table 7.2.

7.3 Results II

We also constrain the parameters using likelihood-free inference through DELFI (see section 5.6). We use the model-0 and model-3, because they were fully constrained through the Metropolis-Hastings algorithm in the last section. For DELFI, we use a `Python` library called `pydelfi`[2], where its methods are fully explained in [22]. As before, we will take the two-point correlation estimator from section 6.5 between 46.6 $[h^{-1}\text{Mpc}]$ and 157 $[h^{-1}\text{Mpc}]$. We also will take the covariance matrix from section 6.6 between the same bins.

For both models, we use the same architecture for the NDE, which consists in 5 MAFs, each one with 5 MADEs. These have two hidden layers with $[40, 40], [40, 40], [50, 50], [60, 60]$ and $[60, 60]$ neurons, and their activation functions between hidden layers are sigmoid (see eq. (5.11)), tanh (see eq. (5.12)), sigmoid, sigmoid, tanh, respectively.

Besides, to ensure the convergence of the algorithm and to find the correct weights, we run 8000 simulations to minimize the negative log-loss, which is defined in eq. (5.22). The convergence is made by splitting the simulations into a train set (90%) and a validation set (10%). The results are shown in fig. 7.5.

Model 0

We use the same parameters and model from eqs. (7.23) and (7.26), respectively. We use flat priors in the analysis, which are detailed in table 7.1. The results are shown as mean values and their 68% confidence intervals in table 7.3 and fig. 7.6, where we also show the comparison, in the contour plot, between the results obtained from Metropolis-Hastings algorithm and from DELFI.

Model 3

We use the same parameters and model from section 7.1.3 and eq. (7.29), respectively. We use flat priors in the analysis, which are detailed in table 7.2. The results are shown as

---

[2] `https://pydelfi.readthedocs.io/en/latest/index.html`

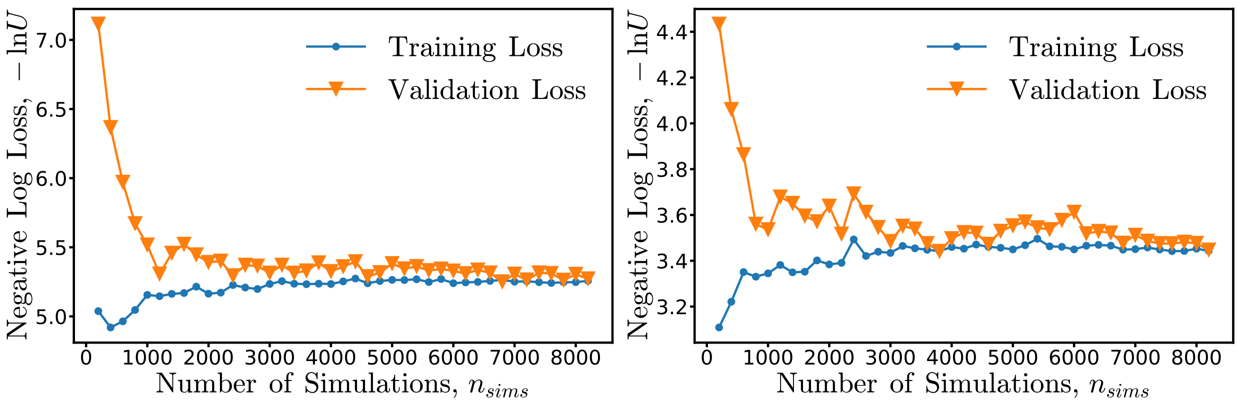


Figure 7.5: The negative log-loss from eq. (5.22) as function of the number of simulations for DELFI in the model-0 (*left*) and the model-3 (*right*).

Parameter	Parameter Constraints
B	$-0.0013^{+0.0004}_{-0.0003}$
γ	$2.855^{+0.248}_{-0.259}$
r_0 $[h^{-1}\mathrm{Mpc}]$	$12.666^{+1.454}_{-1.606}$
r_m $[h^{-1}\mathrm{Mpc}]$	$107.624^{+1.089}_{-1.184}$
N $[h^{-1}\mathrm{Mpc}]$	$0.174^{+0.017}_{-0.015}$
σ $[h^{-1}\mathrm{Mpc}]$	$12.973^{+1.094}_{-0.974}$

Parameter	Parameter Constraints
B	5.03 ± 0.50
α	0.94 ± 0.01
a_0	-0.010 ± 0.001
a_1 $[h^{-1}\mathrm{Mpc}]$	$2.95^{+0.34}_{-0.32}$
a_2 $\left[h^{-2}\mathrm{Mpc}^2\right]$	$-184.1^{+24.4}_{-25.8}$
Σ_{NL} [Mpc]	$12.9^{+2.2}_{-2.4}$

Table 7.3: The parameter constraints derived from DELFI to the empirical model (*left*) described in section 7.1.6 and for the geometrical analysis (*right*) described in eq. (7.13).

mean values and their 68% confidence intervals in table 7.3 and fig. 7.7, where we also show the comparison, in the contour plot, between the results obtained from Metropolis-Hastings algorithm and from DELFI.

7.4 Discussion

In section 7.2, we use several models to constrain the shift parameter α, which shows an excellent concordance between them, except for model-4B. However, this parameter has only 1% of uncertainty (see table 7.2), which is low compared with analysis in almost the same sky patch [210] and redshift [38] that predict a 4% of uncertainty. Several errors in the data production can cause this difference because the models were taken from those works. Then, we highlight the likest next

1. In the covariance matrix production, we only used 40 sky patches, but 45 bins to construct the correlation function. This partition can induce several problems since the number of jackknife cuts must be greater than the number of bins, as we will see in chapter 10. Also, how many patches are necessary to estimate the covariance better, depends on the scale. For example, in an ideal case, we could take one point off to compute the covariance matrix, but because the correlation function depends on hundreds of thousands of points, the different computations would be the same, and then the covariance would be zero, at least in the scales of interest. Therefore, for each scale, there are a number of partitions that maximize the covariance, but it is difficult to determine precisely. In chapter 9, we use 200 jackknife resamplings and 20 bins for the same survey, and in chapter 10, we delve more into this topic.
2. The error generated by the estimation of the precision matrix as the inverse of the covariance matrix estimator. As we explained in section 7.2, this assumption could induce a bias in the parameter to constrain. This also explains the tension between our measurement and those in [210, 38]. The correct estimation needs the Hartlap and tapering corrections (see chapters 9 and 10 for a better handle of this effect).
3. Both results from [210, 38], were obtained through a covariance matrix produced by mock catalogs. This fact could induce a difference, mostly due to the cosmic variance present because we replacement ensemble averages by volume ones, as we mentioned in sections 4.1 and 6.6. In chapter 10, we compare the differences between jackknifes

and log-normal realizations in the estimation of the α parameter.

4. Finally, in [210], they made a data reconstruction algorithm in which the positions of the galaxies are displaced to correct the RSD due to their proper movements before constructing the two-point correlation estimator. Also, this algorithm helps to increase the power of constraining and bias determination, especially in data sets at low redshift.

Besides, since most of the results from different models are statistically consistent, makes us think that effectively the discrepancies with [210, 38] are due to the issues in the data production mentioned above, and not due to model differences.

The addition of the Σ_{NL} parameter shifts a bit the α parameter (see the results from model-3 versus models-$1, 2$). This is expected since the physical problem is preserved, even with an underestimated covariance matrix. For the models-$5, 6, 7$, the k_* parameter, which is analogous to Σ_{NL}, seems not completely constrained for models-$5, 6$, but we can distinguish a peak on its distribution. Nonetheless, adding the nuisance parameters, the constrain of k_* improves and it takes a value statistically consistent with Σ_{NL} (see eq. (7.12)). On the other hand, the addition of the mode-coupling parameter A_{MC} does not improve the constraining (see the results from model-6 versus model-5), because it is not constrained as we can see in its probability distribution function that is still flat like the prior.

For models that consider the RSD (see the results from models-$4, 4B, 4C$), we can see at first instance that by working in the same scales that the rest of the models, we can not constrain the $f\sigma_8$ parameter since the probability distribution function for model-4 is still flat, like the prior. We obtain better results considering the scales of model-$4B$, but they also shift the α parameter considerably. Since its constrain depends on the BAO peak, we can ensure that the results with model-$4B$ scales are completely biased. For those reasons, we also add the model-$4C$, which fixes the α and Σ_{NL} parameters. This model also considers a few more small scales, and in this analysis, we obtain a well constrain of the $f\sigma_8$ parameter. However, there is a great dependence on the $b\sigma_8$ parameter, as we can see in their contour plot. The exclusion of the two aforementioned parameters also reduces the uncertainties, which is most notorious for the nuisance parameters that seem extremely constrained compared to those from models-$4, 4B$. Also, the results for $f\sigma_8$ and $b\sigma_8$ from model-$4C$ are compared to those in [127], which shares the same data from [210] but with a focused analysis in RSD. We obtain consistent constraints to those results, but the uncertainties differ because of the same reasons for the data production highlighted before.

For the empirical model, we obtained well-constrained results recovering the BAO peak from the r_m parameter (see table 7.1). Also, from fig. 7.1, we can observe a high correlation between the γ and r_0 parameters, which is due to the functional form of the model in eq. (7.23).

In section 7.3, we take two models that are well constrained from section 7.2 to compare this inference technique with the previous. Then, we can see from tables 7.1 and 7.3 that the results between both inference methods agree because they are statistically consistent. Their probability distribution functions for the parameters of model-0 are quite similar, as we can observe in the contour plot from fig. 7.6. The log-loss (see fig. 7.5) converges for this model, between the validation and training sets. This indicates that we are not falling into overfitting or underfitting. Therefore, we can claim that DELFI works well for this model, given its results, even when we used an underestimated covariance matrix.

For the model-3 we also obtain statistically consistent results (see tables 7.2 and 7.3) and probability distribution functions quite similar for most of the parameters as we can see from fig. 7.7. However, the α parameter seems slightly shifted using this method, which requires more research in the field since it could be due to the architecture of the neural network or its hyper-parameters. Nonetheless, as for model-0, the log-loss from fig. 7.5 converges between validation and training sets, which indicates that we are not into overfitting or underfitting, but it is more noisy compare to the log-loss from model-0.

Finally, the principal hypothesis of section 7.3 is to determinate if DELFI is good enough to constrain this kind of experiments, and given the results from both models, we can claim that it is true, even when the existing problems in the determination of the covariance matrix that we highlighted at the beginning of this section. Thus, we can add these likelihoods to the DELFI list, which in cosmology includes Supernovae, Weak Lensing, and Lyman-α, as is shown in [22].

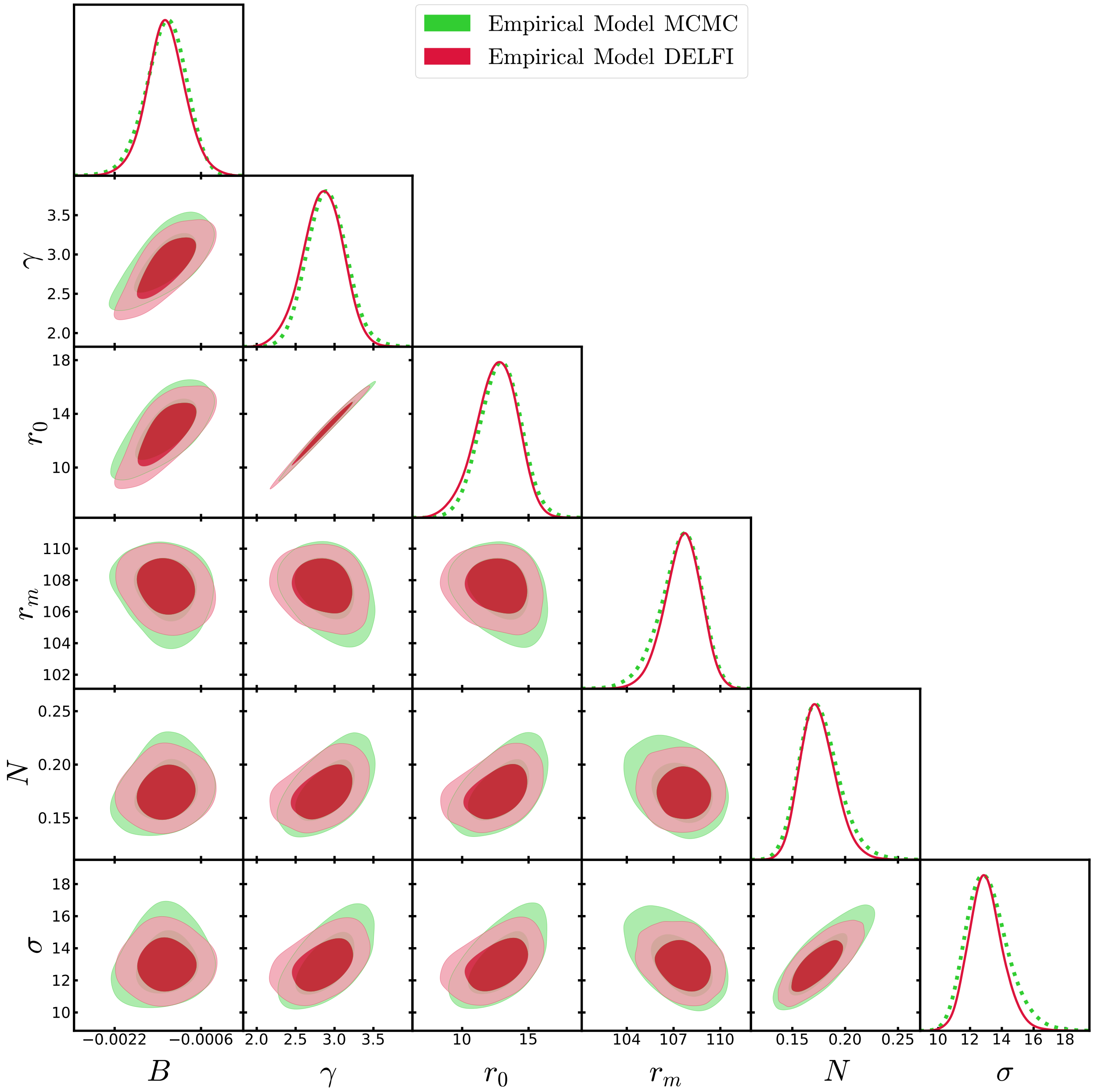


Figure 7.6: The contour plot for the model-0, which shows the comparison of the posterior probability distribution function obtained from the Metropolis-Hastings versus DELFI. The numerical results are presented in tables 7.1 and 7.3.

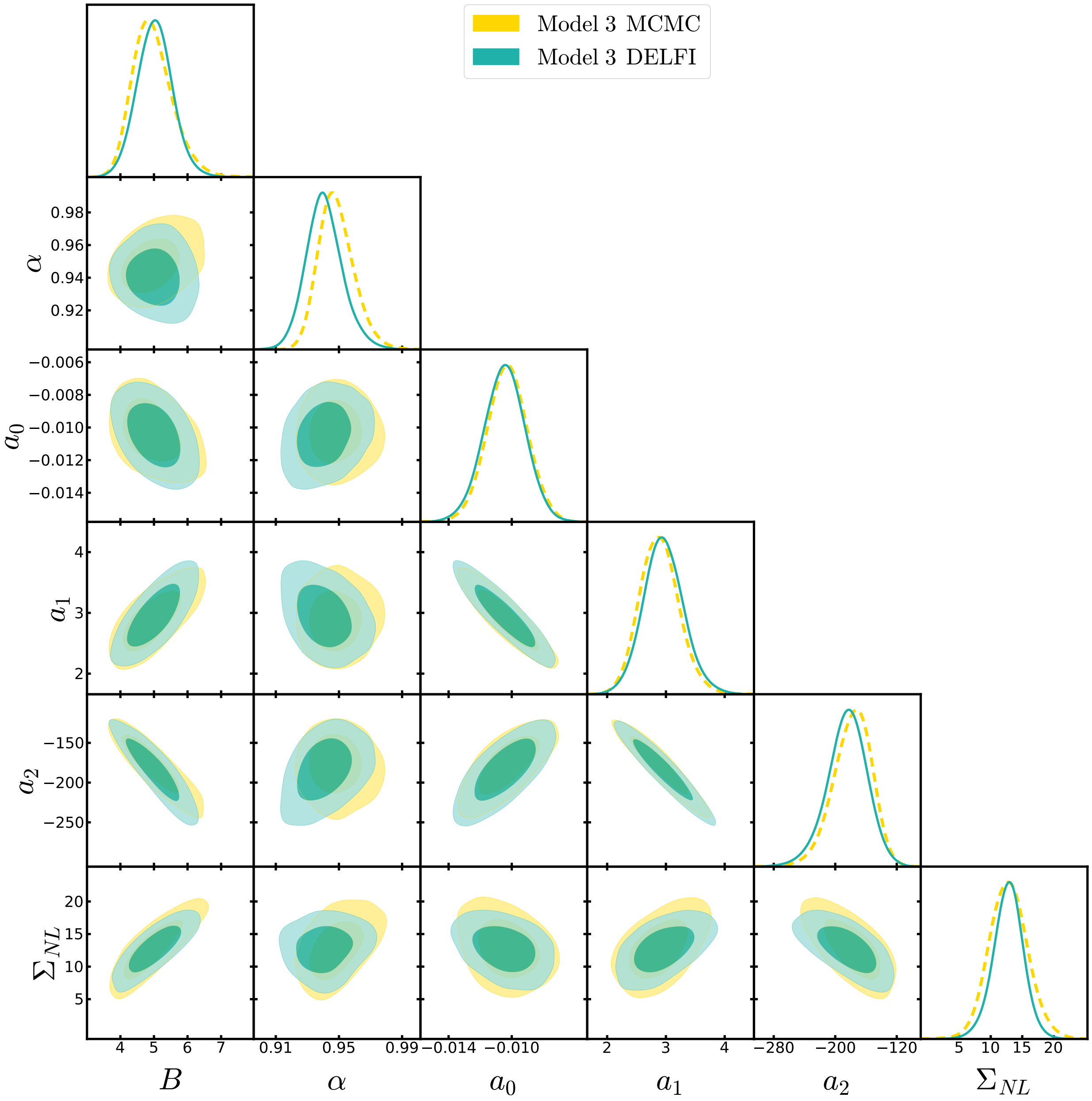


Figure 7.7: The contour plot for the model-3, which shows the comparison of the posterior probability distribution function obtained from the Metropolis-Hastings versus DELFI. The numerical results are presented in tables 7.2 and 7.3.

Chapter 8

Comparing Dark Energy models with Hubble versus Growth Rate data

This chapter presents a research work which corresponds to [213]. This work is a team effort with aid from Bryan Sagredo and Dr. Domenico Sapone, both affiliated by that time to the Cosmology and Theoretical Astrophysics group at the Facultad de Ciencias Físicas y Matemáticas of the Universidad de Chile in Chile.

Abstract

In this work we perform an analysis on the recently proposed conjoined cosmic growth and cosmic expansion diagram [155] to compare several dark energy models using the Figure of Merit showed in [29], which consists in the inverse of the 1σ confidence region in the $f\sigma_8(z)-H(z)$ plot. Our analysis also consists of comparing the models by performing different statistical criteria: Bayes factor [249], the Bayesian Information Criteria [223] and the Akaike Information Criterion [243]. We also developed a 3-dimensional Figure of Merit to account simultaneously for the errors on the growth rate and the Hubble parameter. The main idea is to consider several cosmological models and compare them with the different statistical criteria in order to highlight the differences and the accuracies of each single criterion.

Keywords: cosmology, dark energy, model comparison.

8.1 Introduction

Recent observations [208, 197] pointed out that the Universe seems to be in a phase of accelerated expansion. These evidences have led cosmologists to revise the theory of the expansion of the Universe either by introducing a new component called dark energy [218] or by modifying directly the theory of gravity [251].

Within the framework of Friedmann-Lemaître-Robertson-Walker (FLRW) cosmologies, such accelerated expansion can be generated by adding up a simple cosmological constant Λ to the total budget of the Universe. Even though the latter gives rise to severe coincidence and fine-tuning problems, observations still confirm such an explanation [37, 6, 2]. Over the years a series of dark energy models have been considered in order to solve, or at least alleviate, the theoretical problems related to dark energy. However, none of these explanations seem to be convincing.

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

The important question here is whether the two scenarios can be distinguished. It is well known that any Hubble expansion can be generated by choosing an appropriate equation of state for the dark energy, see [49]. However, over the years there have been claims that it is possible to distinguish alternative theories of gravity from dark energy models by using growth data; the last assumption is not always true unless the expansion history is fixed, [144]. Nonetheless, recent works have proposed to study the cosmic growth versus cosmic expansion history conjoined diagram, the $f\sigma_8 - H$ plot, to put constraints on the parameter space of cosmological models, or to compare different models directly [155]. Model comparison using this approach has already been investigated in [176, 29]. The advantage of the $f\sigma_8 - H$ plot over other probes relies on the degeneracy break of the history curves when comparing different models or the parameter space, since it contrasts a geometrical observable, given by $H(z)$, to a pure gravitational effect, given by $f\sigma_8(z)$.

Using this approach, dark energy models were compared using the $f\sigma_8 - H$ plot [29] through the FoM defined as the inverse area of the 1σ confidence region in the conjoined diagram. In this work, we follow a similar approach and we also compare the models using different statistical tools: the standard Bayesian evidence [249], the Bayesian Information Criteria [223] (BIC), the Akaike Information Criterion [14] (AIC) and the FoM. Furthermore, we considered an extension to the FoM which we define 3-FoM, which considers both errors on $f\sigma_8(z)$ and $H(z)$.

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

The paper is structured as follows: in Section 8.2 we report the basic equations that will be used in our work, whereas in Section 8.3 we list the cosmological models that will be compared, and the link between H and $f\sigma_8$ measurements. In Section 8.4 we show the datasets used in the analysis and the statistical methodology is reported in Section 8.5. In Section 8.6 we report the results of our analysis.

8.2 Basic equations

The evolution of a general fluid can be expressed in terms of its present energy density parameter Ω_0 and its equation of state parameter (EoS) $w(a) = p(a)/\rho(a)$, where p and ρ are the pressure and energy density of the fluid, respectively, and a is the scale factor, normalized to 1 today. The EoS is the key quantity that fully characterizes the fluid at the background level.

Using a general formalism the Hubble parameter in a non-flat cosmology is given by

$$H^2 = H_0^2\left[\Omega_{m_0}a^{-3} + \Omega_{k_0}a^{-2} + \Omega_{\mathrm{de}_0}a^{-3(1+\hat{w})}\right] , \tag{8.1}$$

where H_0 is the Hubble constant, Ω_{m_0}, Ω_{k_0}, Ω_{de_0} are the present-day values of matter, curvature and dark energy densities, respectively. Furthermore, the parameters satisfy the relation $\Omega_{m_0} + \Omega_{k_0} + \Omega_{\mathrm{de}_0} = 1$. The total matter density is:

$$\Omega_m(a) = \left(1 + \frac{1-\Omega_{m_0}}{\Omega_{m_0}}a^{-3\hat{w}}\right)^{-1} . \tag{8.2}$$

The quantity $\hat{w}$ in Eqs. (8.1) and (8.2) is the effective EoS parameter accounting for the time dependence, given by

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

The angular diameter distance is defined as

$$\mathrm{d}_A(z) = \frac{c\,H_0^{-1}}{(1+z)\sqrt{-\Omega_{k_0}}}\sin\left(\sqrt{-\Omega_{k_0}}\int_0^z \frac{H_0}{H(y)}\mathrm{d}y\right) ,$$

which reduces to

$$\mathrm{d}_A(z) = \frac{c}{H_0}\frac{1}{1+z}\int_0^z \frac{H_0}{H(y)}\mathrm{d}y \; ,$$

if the curvature is set to zero.

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

The growth of perturbations for a general fluid is governed, assuming homogeneity and isotropy, by the differential equations [160]

$$\begin{aligned} \delta' = &\; 3 \quad (1+w)\,\phi' - \frac{V}{H\,a^2} - \frac{3}{a}\left(\frac{\delta p}{\rho} - w\delta\right) , && (8.3)\\ V' = &\; - \quad (1-3w)\,\frac{V}{a} + \frac{k^2}{H\,a^2}\frac{\delta p}{\rho} + (1+w)\frac{k^2}{H\,a^2}\psi \\ &\; - \quad (1+w)\frac{k^2}{H\,a^2}\sigma \; , && (8.4) \end{aligned}$$

where the primes denote derivatives with respect to the scale factor a, $\delta = \delta\rho(a,\, k)/\rho(a)$ is the density contrast, $V = i\,k_j T_0^j/\rho(a)$ is the scalar velocity perturbation. The quantities ψ

and ϕ are the gravitational potentials in the Newtonian gauge. These potentials follow

$$k^2\phi = -4\pi G a^2 \sum_j \rho_j \left(\delta_j + \frac{3aH}{k^2} V_j\right) , \tag{8.5}$$

$$k^2\left(\phi - \psi\right) = 12\pi G a^2 \sum_j \left(\rho_j + p_j\right)\sigma_j , \tag{8.6}$$

where the sum runs over all the species in the Universe. We will then have sets of equations of the form of Eqs. (8.3) and (8.4) depending on the number of species present in the Universe. For non-relativistic particles, i.e. cold dark matter and baryons, we just need to set $w = \delta p = \sigma = 0$. However, in this paper we consider general dark energy models as well. There is no unique way to parametrize these quantities as they depend directly on the specific model considered.

For simplicity, in this work, we consider only two components, and they are pressureless dark matter and a dark energy fluid, because we are more interested on how the different criteria reacts to a particular model. In the next section, we describe the different dark energy models.

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

$$f\sigma_8(a) = \sigma_{8,0}\, a \frac{\delta_m'(a)}{\delta(a=1)} , \tag{8.7}$$

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

8.3 Models

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

ΛCDM

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

ΛCDM: this model refers to flat ΛCDM (without spatial curvature) where we set the curvature parameter $\Omega_{k_0} = 0$, hence the Hubble parameter in Eq. (8.1) reads

$$H^2 = H_0^2\left[\Omega_{m_0}a^{-3} + (1 - \Omega_{m_0})\right] .$$

Furthermore, the cosmological constant Λ has zero perturbations, hence the system of equations simplifies and the gravitational potentials only depend on pressureless matter. For small scales, Eqs. (8.3) - (8.6) reduce to a single second-order differential equation for matter density contrast to which an analytical solution[1] can be found, see Appendix 8.7 for more details.

Consequently, we will have one model with two variants: ΛCDM and ΛCDM-a, using the numerical and analytic solution, respectively. However, for consistency reasons, we decided to use the full set of differential equations Eqs. (8.3) - (8.6), leaving to the appendix the results obtained by using the analytical solutions as a further test.

Finally, the parameters of both models are:

$$\boldsymbol{\theta}_{\Lambda\text{CDM}} = (\Omega_{m_0},\, H_0,\, \sigma_{8,0})\,. \tag{8.8}$$

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

$$H^2 = H_0^2\left[\Omega_{m_0}a^{-3} + (1-\Omega_{m_0}-\Omega_{\text{de}_0})a^{-2} + \Omega_{\text{de}_0}\right].$$

The cosmological constant still has zero perturbations, however the differential equation for matter perturbations does not have an analytical solution, hence we solve numerically Eqs. (8.3) - (8.6).

The parameters of the model are:

$$\boldsymbol{\theta}_{\Lambda\text{CDM-nf}} = (\Omega_{m_0},\, \Omega_{\text{de}_0},\, H_0,\, \sigma_{8,0})\,. \tag{8.9}$$

*w*CDM

This model is an extension of the ΛCDM model in which a constant EoS w is set as a free parameter. If the EoS parameter of dark energy is no longer constant and equal to -1, then dark energy may have perturbations and its growth will be fully characterized by the values of w and c_s[2]. Clearly, if dark energy has perturbations, these will affect the growth of matter perturbations through the gravitational potential Eqs. (8.5) - (8.6). We identify four different cases.

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

$$H^2 = H_0^2\left[\Omega_{m_0}a^{-3} + (1-\Omega_{m_0})a^{-3(1+w)}\right]. \tag{8.10}$$

If we decide to ignore *a priori* the dark energy perturbations, then the growth of matter density is still governed by a second order differential equation, and it is still possible to find an analytical solution to the matter density contrast, see Appendix 8.7 for more details. As a consequence we have the wCDM and wCDM-a solutions to this model. As for the ΛCDM

---

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

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

case, we also consider the full numerical solutions from Eqs. (8.3) - (8.6) and leaving the results from the analytical solution to the appendix.

Finally, the parameters of both models are

$$\boldsymbol{\theta}_{\text{wCDM}} = (\Omega_{m_0},\, w,\, H_0,\, \sigma_{8,0}) \, . \tag{8.11}$$

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

$$H^2 = H_0^2\Big[\quad \Omega_{m_0}a^{-3} + (1-\Omega_{m_0}-\Omega_{\text{de}_0})a^{-2} + \\ \Omega_{\text{de}_0}a^{-3(1+w)}\Big] \, . \tag{8.12}$$

Here we set dark energy perturbations to zero. However, due to the complexity of the Hubble parameter, analytical solutions for the matter density contrast do not exist and we solve numerically the system of Eqs. (8.3) - (8.6).

We have the following free parameters for the model:

$$\boldsymbol{\theta}_{\text{wCDM-nf}} = (\Omega_{m_0},\, \Omega_{\text{de}_0},\, w,\, H_0,\, \sigma_{8,0}) \, . \tag{8.13}$$

w**CDM-p:** this model is a flat wCDM for which we allow perturbations (this addition is symbolized by '-p') in the dark energy sector. The Hubble parameter is given by Eq. (8.10). However, we now have two sets of equations (8.3) - (8.4), for pressureless matter and for the dark energy fluid. Analytical solutions can also be found in some special limits, see Appendix 8.7. However, as for the other cases, we also use the full numerical solutions from the equation of perturbations.

As mentioned earlier, the growth of the perturbations of one species depends on the characteristics of the fluid, which are given by w, δp and σ. For pure pressureless matter, $w = \delta p = \sigma = 0$. For a dark energy fluid, we assume zero anisotropic stress $\sigma = 0$, and the pressure perturbation to be given by [143]:

$$\delta p = c_s^2\rho\delta + \frac{3aH(c_s^2 - c_a^2)}{k^2}\rho V \, , \tag{8.14}$$

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

$$c_a^2 = w - \frac{\dot{w}}{3H(1+w)} = w - \frac{w'}{3(1+w)} \, , \tag{8.15}$$

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

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

$$\boldsymbol{\theta}_{\text{wCDM-p}} = (\Omega_{m_0},\, w,\, c_s^2,\, H_0,\, \sigma_{8,0}) \, . \tag{8.16}$$

w**CDM-nf-p:** this model corresponds to a non-flat wCDM for which we allow perturbations in the dark energy sector; the Hubble parameter takes the form in Eq. (8.12) and the perturbations will be solved numerically for both matter and dark energy. Thus, the parameter set of the model is

$$\boldsymbol{\theta}_{\text{wCDM-nf-p}} = (\Omega_{m_0},\, \Omega_{\text{de}_0},\, w,\, c_s^2,\, H_0,\, \sigma_{8,0}) \, . \tag{8.17}$$

Chevallier-Polarski-Linder (CPL)

This class of models [53, 154] can be considered an extension to wCDM models in which the equation of state depends on the scale factor. The simplest extension is a Taylor expansion around the present time $a = 1$, giving

$$w(a) = w_0 + w_a(1 - a)\,. \tag{8.18}$$

Hence, giving two extra parameters: w_0, which is the present time EoS parameter and w_a which represents the variation over time of $w(a)$. We identify four different models using this parametrization.

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

$$H^2 = H_0^2\big[\Omega_{m_0}a^{-3} + (1 - \Omega_{m_0})a^{-3(1+\hat{w}(a))}\big]\,. \tag{8.19}$$

There is no exact analytic expression for the matter density contrast when the EoS parameter takes the form of Eq. (8.18) but only approximated analytical solutions, [51]. Here we solve numerically Eqs. (8.3) - (8.6). This way, the parameters of are

$$\boldsymbol{\theta}_{\text{CPL}} = (\Omega_{m_0},\, w_0,\, w_a,\, H_0,\, \sigma_{8,0})\ . \tag{8.20}$$

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

$$\begin{aligned} H^2 = H_0^2 \quad & \big[\Omega_{m_0}a^{-3} + (1 - \Omega_{m_0} - \Omega_{\text{de}_0})a^{-2} + \\ & \Omega_{\text{de}_0}a^{-3(1+\hat{w}(a))}\big]\ . \end{aligned} \tag{8.21}$$

We set dark energy perturbations to zero and solve numerically the Eqs. (8.3) - (8.6) for pure pressureless matter only. The free parameters of the model are:

$$\boldsymbol{\theta}_{\text{CPL-nf}} = (\Omega_{m_0},\, \Omega_{\text{de}_0},\, w_0,\, w_a,\, H_0,\, \sigma_{8,0})\ . \tag{8.22}$$

CPL-p: the Hubble parameter is given by Eq. (8.19), the equation of perturbations will be solved numerically by using Eqs. (8.3) - (8.6). The characteristics of the dark energy fluid are given by Eq. (8.14), with the further assumption that the adiabatic sound speed $c_a^2 = w$; the former is somehow required in order to stabilize the growth of dark energy perturbations when it crosses the phantom regime [143]. Thus, the parameters of the model are

$$\boldsymbol{\theta}_{\text{CPL-p}} = (\Omega_{m_0},\, w_0,\, w_a,\, c_s^2,\, H_0,\, \sigma_{8,0})\ . \tag{8.23}$$

CPL-nf-p: the Hubble parameter takes the form in Eq. (8.21). We solve numerically Eqs. (8.3) - (8.6) for pressureless matter and dark energy. The characteristic of the dark energy fluid are given by Eq. (8.14), with the further assumption of $c_a^2 = w$ and $w' = 0$ at crossing.

The parameter set for the model is

$$\boldsymbol{\theta}_{\text{CPL-nf-p}} = (\Omega_{m_0},\, \Omega_{\text{de}_0},\, w_0,\, w_a,\, c_s^2,\, H_0,\, \sigma_{8,0})\ . \tag{8.24}$$

8.4 Data

The Hubble parameter data for the analysis are the *cosmic chronometers* compilation used in [163], which consists in 31 independent measurements of $H(z)$, obtained from evolving galaxies at different redshifts [177].

z	$H(z)$	$\sigma_H(z)$	Ref.	z	$H(z)$	$\sigma_H(z)$	Ref.
0.07	69.0	19.6	[273]	0.4783	80.9	9.0	[177]
0.09	69.0	12.0	[229]	0.48	97.0	62.0	[239]
0.12	68.6	26.2	[273]	0.593	104.0	13.0	[174]
0.17	83.0	8.0	[229]	0.68	92.0	8.0	[174]
0.179	75.0	4.0	[174]	0.781	105.0	12.0	[174]
0.199	75.0	5.0	[174]	0.875	125.0	17.0	[174]
0.2	72.9	29.6	[273]	0.88	90.0	40.0	[239]
0.27	77.0	14.0	[229]	0.9	117.0	23.0	[229]
0.28	88.8	36.6	[273]	1.037	154.0	20.0	[174]
0.352	83.0	14.0	[174]	1.3	168.0	17.0	[229]
0.3802	83.0	13.5	[177]	1.363	160.0	33.6	[175]
0.4	95.0	17.0	[229]	1.43	177.0	18.0	[229]
0.4004	77.0	10.2	[177]	1.53	140.0	14.0	[229]
0.4247	87.1	11.2	[177]	1.75	202.0	40.0	[229]
0.4497	92.8	12.9	[177]	1.965	186.5	50.4	[175]
0.47	89.0	49.6	[206]				

Table 8.1: The 31 cosmic chronometer data points used in this analysis along with their related references. The $H(z)$ and $\sigma_H(z)$ data are in units of $\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

The growth rate dataset is based on the compilation used in [214], which is an updated version of the 'Gold-2017' dataset from [181]. The dataset consists of 22 independent measurements of $f\sigma_8(z)$, obtained through baryon acoustic oscillations and weak lensing surveys. Among these surveys, it is important to note that the three WiggleZ [42] and the four SDSS-IV [274] measurements are correlated, and their covariance matrices are

$$\mathbf{C}_{\mathrm{WiggleZ}} = 10^{-3}\begin{pmatrix} 6.400 & 2.570 & 0.000 \\ 2.570 & 3.969 & 2.540 \\ 0.000 & 2.540 & 5.184 \end{pmatrix}, \tag{8.25}$$

$$\mathbf{C}_{\mathrm{SDSS\text{-}IV}} = 10^{-2}\begin{pmatrix} 3.098 & 0.892 & 0.329 & -0.021 \\ 0.892 & 0.980 & 0.436 & 0.076 \\ 0.329 & 0.436 & 0.490 & 0.350 \\ -0.021 & 0.076 & 0.350 & 1.124 \end{pmatrix}. \tag{8.26}$$

8.5 Methodology

To perform the analysis, both datasets are assumed to have Gaussian likelihood distributions, this is the probability of the data given a set of parameters. The datasets are assumed to be independent, thus their conjoined likelihood is the product of each dataset's likelihood. In terms of the traditional chi-squared, defined by $\chi^2 \equiv -2\log L$, where L is the likelihood of the current model, it is simply given by the sum of each dataset's chi-squared, or

$$\chi^2 = \chi^2_H + \chi^2_{f\sigma_8} \tag{8.27}$$

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

z	$f\sigma_8(z)$	$\sigma_{f\sigma_8}(z)$	$\Omega^{\mathrm{ref}}_{m_0}$	Ref.
0.02	0.428	0.0465	0.3	[131]
0.02	0.398	0.065	0.3	[253],[130]
0.02	0.314	0.048	0.266	[70],[130]
0.10	0.370	0.130	0.3	[95]
0.15	0.490	0.145	0.31	[127]
0.17	0.510	0.060	0.3	[235]
0.18	0.360	0.090	0.27	[43]
0.38	0.440	0.060	0.27	[43]
0.25	0.3512	0.0583	0.25	[215]
0.37	0.4602	0.0378	0.25	[215]
0.32	0.384	0.095	0.274	[217]
0.59	0.488	0.060	0.307115	[56]
0.44	0.413	0.080	0.27	[42]
0.60	0.390	0.063	0.27	[42]
0.73	0.437	0.072	0.27	[42]
0.60	0.550	0.120	0.3	[198]
0.86	0.400	0.110	0.3	[198]
1.40	0.482	0.116	0.27	[184]
0.978	0.379	0.176	0.31	[274]
1.23	0.385	0.099	0.31	[274]
1.526	0.342	0.070	0.31	[274]
1.944	0.364	0.106	0.31	[274]

Table 8.2: Compilation of the cosmic growth $f\sigma_8(z)$ measurements used in this analysis along with the reference matter density parameter Ω_{m_0} (needed for the redshift correction) and associated references.

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

$$\mathbf{x}_s = \mathbf{m}_s - \boldsymbol{\mu}_s \ , \tag{8.28}$$

with the subscript 's' denoting the data source: H or $f\sigma_8$. However, in the case of growth measurements, we need to take into account a redshift correction, which is featured in Ref. [181]. This correction consists in the following factor

$$\mathrm{fac}(z_i) = \frac{H(z_i)\mathrm{d}_A(z_i)}{H^{\mathrm{ref},i}(z_i)\mathrm{d}_A^{\mathrm{ref},i}(z_i)} \tag{8.29}$$

where the superscript 'ref, i' indicates that the reference cosmology is taken on the corresponding data point at redshift z_i. With this procedure, we arrive at the corrected growth theoretical prediction:

$$\mu_c^i = \frac{\mu^i_{f\sigma_8}}{\mathrm{fac}(z_i)} \, . \tag{8.30}$$

For all the datapoints, the reference model used is ΛCDM, and one can note that the product $H(z)\mathrm{d}_A(z)$ is independant of H_0 and $\sigma_{8,0}$ for all models considered. We list the reference values for Ω_{m_0} of each datapoint in Table 8.2. Using the corrected prediction, the data vector for $f\sigma_8$ is

$$\mathbf{x}_{f\sigma_8} = \mathbf{m}_{f\sigma_8} - \boldsymbol{\mu}_c \, . \tag{8.31}$$

Therefore, the chi-squared are constructed through

$$\chi^2_s = \mathbf{x}_s^T \mathbf{C}_s^{-1} \mathbf{x}_s \, , \tag{8.32}$$

where $\mathbf{C}_s^{-1}$ the inverse of the covariance matrix of the dataset. In the case of cosmic expansion, the covariance matrix is diagonal and equal to each datapoint's variance. The total cosmic growth covariance matrix is given by a diagonal matrix with the measurements' variance, with the insertion of the WiggleZ matrix and SDSS-IV matrices, given by Eqs. (8.25) and (8.26).

Parameter	Flat prior limits
Ω_{m_0}	$[0, 1]$
Ω_{de_0}	$[0, 1.7]$
w_0	$[-3.5, -1/3]$
w_a	$[-2.5, -1/3 - w_0]$
c_s^2	$[0, 1]$
H_0 [Mpc/km/s]	$[35, 110]$
$\sigma_{8,0}$	$[0.3, 1.5]$

Table 8.3: Ranges of the flat priors used for each parameter. Note that w_a depends on the value of w_0 to define its upper bound. This is to ensure that $w(a) < -1/3$ in order to have acceleration on the expansion of the Universe.

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

Evidence. The first method is the standard Bayesian model comparison via evidence computation $\log(E)$ [249], where the evidence is defined via

$$E(\mathbf{m}|\, M) = \int L(\mathbf{m}|\, \boldsymbol{\theta}_M, M)\, \pi(\boldsymbol{\theta}_M|\, M)\, \mathrm{d}\boldsymbol{\theta}_M \, . \tag{8.33}$$

The former quantity determines the probability of a given model M to be true, given the data $\mathbf{m}$. As already mentioned, the likelihood function $L(\mathbf{m}|\, \boldsymbol{\theta}_M, M)$ is Gaussian on the data $\mathbf{m}$, and the prior probability for the parameters, $\pi(\boldsymbol{\theta}_M|\, M)$. If we assume the prior probabilities $\pi(M)$ to be the same for each model, then the Evidence completely defines the ranking of the cosmological models.

All throughout the analysis, and specifically for the evidence computation, we adopted standard flat priors for all the parameters, with boundaries reported in Table 8.3. Despite that, the only special treatment was made on w_a, for which we used an upper bound that depends on the value of w_0 in order to guarantee a phase of accelerated expansion [53]. Furthermore, we use the same priors for all models that have the free parameter, as we are more concerned about the statistical methods used.

The computation is performed using the package Nestle [27], a Python implementation of the MultiNest algorithm [227, 98]. This algorithm is an efficient and robust way of computing the evidence integral, a numeric task that becomes too large to be grid-integrated. MultiNest also produces a Markov chain that can be reused as the MCMC-sample for the next method below.

Figure of Merit. With this method, the models are ranked by their FoM's defined in [29], which corresponds to the inverse of the 1σ confidence region area in the conjoined $f\sigma_8 - H$ plot given a redshift range. The likelihood is used to MCMC-sample in the parameter space of each model, and this parameter chain is used to get the 1σ range of $f\sigma_8(z_i)$ for $i \in 1, ..., n$. If there are sufficient z_i points, a spline can be constructed to connect the points in the $f\sigma_8 - H$ plane, keeping $H(z_i)$ fixed to its mean value. This method is viable because $f\sigma_8$ is much less constrained than H in all the models tested, and $H(z)$ increases monotonically with z for each model. The redshift range is, in principle, defined between $z = 0$ to $z_{max} = 2$, to include the whole redshift data range. We will also show how the FoM varies when z_{max} changes.

3-FoM. Here, we propose an extension of the previous method, in which we now consider the 1σ range of $H(z_i)$ (as opposed as in the last method where it was omitted). For a z_i point we obtain the values plus the associated confidence levels of the Hubble parameter and the growth rate, i.e.

$$H(z_i)^{+\sigma_{H(z_i)+}}_{-\sigma_{H(z_i)-}} \quad \text{and} \quad f\sigma_8(z_i)^{+\sigma_{f\sigma_8(z_i)+}}_{-\sigma_{f\sigma_8(z_i)-}} .$$

With these values we compute the ellipsoidal area on each redshift point z_i, as an approximation for the 2-dimensional confidence region in the $f\sigma_8(z_i), H(z_i)$ space.

$$A_{\mathrm{e}}(z_i) = \frac{\pi}{4}\left(\sigma_{H(z_i)+} + \sigma_{H(z_i)-}\right)\left(\sigma_{f\sigma_8(z_i)+} + \sigma_{f\sigma_8(z_i)-}\right) .$$

The 3-FoM is defined as the inverse of the ellipsoidal volume quantity in the $f\sigma_8, H, \sigma_H$ space, being

$$\begin{aligned} V_{\mathrm{e}} &= \int A_{\mathrm{e}}(z)\mathrm{d}H(z) \\ &= -\int_{z=0}^{z=2} A_{\mathrm{e}}(z)\frac{H'(z)}{(1+z)^2}\mathrm{d}z \\ &\simeq -\sum_i A_{\mathrm{e}}(z_i)\frac{H'(z_i)}{(1+z_i)^2}\Delta z . \end{aligned} \tag{8.34}$$

If there are many equispaced z_i points, the previous quantity corresponds to the volume enclosed in Fig. 8.1.

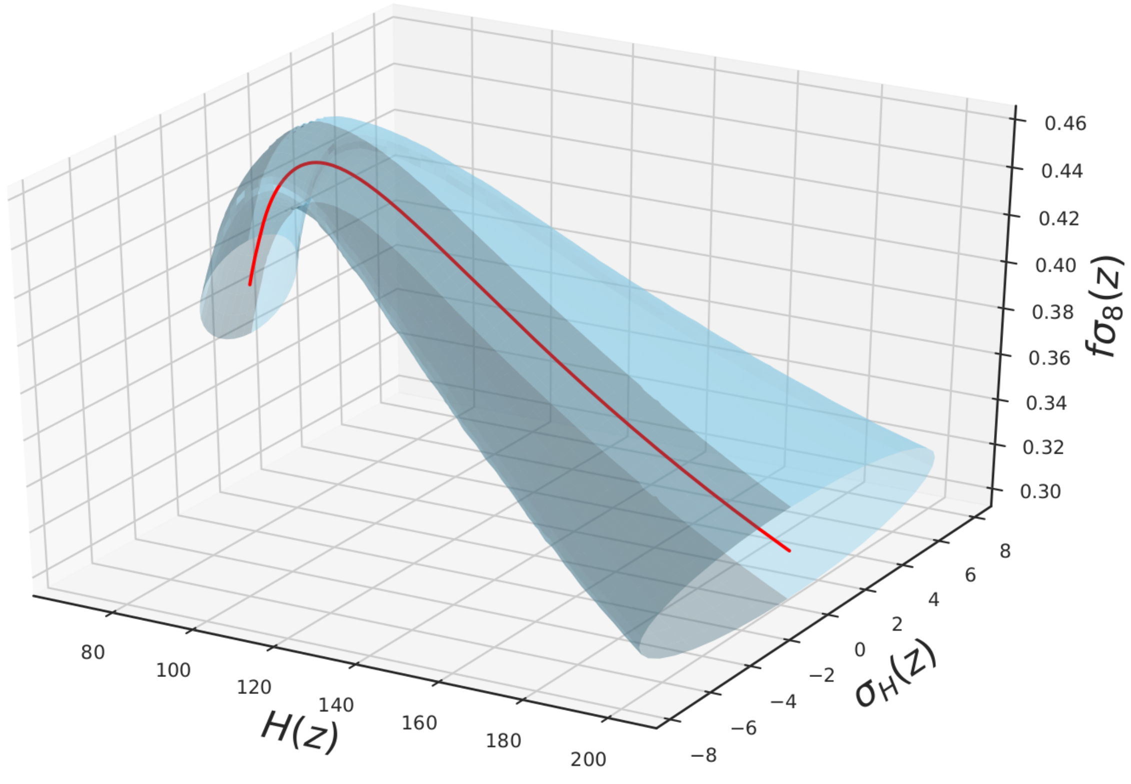


Figure 8.1: 3-FoM plot.

BIC. The fourth method is the Bayesian Information Criterion [223, 152] which is given by:

$$\mathrm{BIC} = 2\ln(N_{\mathrm{data}})n_{\mathrm{pars}} - 2\ln L_{\max}\,. \tag{8.35}$$

This method still considers the maximum likelihood $L_{\max}$, however it tends to penalize models with several parameters through the direct dependence of n_{pars} . Its formulation aims at approximating the evidence (specifically, $-2\log(E)$) of the model to be tested, hence the favored model is the one with the lowest BIC value.

AIC$_c$. The last statistical method is the corrected Akaike Information Criterion (AIC$_c$) [243]. This method is similar to the BIC method because it still penalizes models with several parameters, however the penalisation is weighted with the number of data. Contrary to the BIC test, the AIC$_c$ tends to favor one model if the data set is large enough. The criterion is given by:

$$\mathrm{AIC}_c = 2n_{\mathrm{pars}} - 2\ln L_{\max} + \frac{n_{\mathrm{pars}}(n_{\mathrm{pars}}+1)}{N_{\mathrm{data}} - n_{\mathrm{pars}} - 1}\,. \tag{8.36}$$

This equation, derived in [243], accounts for a correction term when the number of data is small, unlike the original Akaike Information Criterion [14]. As before, the test should also be similar to the value of $-2\log(E)$, which means that the lower AIC$_c$ is, the more favored is the model.

8.6 Results and Discussion

In this section, we discuss the results found for each model and we compare the values of the criteria used. As mentioned previously the goal of the paper is to accurately test the common criteria found in literature and to highlight their differences.

In general, we are not interested in the specific value of the criterion found for a particular model but rather their difference between two models. This difference will tell us which is the model that is able to better reproduce the data.

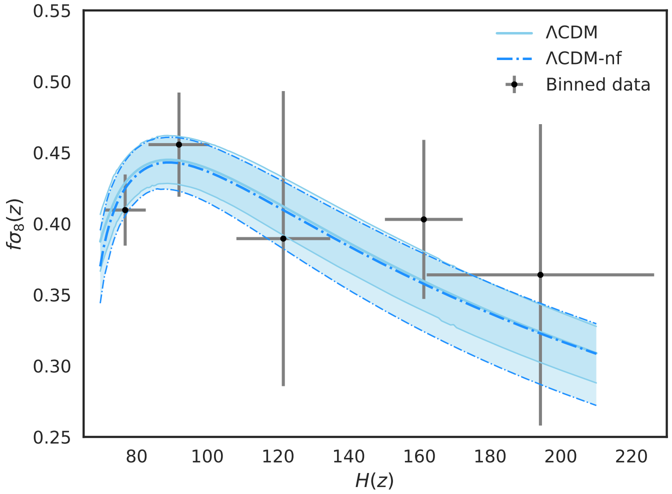
For the first criterion, i.e. the evidence E, we use Jeffrey's scale which is defined as the difference of the logarithmic evidences for two particular models, we report it in Tab. 8.6 for completeness.

The other two criteria, i.e. BIC and AIC_c, are directly connected to the likelihood of the models and hence they can be used as model selection tests. Since they come from a Taylor expansion around the maximum likelihood estimator of the likelihood function, they can be connected to Jeffrey's scale, however, this interpretation must be taken with care, see [179] for a detailed discussion. Generally, we can still consider the difference $|\Delta\mathrm{BIC}| = |\mathrm{BIC}_2 - \mathrm{BIC}_1|$, where the index 2 refers to the model with the higher value of BIC and the index 1 to the one with the lower, as a good model selection test. Specifically, if $|\Delta\mathrm{BIC}| \leq 2$, then there is no evidence in support of a model, if $2 < |\Delta\mathrm{BIC}| \leq 6$, then there is a positive evidence in favor of the model with the smaller value, whereas if $|\Delta\mathrm{BIC}| > 6$, the evidence is considered to be strong. The same discussion applies to the AIC_c criterion, where in this case we have: if the difference is less than 2, then both models are able to reproduce the data with the same accuracy, if $|\Delta\mathrm{AIC}|$ is between 2 and 4, then there is a positive evidence for the model with the lower AIC_c, instead if $|\Delta\mathrm{AIC}| > 10$, then the model with the larger AIC_c is strongly disfavored, see [204].

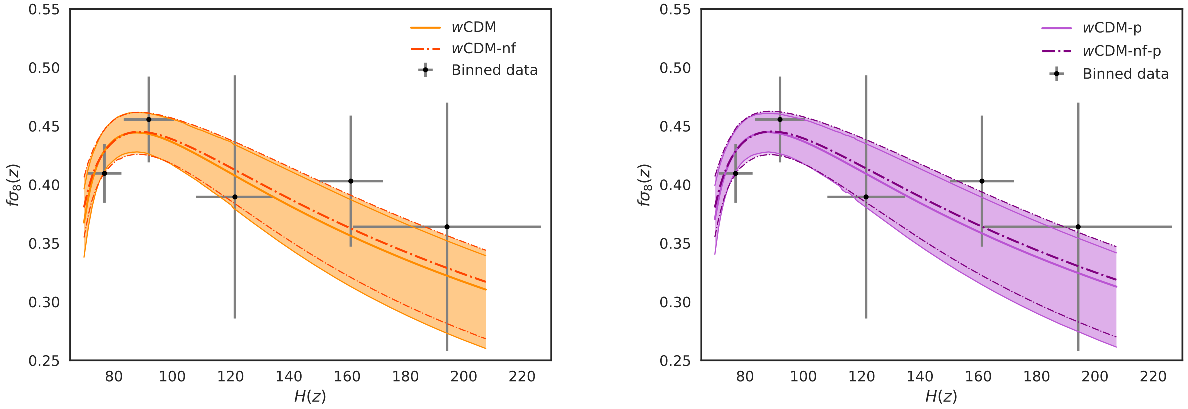
The last two criteria considered in this work are the FoM, defined as the inverse of the enclosed area at 1σ level for $f\sigma_8(z)$, and the 3-FoM defined as the inverse of the enclosed volume at 1σ level in both $f\sigma_8(z)$ and $H(z)$. It is clear the FoM and its extension (3-FoM) are not criteria able to favor/disfavor a model, but rather they give an estimation on the sensitivity of the parameters according to the data used. In practice, a larger FoM and/or 3-FoM means that the model is better constrained by the data.

Fig. 8.2a shows the reconstruction of the $H(z) - f\sigma_8(z)$ assuming flat and non-flat ΛCDM as the cosmological model. The shaded areas are obtained directly from the 1σ errors of the parameters given by the MCMC samples. The best fit of the parameters are reported in Tab. 8.7. For this particular model the addition of an extra parameter, Ω_{de_0}, alters the results and the two shaded areas differ, specially at high redshift where the lower limit of the errors are larger for not flat ΛCDM: as a consequence the FoM and 3-FoM decrease of about 35% and 125%, respectively. The BIC and AIC_c criteria used in this analysis increase of about 13% and 5% when the curvature parameter is considered, see Tab. 8.4.

As for the model comparison, the evidence gives inconclusive results, the AIC_c criterion favors positively the flat ΛCDM over the non-flat ΛCDM model, the BIC criterion instead shows a strong evidence in favor of ΛCDM.

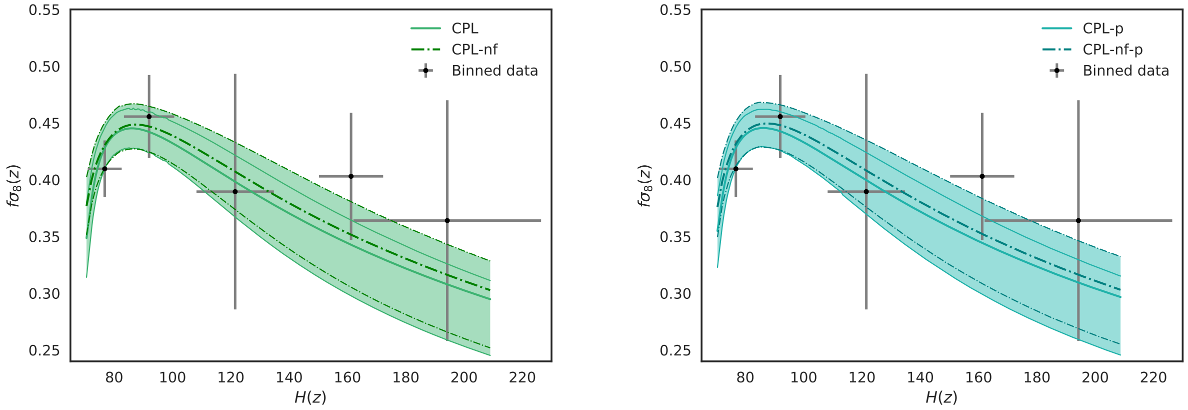


(a) ΛCDM with ΛCDM-nf.



(b) wCDM with wCDM-nf.

(c) wCDM-p with wCDM-nf-p.



(d) CPL with CPL-nf.

(e) CPL-p with CPL-p-nf.

Figure 8.2: The conjoined plots of the cosmic growth $f\sigma_8(z)$ versus the cosmic expansion $H(z)$ for differents models described in the text. Also the 1σ error regions (shaded areas) and the real binned data (gray points) are shown.

Model	$\log(E)$	FoM	3-FoM	BIC	AIC$_c$	H_{max}
ΛCDM	-21.87	0.192	0.027	51.35	34.02	201.71
ΛCDM-nf	-22.09	0.145	0.012	58.84	35.91	210.06
wCDM	-23.50	0.124	0.013	59.31	36.39	204.89
wCDM-p	-23.01	0.124	0.014	67.21	38.78	204.47
wCDM-nf	-23.33	0.125	0.010	66.57	38.14	207.54
wCDM-nf-p	-23.29	0.122	0.010	74.80	40.99	207.36
CPL	-24.16	0.129	0.014	67.18	38.75	208.97
CPL-p	-24.14	0.127	0.014	74.87	41.05	208.69
CPL-nf	-24.53	0.119	0.010	74.51	40.69	204.65
CPL-nf-p	-24.53	0.120	0.010	82.62	43.52	204.33

Table 8.4: Results of the different methods for each model. We also show $H_{max} = H(z = 2)$ to compare the extension of the integration in the H-dimension for the FoM and 3-FoM methods.

Model	$\log(E)$	FoM	3-FoM	BIC	AIC$_c$	H_{max}
ΛCDM	-8.21	0.192	0.026	23.88	6.55	202.76
ΛCDM-nf	-8.61	0.151	0.015	31.84	8.91	204.05
wCDM	-9.43	0.138	0.016	31.81	8.88	204.67
wCDM-p	-9.61	0.140	0.017	39.80	11.38	204.48
wCDM-nf	-9.76	0.130	0.013	40.06	11.64	201.14
wCDM-nf-p	-9.91	0.127	0.012	47.80	13.99	200.95
CPL	-10.63	0.134	0.015	39.85	11.42	208.74
CPL-p	-10.67	0.129	0.015	47.83	14.02	208.57
CPL-nf	-10.69	0.131	0.015	48.05	14.23	199.57
CPL-nf-p	-10.85	0.134	0.014	55.90	16.80	199.75

Table 8.5: Results of the different methods for each model using the Mock Catalogue. We also show $H_{max} = H(z = 2)$ to compare the extension of the integration in the H-dimension for the FoM and 3-FoM methods.

$\lvert\Delta\log(E)\rvert$	Probability	Evidence
$0 \leq \lvert\Delta\log(E)\rvert < 1.0$	$0 \leq P_1 < 0.75$	Inconclusive
$1.0 \leq \lvert\Delta\log(E)\rvert < 2.5$	$0.75 \leq P_1 < 0.923$	Weak
$2.5 \leq \lvert\Delta\log(E)\rvert < 5.0$	$0.923 \leq P_1 < 0.993$	Moderate
$5.0 \leq \lvert\Delta\log(E)\rvert$	$0.993 \leq P_1$	Strong

Table 8.6: Jeffrey's Scale as in Ref. [250], which compares the logarithmic Evidence difference between the two models. The different levels represent different degrees of belief in that one is the true theory.

In Fig. 8.2b are shown the reconstruction of the $H(z) - f\sigma_8(z)$ assuming flat and non-flat wCDM. These models have one parameter more with respect to the corresponding ΛCDM models discussed above. The addition of w as a free parameter increases the confidence regions substantially, as it can be seen from the figures and also reported in Tab. 8.4, where the FoM decreases compared to previous cases. Here the variation is due to the parameter itself rather than the addition of an extra parameter; in fact, if we consider the non-flat ΛCDM model, which has the same number of parameters as wCDM model, the FoM reduces from 0.145 to 0.124 which corresponds to almost 15%. However, the 3-FoM manifests an opposite behavior, it increases of about 8%. The reason is that $f\sigma_8(z)$ is sensitive to the variation of the parameters almost at any redshift, whereas the Hubble parameter is more sensitive at high redshifts (fixing one value of H_0, the variation on $H(z)$ can only appear when the z is increased). For the non-flat ΛCDM model the area enclosed by $f\sigma_8(z)$ is smaller than the era enclosed for the wCDM model, hence giving a lager FoM. However, the maximum value of the Hubble parameter is larger for non-flat ΛCDM model, 210.06 against 204.89 for the wCDM model. This effect is taken into account in the 3-FoM, where the errors on $H(z)$ are considered. The two effects are counterbalanced, giving almost the same value in the 3-FoM.

The evidence is weakly in support of the non-flat ΛCDM model over wCDM and the same is found for the BIC and AIC$_c$ criteria. Adding curvature to the wCDM model makes the FoM increase of about 1% meaning that the 1σ errors are almost the same, however, their best-fits differ. The 3-FoM decreases for the non-flat model showing that their errors are less constrained. The BIC and AIC$_c$ supports the flat model but the evidence is inconclusive.

In Fig. 8.2c we show the reconstruction of the $H(z) - f\sigma_8(z)$ assuming flat and non-flat wCDM with the further addition of perturbations in the dark energy sector parameterized with c_s^2 as an extra free parameter. If we compare the latest results with the former case we realize that the FoM does not change from wCDM to wCDM-p, whereas it decreases of about 2.4% from wCDM-nf to wCDM-nf-p . These negligible variations are repeated for the 3-FoM that does not change from wCDM-nf to wCDM-nf-p and it increases of about 7.1% from flat wCDM to wCDM-p. As expected, dark energy perturbations are weakly constrained with the data available (dark energy perturbations affect only the growth of matter). This is shown in Tab. 8.7 where the best fits of the models with and without dark energy perturbations are basically the same. This behavior is shown in all the criteria used in this work, except for the BIC criterion which indeed favors the model without dark energy perturbations. However, this is a pure mathematical effect as the BIC criterion always penalizes the model with extra parameters.

In Fig. 8.2d are shown the reconstructions of the $H - f\sigma_8(z)$ assuming flat and non-flat CPL. If we look at Tab. 8.4, we realized that the FoM constrains better CPL over wCDM which might sound peculiar because one would naively expect that a model with more parameters has larger 1σ errors. Here, the difference in the FoM comes from the asymmetric values of the errors on w_a; this asymmetry is due to the choice of the prior for w_a, for which we chose to bind it to w_0 in order to guarantee an accelerated expansion. This asymmetry led to a smaller area in the upper part, reducing the enclosed 1σ area of $f\sigma_8(z)$. The 3-FoM is more stable and this is again due to the value of the Hubble parameter at high redshifts: for the wCDM model $H(z = 2) = 204.89$, whereas for CPL model is 208.97. This

2% difference is accounted in the final 3-FoM which decreases with respect to its companion. The evidence gives inconclusive results, manifesting the negligible effects of w_a on the two observables; the same conclusion is obtained with the AIC$_c$ criterion. However, the BIC criterion strongly penalizes CPL just because of the extra parameter in the model.

The CPL, wCDM-nf, and wCDM-nf models have the same number of parameters, thus BIC and AIC$_c$ criteria change less than 1% between them, showing again that they depend strongly on the number of parameters. The FoM and 3-FoM show that CPL is better constrained but, as mentioned, this is due to the priors on w_a used. The evidence weakly favors the wCDM-p model over CPL but it is inconclusive with respect of wCDM-nf.

The non-flat CPL have the same number of parameters as wCDM-nf-p, but the FoM shows that CPL is better constrained by the data whereas the 3-FoM does not change. The BIC and AIC$_c$ change less than 1% and the evidence weakly favors wCDM-nf-p model.

In Fig. 8.2e we show the reconstruction of the $H - f\sigma_8(z)$ assuming flat and non-flat CPL-p models. The behavior is similar to the previous case (CPL versus CPL-nf). By adding the curvature parameter the evidence is inconclusive and the other indicators favor the flat model because it has one parameter less. When we take into account dark energy perturbations into the CPL models, we obtain a similar behavior as seen for the wCDM models. Again, with the available data we are not able to constrain c_s^2, hence all the criteria are insensitive to the variation of the sound speed. The only exceptions are BIC and AIC$_c$ criteria, which penalize the addition of the sound speed into the analysis.

For completeness we also performed our analysis using the analytical solutions for the growth rate of matter, the models are ΛCDM, wCDM and wCDM with dark energy perturbations. The results are reported in the Appendix 8.7 and the results are shown in Tab. 8.9, whereas the best fit of these three models can be found in Tab. 8.10. All the three analytical models give results in excellent agreement with the full numerical analysis, demonstrating that the analytical solutions found in the literature are consistent and they can be safely used.

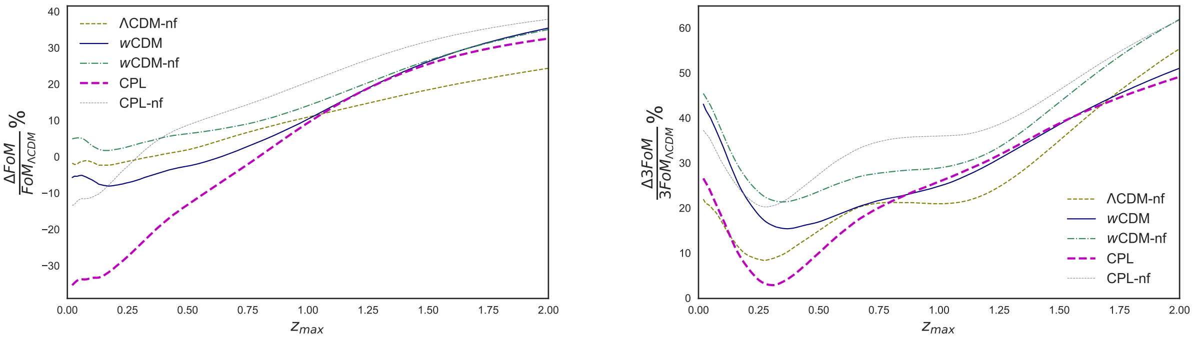


Figure 8.3: These figures show the percentage difference of FoM (upper panel) or 3-FoM (lower panel) between a model and ΛCDM. We only present the models without perturbations in the dark sector. Here, $\Delta\mathrm{FoM} = \mathrm{FoM}_{\Lambda\mathrm{CDM}} - \mathrm{FoM}_{\mathrm{model}}$ and likewise for the 3-FoM.

Model	Ω_{m_0}	Ω_{de_0}	w_0	w_a	c_s^2	H_0	σ_8
ΛCDM	$0.286^{+0.032}_{-0.038}$	$\mathbf{1-\Omega_{m_0}}$	-	-	-	69.7 ± 2.3	0.779 ± 0.039
ΛCDM-nf	0.37 ± 0.16	$0.83^{+0.29}_{-0.24}$	-	-	-	70.4 ± 3.1	$0.762^{+0.044}_{-0.084}$
wCDM	$0.280^{+0.045}_{-0.039}$	$\mathbf{1-\Omega_{m_0}}$	$-1.11^{+0.38}_{-0.30}$	-	-	$70.6^{+4.2}_{-4.7}$	$0.782^{+0.045}_{-0.11}$
wCDM-p	$0.278^{+0.044}_{-0.037}$	$\mathbf{1-\Omega_{m_0}}$	$-1.09^{+0.38}_{-0.30}$	-	0.50 ± 0.29	70.5 ± 4.4	$0.788^{+0.045}_{-0.11}$
wCDM-nf	$0.34^{+0.18}_{-0.22}$	$0.86^{+0.34}_{-0.41}$	$-1.08^{+0.49}_{-0.18}$	-	-	69.7 ± 4.2	$0.790^{+0.045}_{-0.11}$
wCDM-nf-p	$0.34^{+0.18}_{-0.22}$	$0.85^{+0.34}_{-0.42}$	$-1.07^{+0.50}_{-0.16}$	-	0.51 ± 0.29	$69.6^{+3.7}_{-4.7}$	$0.795^{+0.048}_{-0.11}$
CPL	$0.294^{+0.047}_{-0.041}$	$\mathbf{1-\Omega_{m_0}}$	-1.20 ± 0.34	$-0.50^{+0.99}_{-0.46}$	-	71.9 ± 4.5	$0.747^{+0.026}_{-0.099}$
CPL-p	$0.293^{+0.046}_{-0.041}$	$\mathbf{1-\Omega_{m_0}}$	-1.17 ± 0.33	$-0.50^{+1.0}_{-0.49}$	0.50 ± 0.29	71.7 ± 4.4	$0.751^{+0.030}_{-0.10}$
CPL-nf	$0.27^{+0.12}_{-0.24}$	$0.72^{+0.21}_{-0.45}$	$-1.27^{+0.63}_{-0.27}$	$-0.42^{+1.0}_{-0.47}$	-	$70.5^{+4.0}_{-4.6}$	$0.778^{+0.048}_{-0.11}$
CPL-nf-p	$0.27^{+0.11}_{-0.26}$	$0.72^{+0.23}_{-0.45}$	$-1.28^{+0.65}_{-0.27}$	$-0.38^{+0.98}_{-0.45}$	0.50 ± 0.29	$70.4^{+3.9}_{-4.5}$	$0.778^{+0.054}_{-0.10}$

Table 8.7: Parameter constraints derived from Nested Sampling to each (non-analytical) model described in the text.

8.7 Conclusions

In our work we implemented the conjoined $H(z) - f\sigma_8(z)$ method in order to test an entire family of ten dark energy models; we started with the simplest model, ΛCDM which is described by three parameters only, and we systematically increased the level of complexity of the model by adding extra parameters, being the non-flat CPL with dark energy perturbation the most complex model (with seven parameters).

For each model, we first found the best fit using MCMC analysis by combining the most recent cosmic chronometer and growth data available. Subsequently, we compared the dark energy models with five different statistical criteria, aiming at highlighting the potentiality and the weakness of each criterion.

As expected, we found that the evidence is the most accurate statistical test to compare different models as it takes into account the information of the entire likelihood of the parameters and it does not always penalize a model with extra parameters. The 3-FoM better characterizes the sensitivity of the parameters according to the data used. This criterion takes into account simultaneously the errors from both $f\sigma_8(z)$ and $H(z)$; in particular, we showed that the errors of the Hubble parameter increase with redshift and this has an important effect on the constraining power of the test. The FoM instead is limited only to $f\sigma_8(z)$, hence neglecting the information from $H(z)$, which might be crucial if the analysis is extended at high redshift. As a complementary test, we performed the same analysis in the same redshift range as in [29] and we found consistent results.

For the last two criteria, BIC and AIC$_c$, we showed that they always penalize the addition of extra parameters; in fact, if we consider the two extreme models, i.e. ΛCDM with only three parameters and non-flat CPL with dark energy perturbations, which has seven parameters, we find that ΔBIC ~ 40 manifesting a *very* strong evidence in favor of the ΛCDM model. Similarly, but less decisive is ΔAIC$_c$ for which we find a value of ~ 10, which still favors strongly ΛCDM but more moderately than BIC.

To demonstrate the power of the 3-FoM, we compute the FoM and 3-FoM at different redshifts starting from $z = 0$ up to the z_{max}. These results are shown in Fig. 8.3 where we plotted the relative difference of the FoM (top panel) and the 3-FoM (lower panel) for each model with respect to ΛCDM. It is interesting to notice that at low redshifts the FoM for wCDM, wCDM-nf, CPL, and CPL-nf is larger than ΛCDM, meaning that the former is better constrained than the latter. This effect is not manifested in the 3-FoM which is always larger for the ΛCDM model.

Redshift bin	$H(z)$ [km s^{-1} Mpc^{-1}]	$f\sigma_8(z)$
$0 < z \leq 0.4$	76.8 ± 5.8	0.410 ± 0.025
$0.4 < z \leq 0.8$	92.0 ± 8.6	0.456 ± 0.037
$0.8 < z \leq 0.12$	121.5 ± 13.3	0.390 ± 0.104
$0.12 < z \leq 0.16$	161.2 ± 11.0	0.404 ± 0.056
$1.16 < z \leq 1.2$	194.2 ± 32.2	0.364 ± 0.106

Table 8.8: Binned measurements of $H(z)$ and $f\sigma_8(z)$ with equispaced redshifts points and its uncertainties. These are the gray points shown in Fig. 8.2.

Acknowledgments

BS acknowledges help from CONICYT. JS and DS acknowledge financial support from the Fondecyt project number 11140496. We also thank Savvas Nesseris for useful discussions.

Appendix

Comparison with analytical solutions

The second order differential equation for the matter density contrast at small scales, without dark energy perturbations is given by, [51]

$$a^2\delta_m'' + (3 - \epsilon(a))\, a\delta_m' - \frac{3}{2}\Omega_m(a)\delta_m(a) = 0 \ ,$$

with $\epsilon(a) = -\mathrm{d}\log H(a)/\mathrm{d}\log a$. As we are describing late time solutions we will always take the growing mode solution given by [51]

$$\delta(a) = a_2F_1\left(\frac{w-1}{2w}, -\frac{1}{3w}, 1 - \frac{5}{6w}, 1 - \Omega_m^{-1}(a)\right) ,$$

where we omitted the integration constant because it will cancel out when we evaluate $f(a)$. The result to ΛCDM is given by setting $w = -1$.

There exist analytical solution for the matter density contrast when dark energy perturbations are included, see [180] for mode details. The joint solution for the density contrast is given by

$$\begin{aligned}\delta(a) = a_2F_1\biggl(\frac{1}{4} &- \frac{5}{12w} + B, \frac{1}{4} - \frac{5}{12w} - B, \\ 1 &- \frac{5}{6w}, 1 - \Omega_m^{-1}(a)\biggr) ,\end{aligned}$$

where B is used as B_{joint} in [180], which corresponds to:

$$B = \frac{1}{12w}\sqrt{(1-3w)^2 + 24\frac{1+w}{1-3w+\frac{2}{3}\frac{k^2c_s^2}{H_0^2\Omega_{m_0}}}}\,. \tag{8.37}$$

Model	$\log(E)$	FoM	3FoM	BIC	AIC$_c$	H_{max}
ΛCDM-a	-22.07	0.192	0.027	51.34	34.01	202.05
wCDM-a	-23.28	0.125	0.014	59.32	36.39	204.80
wCDM-p-a	-23.23	0.126	0.014	67.22	38.79	204.66

Table 8.9: Results of the different methods for each analytic model. These are almost equal to their numerical versions.

	ΛCDM-a	wCDM-a	wCDM-p-a
Ω_{m_0}	$0.286^{+0.033}_{-0.038}$	$0.281^{+0.044}_{-0.039}$	0.281 ± 0.044
w_0	-	$-1.10^{+0.36}_{-0.31}$	$-1.10^{+0.36}_{-0.30}$
c_s^2	-	-	0.50 ± 0.29
H_0	69.8 ± 2.4	70.4 ± 4.4	$70.5^{+4.1}_{-4.7}$
σ_8	0.780 ± 0.040	$0.784^{+0.042}_{-0.11}$	$0.784^{+0.044}_{-0.11}$

Table 8.10: Parameter constraints derived from Nested Sampling to each analytical model described in the text.

Chapter 9

Cosmological constraints from galaxy multi-tracers in the nearby Universe

This chapter presents a research work which corresponds to [92]. This work is a team effort with aid from Dr. Ginevra Favole affiliated by that time to the Institute of Cosmology & Gravitation at the University of Portsmouth in UK, and Dr. Domenico Sapone affiliated to the Cosmology and Theoretical Astrophysics group at the Facultad de Ciencias Fśicas y Matemáticas of the Universidad de Chile in Chile.

Abstract

The Baryon Acoustic Oscillation (BAO) scale in the clustering of galaxies is a powerful standard ruler to measure cosmological distances and determine the geometry of the Universe. Past surveys have detected the BAO feature in the clustering of different galaxy samples, most of them composed of redder, quiescent galaxies and bluer, star-forming ones out to redshift $z \sim 1$. Besides these targets, new upcoming surveys will observe high-redshift galaxies with bright nebular emission lines out to $z \sim 2$, quasars and Lyman-α quasars at $z > 2$. All these different galaxy targets will be used as multi-tracers of the same underlying dark matter field. By combining them over wide cosmological volumes, we will be able to beat cosmic variance and measure the growth of structure with unprecedented accuracy. In this work, we measure the BAO scale in the two-point auto- and cross-correlation functions of three independent populations of multi-tracers extracted from the SDSS DR7 Main galaxy sample at redshift $0.02 < z < 0.22$. Combining their covariances, we find accurate constraints on the shift parameter $\alpha = 1.00 \pm 0.04$ and $D_{\mathrm{V}}(z = 0.1)/r_{\mathrm{s}} = 2.92 \pm 0.12$.

9.1 Introduction

During the last decades, observations have led to the general acceptance that the Universe is in a phase of accelerated expansion. In a homogeneous and isotropic Universe, the simplest way to account for such expansion is to introduce a constant term in the Einstein equations, dubbed as cosmological constant (Λ). Based on recent observations [10, 86, 41] and on the simplicity of the model, the cosmological constant is still the most accepted dark energy candidate responsible for the acceleration of the Universe. Despite of its simplicity, such a scenario has raised several theoretical issues, which have led cosmologists to invoke more sophisticated dark energy models without succeeding on the task, see [218]. An alternative approach suggests that we would need to modify the laws of gravity and make it weaker at larger scales to mimic, phenomenologically, the observed expansion [144, 251].

The fundamental observables that trace the dynamics of the Universe are the Hubble parameter $H(z)$ and the angular diameter distance $D_\mathrm{A}(z)$, which are directly connected to the properties of matter and quantify the overall expansion of the Universe. Observational exploration is necessary to provide an indication about the dynamics of the Universe. One way of understanding this is to measure distances at different epochs. Modern cosmology has been revolutionised when the definition of *standard ruler* [226] was introduced: a distance scale in the Universe whose size and evolution with redshift are known. An ideal candidate is the Baryon Acoustic Oscillation scale (BAO; [86]) observed at the last scattering surface in the Cosmic Microwave Background (CMB) radiation. This feature represents the width of the primordial density fluctuations that propagate as acoustic waves in the early baryon-photon fluid. Such a distance can be decomposed into a radial, $H(z)$, and a transverse, $D_A(z)$, direction, which allow us to measure the expansion history of the Universe [109]. If the standard cosmological model, i.e. the structures that we see today, have been generated by gravitational collapse of the primordial seeds in an expanding, homogeneous and isotropic Universe, then we should see an excess of baryonic matter in the distribution of galaxies at the same comoving scale. This excess of baryonic matter is visible as a prominent peak around $110\,h^{-1}$Mpc in the galaxy two-point correlation function. The first detection of the BAO peak happened in SDSS [86] and was then confirmed by 2dFGRS [57], BOSS [72] and WiggleZ Dark Energy Survey [41], VIPERS [75] and eBOSS [73].

New upcoming surveys, such as the Dark Energy Spectroscopic Instrument (DESI) [9, 8], Euclid [149, 24], Subaru Prime Focus Spectrograph (PFS) [242, 232], the Large Synoptic Survey Telescope (LSST) [4], or the Wide Field Infrared Survey Telescope (WFIRST) [109, 236], will observe tens of hundreds of millions of galaxies positions and spectra covering enormous cosmological volumes and extend the observations at very high redshifts ($z \sim 2{-}3$). These observations will map the late time dynamics of the Universe with unprecedented precision (few percents on the final cosmological parameters [44]). It is therefore imperative to gain as much information as possible from these data sets.

One statistical limitation of measuring the cosmological parameters is due to the cosmic variance effects in the survey volume. Recently, it was shown [5, 173] that by cross-correlating different dark matter tracers over wide cosmological volumes it is possible to beat cosmic variance, dramatically reducing the uncertainties on the observables. In fact, while the effective volume still remains a limitation, the relative information between different species is not. In this *Letter*, we show that using a multi-tracer approach we can lower the errors on the scale distortion parameter α. In particular, we use luminous red galaxies and emission line galaxies as tracers to map the dynamics of the Universe at $z \sim 0.1$. Throughout this work, We adopt a Planck cosmology [6] with $\Omega_m = 0.3071$, $\Omega_\Lambda = 0.6929$, $h = 0.6777$, $n = 0.96$ and $\sigma_8 = 0.8228$.

9.2 Data

We analyse three independent galaxy populations selected from the SDSS DR7 Main galaxy sample [241], each one composed of a different tracer, and all of them covering the redshift range $0.02 < z < 0.22$. Specifically, these samples are: two selections of emission line galaxies (ELGs), one of [O II] [91] and another one of Hα emitters [89], plus a selection of luminous red galaxies (LRGs) [85]. These are the only galaxy multi-tracers currently available in the nearby Universe. The SDSS Main parent sample, which is brighter than $r = 17.77$, and covers and effective area of $7300\,\mathrm{deg}^2$ [114], was extracted from the NYU-Value Added Galaxy Catalogue[1] [45] and it was spectroscopically matched (i.e. matching the redshifts) to the MPA-JHU[2] DR7 release of spectral measurements to assign emission line properties.

We consider only ELGs with well measured spectra, i.e. those with the flag `ZWARNING = 0`, and with good signal-to-noise, i.e. S/N > 5. Both [O II] and Hα ELG samples have specific star formation rate of

$$\log\left(\mathrm{sSFR/M_{\odot}yr^{-1}}\right) > -11\,,$$

and line equivalent width of

$$\mathrm{EW} > 10\,\mathrm{\AA}\,,$$

to guarantee that we are selecting only very star-forming galaxies. In addition, they are both limited in flux at $2 \times 10^{-16}\mathrm{erg\,cm^{-2}\,s^{-1}}$ to match the Euclid nominal expected depth and flux limit [166] at higher redshift. The LRG sample includes only galaxies with

$$\log\left(\mathrm{sSFR/M_{\odot}yr^{-1}}\right) < -11\,,$$

which are all quiescent.

The observed (i.e. attenuated by dust) [O II] and Hα ELG luminosities are computed from the corresponding flux densities F as

$$L\,[\mathrm{erg\ s^{-1}}] = 4\pi D_{\mathrm{L}}^2 10^{-0.4(r_{\mathrm{P}}-r_{\mathrm{fib}})} F\,,$$

where D_{L} is the luminosity distance as a function of cosmology and the exponent is the SDSS fibre aperture correction written in terms of the r-band petrosian and fibre magnitudes. For further details on the luminosity calculations, we refer the reader to [91, 89].

Fig. 9.1 shows the galaxy number density of the SDSS Hα, [O II] and LRG samples as a function of redshift. Compared to the SDSS Main sample in [210], our galaxy selections cover a larger area (we consider both North and South Galactic Caps) and span a slightly different redshift range. Therefore, it is not surprising that our galaxy number densities differ from [210]. In particular, our $n(z)$ values are large enough to ensure that the SDSS Hα, [O II] and LRG samples are limited by cosmic variance (i.e., $n(z)\,P(k) > 1$) at $z < 0.22$ for $k < 0.095$, 0.093, $0.071\,h\,\mathrm{Mpc}^{-1}$, respectively.

9.3 Measurements

In Large Scale Structure (LSS) analysis, galaxies can be thought as point-like objects in space-time that move along with the expansion of the Universe, forming bounded structures

---

[1] `http://cosmo.nyu.edu/blanton/vagc/`

[2] `http://www.mppg.de/SDSS/DR7/`

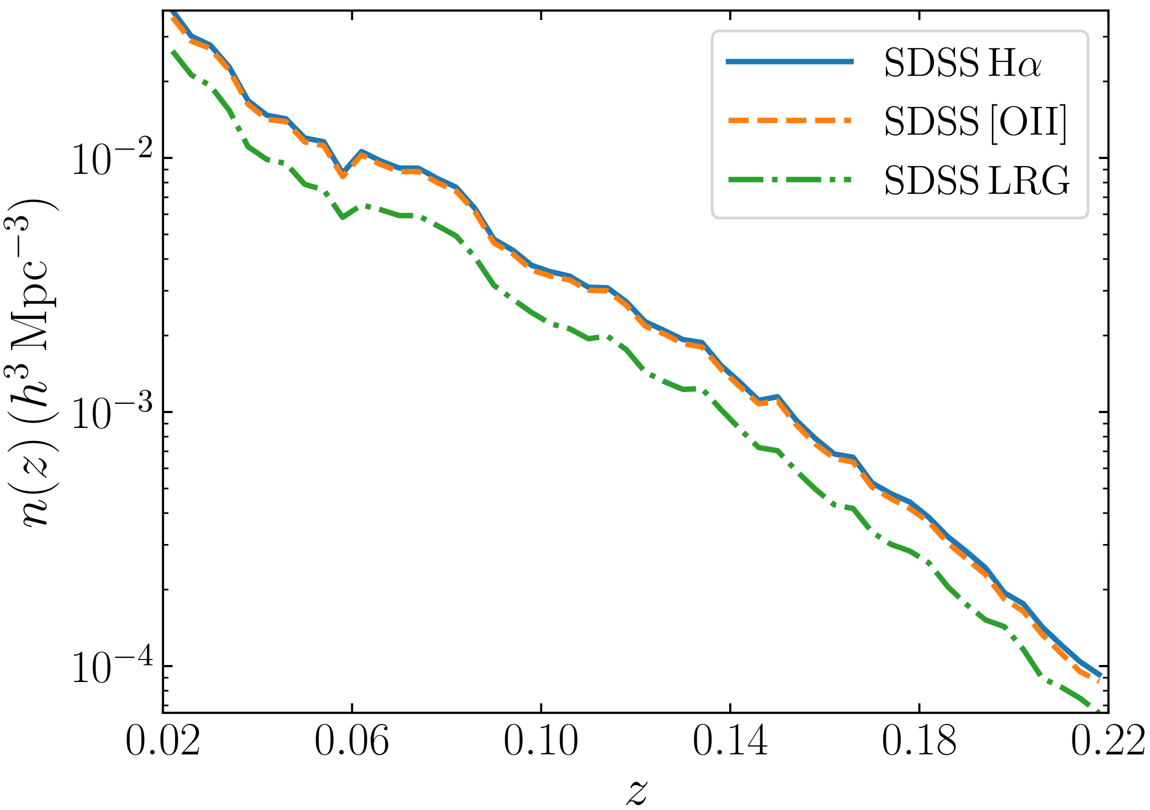


Figure 9.1: Galaxy number density, as a function of redshift, of the SDSS Hα, [O II] and LRG samples at $0.02 < z < 0.22$.

due to their gravitational interaction. By analysing their positions, we can gain information on the underlying theory of gravity. One method consists in quantifying how many objects are present in a given cosmological volume. This method relies on the two-point correlation function (2PCF), which is the excess probability over randoms of finding two galaxies separated by a distance s in redshift-space.

We measure the two-point auto- and cross-correlation functions of the three galaxy tracers defined in Sec. 9.2 using the Landy and Szalay estimator:

$$\xi_{\mu\nu}(s) = \frac{D_\mu D_\nu(s) - D_\mu R_\nu(s) - D_\nu R_\mu(s)}{R_\mu R_\nu(s)} + 1\,,$$

where $s = \sqrt{\pi^2 + r_\mathrm{p}^2}$ represents the redshift-space distance as a function of the parallel (π) and perpendicular (r_p) components to the line of sight, while μ and ν are the tracers. The DD, DR and RR terms are the normalised and weighted data-data, data-random and random-random pair counts formed from the observed galaxies and the synthetic randoms. We use the equal surface density randoms from the NYU-VAGC. The weighting scheme adopted for the pair counts is $w = w_\mathrm{fc} w_\mathrm{ang} w_\mathrm{FKP}$ for data and $w = w_\mathrm{FKP}$ for randoms. The w_fc term represents the fibre collision weight (in SDSS fibres cannot be placed closer than

55"). The angular weight, $w_{\rm ang} = 1/f_{\rm got}$, accounts for the angular sector completeness, and the FKP weight [96], $w_{\rm FKP} = 1/(1+\bar{n}(z)P_0)$, corrects for any fluctuation in the number density of tracers. Here, $\bar{n}(z)$ is the expected number density of a galaxy at redshift z and we set $P_0 = 16000\,h^{-3}{\rm Mpc}^3$, which is close to the amplitude of the SDSS power spectrum at $k = 0.1h{\rm Mpc}^{-1}$ [210].

We estimate the uncertainties on the SDSS clustering measurements via 200 jackknife resamplings [167, 183, 182, 115, 90] containing about the same number of data (randoms) each. The area of the individual jackknife region is $A = 36.5\,{\rm deg}^2$, meaning that its comoving size at z=0.1 in Planck cosmology [6] is $S_{\rm JK} \sim 28\,h^{-1}{\rm Mpc}$. As demonstrated by [93], it is not necessary to design jackknife cells larger than the maximum 2PCF scale that we measure to have accurate covariance estimates. Instead, it is better to prioritise a high number of resamplings to reduce the noise in the covariances.

The jackknife covariance matrix for each SDSS galaxy clustering measurement is calculated as [82, 183, 90]:

$$\hat{C}_{ij} = \frac{N_{\rm res}-1}{N_{\rm res}} \sum_{a=1}^{N_{\rm res}} (\xi_i^a - \bar{\xi}_i^a)(\xi_j^a - \bar{\xi}_j^a)\,, \tag{9.1}$$

where the pre-factor takes into account that in every resampling $N_{\rm res} - 2$ sub-volumes are the same [182], and $\bar{\xi}_i$ is the mean jackknife correlation function in the $i^{\rm th}$ bin:

$$\bar{\xi}_i = \sum_{a=1}^{N_{\rm res}} \xi_i^a / N_{\rm res}\,.$$

The full covariance matrix for all the tracers is built by combining the individual ones in Eq. (9.1) as:

$$\hat{C} = \begin{pmatrix} \hat{C}_{\rm H\alpha-H\alpha} & \hat{C}_{\rm H\alpha-[OII]} & \hat{C}_{\rm H\alpha-LRG} \\ \hat{C}_{\rm [OII]-H\alpha} & \hat{C}_{\rm [OII]-[OII]} & \hat{C}_{\rm [OII]-LRG} \\ \hat{C}_{\rm LRG-H\alpha} & \hat{C}_{LRG-{\rm [OII]}} & \hat{C}_{\rm LRG-LRG} \end{pmatrix}\,. \tag{9.2}$$

The inverse of the covariance matrix, the so-called “precision matrix" $\hat{\Psi} \equiv \hat{C}^{-1}$, requires some corrections. In fact, Eq. (9.2) is obtained from a limited set of resamplings, $N_{\rm res}$, and it has an associated error which propagates into the precision matrix. Following [192], we implement two corrections to obtain an unbiased estimate of the precision matrix and to reduce the noise in its off-diagonal terms. The bias correction consists in multiplying $\hat{C}^{-1}$ by the Hartlap factor [119], which accounts for the limited number of resamplings and the number of bins $n_{\rm b}$ in our measurements of ξ, **i.e** $\hat{\Psi} = [1 - (n_{\rm b}+1)/(N_{\rm res}-1)]\,\hat{C}^{-1}$.

The noise correction, also known as “covariance tapering" [138], can be applied to both covariance and precision matrices. The idea of this method is to neglect the correlation between data pairs far apart through a kernel function of the Matérn class. Such a correction relies on the tapering matrix $T_{ij} \equiv K\left(||s_i - s_j||\right)$, which is defined as a monoparametric Kernel function [192, 260, 261] depending on the physical scale of the tracers that we are correlating. This kernel also includes a tapering parameter $T_{\rm p}$, which identifies the interval where $K(x)$ takes non-zero values, guaranteeing a vanishing correlation between pairs for larger distances. The final corrected precision matrix is:

$$\hat{\Psi} = \left(1 - \frac{n_{\rm b}+1}{N_{\rm res}-1}\right)\left(\hat{C}\circ T\right)^{-1}\circ T\,, \tag{9.3}$$

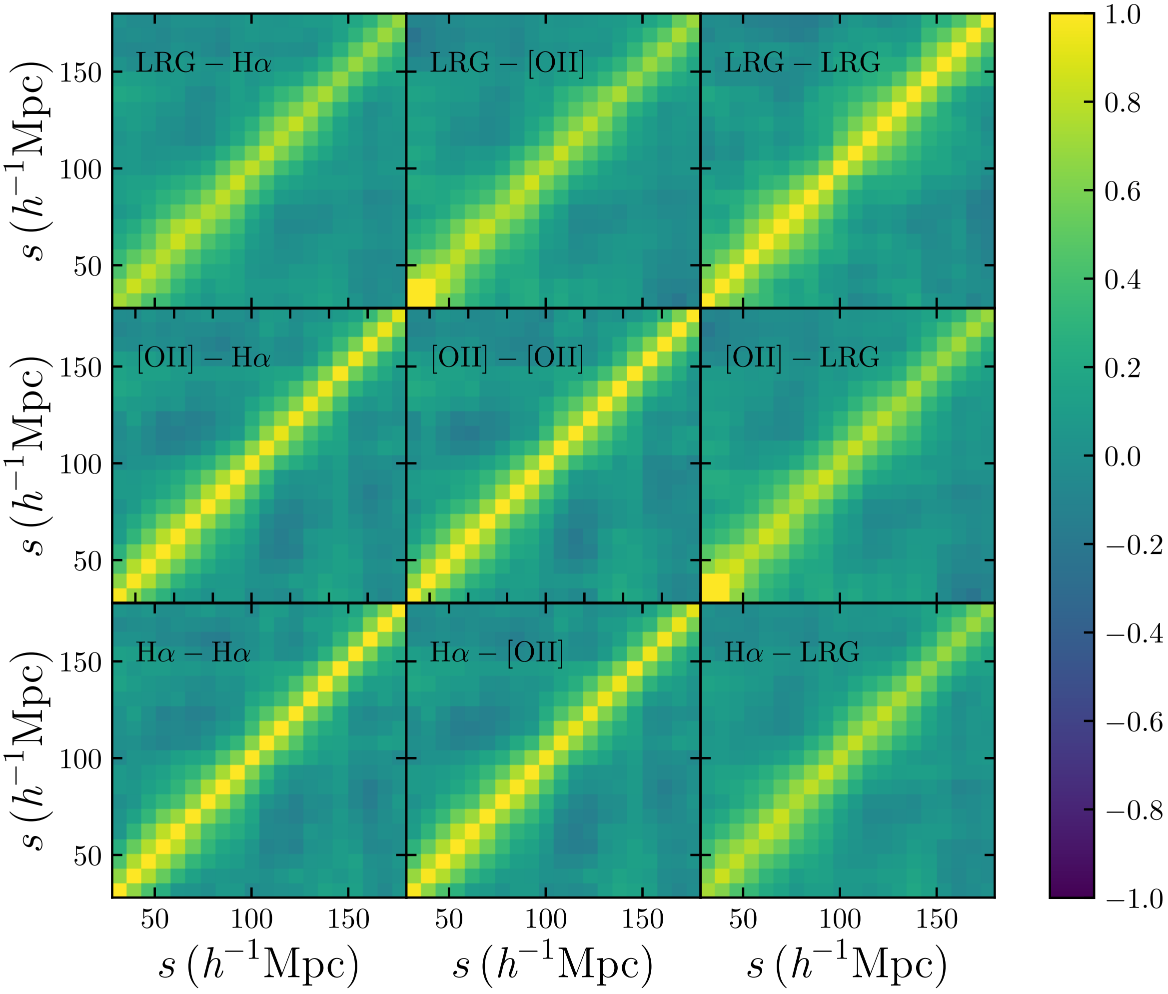


Figure 9.2: Normalised covariance matrix obtained from the monopole auto- and cross-correlation functions of the three SDSS galaxy tracers.

where the ∘ symbol indicates the Hadamard product. In our analysis we assume a tapering parameter $T_\mathrm{p} = 50\,h^{-1}\mathrm{Mpc}$ to ensure that the entire covariance matrix is positive semi-definite. In Fig. 9.2, we present the normalised covariance matrix obtained from the monopole auto- and cross-correlation functions of the SDSS Hα, [O II] and LRG tracers. The normalisation is computed as $C_\mathrm{ij}^\mathrm{norm} = C_\mathrm{ij}/\sqrt{C_\mathrm{ii}\,C_\mathrm{jj}}$, where Cij is given by Eq. 9.1.

9.4 Methodology

The real-space two-point correlation function $\xi(r)$ is the spatially isotropic Fourier transform of the matter power spectrum $P(k)$ defined as:

$$\xi(r) = \frac{1}{2\pi^2}\int P(k)\frac{\sin(kr)}{kr}k^2\mathrm{d}k\,. \tag{9.4}$$

The position of the BAO peak inferred from Eq. (9.4) is expected to appear around $110\,h^{-1}$Mpc, which is well beyond the scales of virialised objects. This implies that the non-linear gravitational effects can be safely ignored. For the power spectrum we use the template [185]:

$$P(k) = [P_{\rm lin}(k) - P_{\rm dw}(k)]\, {\rm e}^{-k^2\Sigma_{\rm nl}^2/2} + P_{\rm dw}(k)\,, \tag{9.5}$$

where $P_{\rm lin}(k)$ is the linear matter power spectrum calculated using the Boltzmann code CLASS [151], and $P_{\rm dw}(k)$ is the de-wiggled power spectrum [83], both using Planck 2015 [6] fiducial cosmology. The $\Sigma_{\rm nl}$ parameter accounts for the smoothing of the BAO peak due to non-linear effects [65].

We compute the theoretical correlation functions needed to fit the SDSS multi-tracer measurements by applying Eq. (9.4) with the power spectrum given in Eq. (9.5). We model the BAO signal as [266]:

$$\xi_{\rm model}(s) = B\xi(\alpha s) + a_0 + \frac{a_1}{s} + \frac{a_2}{s^2}\,, \tag{9.6}$$

where a_1, a_2, a_3 are linear nuisance parameters and B accounts for all possible effects on the clustering amplitude, such as the linear bias, the normalisation of the power spectrum, σ_8, and the redshift space distortions [266]. In addition, we introduce the shift parameter α which takes into account the distortion between distances measured in the data due to the fiducial cosmology chosen to build the estimator. This is defined as [185]:

$$\alpha = \frac{D_{\rm V}}{r_{\rm s}}\frac{r_s^{\rm fid}}{D_V^{\rm fid}}\,, \tag{9.7}$$

where r_s represents the sound-horizon [128, 83] and D_V is the volume-averaged distance defined as [86]:

$$D_{\rm V}(z) = \left[cz(1+z)^2 D_{\rm A}^2(z) H^{-1}(z)\right]^{1/3}\,,$$

where $D_{\rm A}(z)$ and $H(z)$ are the angular diameter distance and the Hubble parameter at redshift z, respectively.

We use a Monte Carlo Markov Chain (MCMC) based on a Metropolis-Hastings algorithm[3] to find the optimal parameter values. We assume a likelihood function of the form $\mathcal{L} \propto \exp(-\chi^2/2)$, where the χ^2 is computed as $\chi^2(\alpha, B) = (\vec{\xi}_{\rm model} - \vec{\xi}_{\rm obs})^T \hat{\Psi}(\vec{\xi}_{\rm model} - \vec{\xi}_{\rm obs})$: In the equation above, $\vec{\xi}_{\rm model}$ is the theoretical correlation function given in Eq. (9.6), $\vec{\xi}_{\rm obs}$ is the observed one, both grouped in a vector at each position, and $\hat{\Psi}$ is the precision matrix given in Eq. (9.3).

9.5 Results and discussions

In this analysis, we have considered different model scenarios with an increasing level of complexity. Our main results are summarised in Tab. 9.1 and shown in Fig. 9.3, together with previous results from literature. First, we use the model given in Eq. (9.6), which has 5 parameters (5p) common to all the targets: (B, α, a_0, a_1, a_2). This is equivalent to assume that all the targets respond in the same way to the gravitational interaction and expansion.

---

[3] https://emcee.readthedocs.io/en/stable/

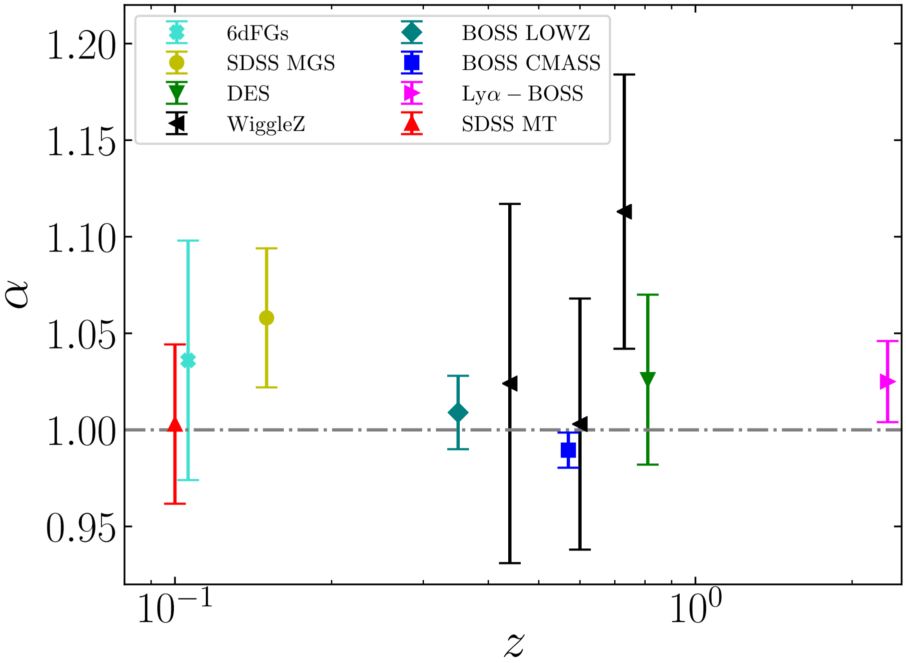


Figure 9.3: Shift parameter α as a function of redshift from different BAO measurements: 6DFGs [38], MGS [210], DES [3], WiggleZ [41], Lowz-BOSS [105], CMASS-BOSS [266] and Lyα-BOSS [79].

The second model we test is a modification of Eq. (9.6), with 13 independent parameters (13p): α, common to all the targets, plus three different sets of (B, a_0, a_1, a_2). The last model used is again a modification of Eq. (9.6), with three different sets of $(B, \alpha, a_0, a_1, a_2)$ i.e., 15 parameters (15p) in total. As a test, we also report the analysis performed using only LRG target. In the 5p scenario, we find a shift parameter of $\alpha = 1.00 \pm 0.04$, while in the 13p model $\alpha = 1.01 \pm 0.04$. In the 15p scenario, we find $\alpha = 1.02 \pm 0.04$ for both Hα and [O II], and $\alpha = 0.97 \pm 0.05$ for LRGs. The latter is consistent with $\alpha = 0.96 \pm 0.07$ found using only LRGs, which is directly comparable to the results at $z \sim 0.15$ by [210] and at $z \sim 0.35$ by [54].

Models:	5p	13p	15p	LRG
α_{H_α}	-	-	$1.02^{0.04}_{0.03}$	-
α_{OII}	-	-	$1.02^{0.04}_{0.04}$	-
α_{LRG}	-	-	$0.97^{0.05}_{0.05}$	$0.96^{0.07}_{0.06}$
α	$1.00^{0.04}_{0.04}$	$1.01^{0.04}_{0.04}$	-	-

Table 9.1: Best-fit constraints from our models.

For our fiducial cosmology, we find a volume-averaged distance of $D_{\rm V}^{\rm fid}(z = 0.1) = 429.90\,{\rm Mpc}$, and $D_{\rm V}^{\rm fid}(z = 0.1)/r_{\rm s}^{\rm fid} = 2.92$. By combining Eq. (9.7) with the constraints obtained on α, we find that the 1σ uncertainty on $D_{\rm V}/r_{\rm s}$ is ~ 0.12. Assuming $r_{\rm s} = 147.41\,{\rm Mpc}$ [6], we find $D_{\rm V} = (435.07 \pm 17.14)\,{\rm Mpc}$.

Our study relies on the jackknife covariance matrices from SDSS data, corrected from bias and noise (see Sec. 9.3), whereas [210] use covariances from synthetic mock catalogues. As shown by [183] and [199], jackknife returns reliable covariance estimates only on large scales, that are the scales of interest in our cosmological analysis. Hence, we do not expect our results to change substantially if covariance matrices from mocks were used.

Another difference between our analysis and [210] is the fact that we do not reconstruct the density field. The main idea of BAO reconstruction [84, 187] is to smooth the linear matter density field and to sharpen the acoustic peak in the correlation function. This method has the advantage of accurately constraining the non-linear parameter $\Sigma_{\rm nl}$. In our analysis we find $\Sigma_{\rm nl} \sim 20\,{\rm Mpc}\,h^{-1}$, while [210] find $\Sigma_{\rm nl} \sim 5\,{\rm Mpc}\,h^{-1}$. A lower value of $\Sigma_{\rm nl}$ provides a better signal and tighter constraints on both B and α. Our results are fairly compared and in agreement within 1σ with the pre-reconstruction value of $\alpha = 1.01 \pm 0.09$ from [210] for LRGs only. The uncertainty we find on α using our multi-tracer analysis is 0.04, identical to the post-reconstruction estimate by [210], and 44% smaller than their pre-reconstruction value. Hence, we expect that by implementing reconstruction on mocks for galaxy multi-tracers, we will be able to significantly reduce our current error. This result highlights the great potentiality of combining different tracers to constrain more accurately the cosmological parameters.

We remind the reader that for all the models used in this work we assumed flat priors, differently from [210], where Gaussian priors are considered. Flat priors are less informative, but they do not rule out any region of the parameter space.

We have performed a cosmological analysis on the LRG, Hα and [O II] ELG multi-tracers currently available at $z \sim 0.1$. For the future, we plan to extend the multi-tracer methodology tested here to the upcoming data sets from the new spectroscopic surveys, such as DESI or Euclid. Our ultimate goal is to include synthetic mock catalogues for galaxy multi-tracers testing the impact of BAO reconstruction on our results. This will enable us to improve the multi-tracer covariance estimates on all scales, and hopefully we will be able to put tight constraints on the non-linear redshift-space distortions.

Acknowledgments

GF is funded through a Dennis Sciama fellowship at the Institute of Cosmology and Gravitation (ICG), at Portsmouth University. DS acknowledges financial support from the Fondecyt Regular project number 1200171. JS acknowledges financial support from CONICYT. The authors are grateful to Melita Carbone, Philipp Sudek and Florian Beutler for insightful discussions and to Daniel Eisenstein for providing useful comments during the preparation of this work.

Chapter 10

Does jackknife scale really matter for accurate large-scale structure covariances?

This chapter presents a research work which corresponds to [93]. This work is a team effort with aid from Dr. Ginevra Favole affiliated by that time to the Institute of Cosmology & Gravitation at the University of Portsmouth in UK, Dr. Benjamin R. Granett affiliated at the Università degli Studi di Milano in Italy, and Dr. Domenico Sapone affiliated to the Cosmology and Theoretical Astrophysics group at the Facultad de Ciencias Fśicas y Matemáticas of the Universidad de Chile in Chile.

Abstract

The jackknife method gives an internal covariance estimate for large-scale structure surveys and allows model-independent errors on cosmological parameters. Using the SDSS-III BOSS CMASS sample, we study how the jackknife size and number of resamplings impact the precision of the covariance estimate on the correlation function multipoles and the error on the inferred baryon acoustic scale. We compare the measurement with log-normal mock galaxy catalogues with the same survey geometry. We build several jackknife configurations that vary in size and number of resamplings. We find that it is useful to apply the tapering scheme to estimate the precision matrix from a limited number of jackknife resamplings. The results from CMASS and mock catalogues show that the error estimate of the baryon acoustic scale does not depend on the jackknife scale. For the shift parameter α, we find an average error of 1.6% and 1.2%, respectively from CMASS and mock jackknife covariances, consistent with pre-reconstruction analyses. However, when relatively few resamplings are used, the jackknife error estimate becomes unreliable. Future large-scale structure surveys will map even greater volumes allowing percent-level estimation of the covariance matrix using a jackknife approach.

Keywords: cosmology: large-scale structure of Universe; cosmological parameters; observations; theory – galaxies: statistics; haloes.

10.1 Introduction

The most popular methods to estimate the uncertainties on the galaxy two-point correlation function (2PCF) internally in a survey are bootstrap [81, 71, 183, 182] and jackknife [205, 168, 252, 183, 182, 199] resampling. Bootstrap resampling is carried out by randomly selecting N_{b} sub-volumes, with replacement, from the original sample. Then the galaxy clustering measurement is performed in each resampling, which is associated a constant weight equal to the number of times that the sub-volume has been selected [183]. Similarly, the jackknife method uses N_{jk} regions in the survey footprint, each with approximately the same galaxy number density. The correlation function is measured on the survey multiple times, each time removing a different jackknife region [183]. The covariance matrix is finally inferred from the 2PCF measurements and the 1σ errors are derived as the square root of the diagonal elements.

Internal methods for error estimation are computationally inexpensive and are derived directly from the measurements. Therefore, the analysis does not depend on an assumed cosmological model, which is an attractive feature when testing alternative models such as dark energy or modified gravity. Indeed, jackknife resampling has been widely used to estimate the uncertainties on the galaxy clustering measurements from large-volume spectroscopic surveys [e.g. 270, 271, 272, 112, 209, 25, 113, 114, 90, 88, 108, 92].

Jackknife resampling has two main disadvantages: (i) it tends to overestimate the errors due to the lack of independence between the N_{JK} copies; and (ii) it is necessary to balance the number and size of resamplings to be drawn in the survey footprint. The last issue is driven by several factors. First of all, in order to have covariances with reduced noise in their off-diagonal terms, we need a large number of jackknife resamplings. This limits the size of the jackknife regions and also the maximum scales that can be probed in the galaxy clustering observables. It is often assumed in the literature that the jackknife cell size S_{JK} should embed the maximum scale measured in the two-point correlation functions. At the same time, to have an invertible (i.e. non-singular) covariance matrix, the number of resamplings should be larger than the number of bins in the measured 2PCF. These conditions are difficult to satisfy in galaxy samples with limited area [e.g. 38, 124]. Due to the finite size of any survey footprint, the more resamplings we draw, the smaller their size and the variation between one copy and the next one [183, 182].

The issues discussed above have discouraged some cosmologists to use jackknife resampling in favour of estimating the galaxy clustering uncertainties from large sets of independent synthetic galaxy catalogues. In the last years, the advent of efficient codes based on fast gravity solvers has considerably reduced the computational time needed for massive mock production, making available many different realisations of accurate, independent mocks for covariance estimates. Among these methods, PTHALOS [225, 162], PINOCCHIO [172, 170], PATCHY [140] and HALOGEN [26] are all based on Lagrangian perturbation theory (LPT). Others, such as QPM [263], FastPM [97] or PPM-GLAM [141], use a quick particle mesh approach. Algorithms such as EZ-MOCKS [55] adopt the effective Zel'dovich approximation, while COLA [247, 142], L-PICOLA [126] or ICE-COLA [134] combine LPT with N-body solvers to speed up the computational time. Finally, high-fidelity mocks can also be obtained from multiple realisations of a log-normal density field [58, 38, 124, 11, 157].

All these fast mocking approaches are extremely convenient compared to running a full

N-body code, but they are generally limited to predicting the dark matter distribution. On top of the dark matter field, it is necessary to model the galaxy distribution by properly accounting for the different baryonic components and the complex physics of galaxy formation and evolution [e.g. 169, 178].

Simulating baryons is a non-trivial task, which requires advanced computational techniques and resources. Semi-analytic models of galaxy formation and evolution [SAMs; 264, 137, 30, 35, 34, 234, 62, 63, 67, 68, 102, 60] and hydrodynamical simulations [e.g. 267, 238, 269, 237, 256, 257, 103, 64, 221, 201] are now able, with different degrees of accuracy, to incorporate the multitude of ingredients and physical processes that contribute to shape the formation and evolution of galaxies within their host dark matter haloes. Some of these processes are: gas accretion [116, 121, 123] and cooling [77, 171, 125], star formation [145], stellar winds [146], stellar evolution Tonini et al. [248], Henriques et al. [120], Gonzalez-Perez et al. [106], AGN feedback Bower et al. [50], Croton et al. [67] or environmental processes Weinmann et al. [259], Font et al. [99], Stevens and Brown [240], Collacchioni et al. [60].

Analogously, one should also account for the effect of massive neutrinos on the growth of cold dark matter perturbations, which are responsible of suppressing the matter power spectrum at intermediate and small scales [19, 265, 191].

All of these assumptions and prescriptions have uncertainties which become significant on non-linear scales and limit the accuracy of the covariance estimate.

Upcoming surveys, such as the Dark Energy Spectroscopic Instrument[1] [DESI; 222], Euclid[2] [148, 219] and the Large Synoptic Survey Telescope[3] [LSST; 132], will bring us to the era of high precision cosmology. In order to prepare to this new phase, it is imperative to improve and compare different methods to construct accurate covariances able to capture the hidden physical process of gravitational collapse. These methods have to carefully optimise the specific binning scheme adopted in order to minimise the noise in the measurements.

For the reasons above, in this work we aim to rehabilitate the use of jackknife resamplings versus mocks for estimating covariances. We explore how varying the size ($S_{\rm JK}$) and number ($N_{\rm JK}$) of jackknife regions impacts the precision in the error estimates of galaxy clustering and on the baryon acoustic oscillation scale through the α shift parameter. In concrete, we measure the monopole and quadrupole two-point correlation functions of the BOSS CMASS DR12 galaxies and we compute their covariances using four different jackknife configurations, coupled with two binning schemes. We compare these results with those obtained from 200 independent log-normal mock light-cones with the same volume of CMASS, and with the covariance obtained by performing jackknife resampling on a single light-cone.

From these covariances, we build the precision matrices needed to estimate the α shift parameter through a Monte Carlo Markov Chain (MCMC) algorithm. We reduce the noise in their off-diagonal terms by applying the tapering correction [138]. We study the impact of a variation in the tapering parameter, $T_{\rm p}$, on the α results. These estimates of α will be directly compared with the galaxy clustering pre-reconstruction results from the BOSS

---

[1] `https://www.desi.lbl.gov`

[2] `https://www.euclid-ec.org`

[3] `https://www.lsst.org`

collaboration [211].

The paper is organised as follows: in Sec. 10.2, we introduce the observational galaxy sample used in our analysis, SDSS-III/BOSS CMASS DR12; in Sec. 10.3, we present the galaxy clustering measurements performed, together with the jackknife methodology and schemes used to estimate their uncertainties. Sec. 10.4 describes the models used to analyse the CMASS observations: the log-normal mock galaxy catalogues and light-cones (§ 10.4.1), and the analytic models used in the Monte Carlo runs (§ 10.4.2). In Sec. 10.5, we explain the Monte Carlo algorithm used to extract the α BAO parameter. Sec. 10.6 presents our main results, which are discussed in Sec. 10.7 together with our conclusions.

Throughout the paper we adopt a Planck et al. [202] cosmology with $\Omega_m = 0.307115$, $\Omega_\Lambda = 0.692885$, $h = 0.6777$, $n = 0.96$ and $\sigma_8 = 0.8228$.

10.2 Observed galaxy sample: BOSS CMASS DR12

The SDSS-III BOSS survey observed about 1.2 million galaxies over an effective area of 9329 deg^2 [17], using the 2.5m Sloan Telescope [111] at the Apache Point Observatory in New Mexico. It used a drift-scanning mosaic CCD camera with five photometric bands, *ugriz* [110, 101], and two spectrographs covering the wavelength range 3600 − 10000Å with a resolving power of 1500 to 2600 [231]. Spectroscopic redshifts were measured using the minimum-χ^2 template-fitting procedure by [13], with templates from [47].

BOSS targeted galaxies into two main samples: LOWZ at $z < 0.43$ and CMASS at $0.43 < z < 0.7$ [12]. For our analysis, we use the data from the BOSS CMASS DR12 sample [15, 207, 17], which is defined through a number of magnitude and colour cuts aimed at obtaining a selection of galaxies with approximately constant stellar mass. These cuts are:

$$\begin{aligned} 17.5 < i_{\mathrm{cmod}} &< 19.9 \\ r_{\mathrm{mod}} - i_{\mathrm{mod}} &< 2 \\ i_{\mathrm{fib2}} &< 21.5 \\ \mathrm{d}_\perp &> 0.55 \\ i_{\mathrm{cmod}} &< 19.86 + 1.6\,(\mathrm{d}_\perp - 0.8), \end{aligned} \tag{10.1}$$

where i_{cmod} represents the SDSS i-band cmodel magnitude, i_{mod} and r_{cmod} are model magnitudes, i_{fib2} is the fibre magnitude within 2" aperture, and $\mathrm{d}_\perp$ is the following colour combination:

$$\mathrm{d}_\perp = r_{\mathrm{mod}} - i_{\mathrm{mod}} - (g_{\mathrm{mod}} - r_{\mathrm{mod}})/8. \tag{10.2}$$

In addition, the CMASS sample satisfies also the star-galaxy separation cuts:

$$\begin{aligned} i_{\mathrm{psf}} - i_{\mathrm{mod}} &> 0.2 + 0.2\,(20 - i_{\mathrm{mod}}) \\ z_{\mathrm{psf}} - z_{\mathrm{mod}} &> 9.125 - 0.46 z_{\mathrm{mod}}, \end{aligned} \tag{10.3}$$

where the subscrit “psf” stands for PSF magnitude.

10.3 Measurements

10.3.1 Two-point correlation functions

We measure the two-point correlation function, $\xi(s,\mu)$, of the galaxy sample described in Sec. 10.2 using the code from Favole et al. [88]. This is based on the Landy-Szalay estimator [147],

$$\xi(s,\mu) = \frac{DD(s,\mu) - 2DR(s,\mu) + RR(s,\mu)}{RR(s,\mu)}, \tag{10.4}$$

where s is the redshift-space distance and μ is the cosine of the angle between s and the line of sight.

In the expression above, DD, DR and RR are respectively the data-data, data-random and random-random pair counts that we can form between the galaxy and random catalogues. The latter is built to have the same angular footprint and radial distribution of the CMASS observations. All the pairs above are properly normalised by the number of galaxies (randoms) and weighted to correct from different systematic effects [see e.g. 216, 209, 90]. In particular, we weight the observed data for potential fibre collisions ($w_{\rm fc}$) and for redshift failures ($w_{\rm zf}$). We also account for possible variation in the galaxy (random) number densities assuming the FKP [96] weight:

$$w_{\rm FKP} = \frac{1}{1+n(z)P_0}, \tag{10.5}$$

where $n(z)$ is the galaxy (random) number density at redshift z and P_0 is a constant quantity that roughly corresponds to the amplitude of the CMASS power spectrum at $k = 0.1\,h{\rm Mpc}^{-1}$. We assume $P_0 = 20000\,h^3{\rm Mpc}^{-3}$ as in Anderson et al. [25].

From Eq. 10.4, we compute the multipoles of the CMASS correlation function as:

$$\xi_l(s) = \frac{2l+1}{2}\int_{-1}^{1} \xi(s,\mu)P_l(\mu)\,{\rm d}\mu, \tag{10.6}$$

where $P_l(\mu)$ are the Legendre polynomials. In this study, we focus only on the first two even multipoles of the 2PCF, i.e. the monopole $\xi_0(s)$ and the quadrupole $\xi_2(s)$. We explore two different binning schemes, both centered on the BAO distance scale, coupled with the jackknife configurations defined in Sec. 10.3.2: (i) 20 linear bins in $24 < s < 184\,h^{-1}$Mpc and 120 linear bins in $0 < \mu < 1$; (ii) 10 linear bins in $24 < s < 184\,h^{-1}$Mpc and 120 linear bins in $0 < \mu < 1$.

10.3.2 Jackknife configurations and covariances

We implement jackknife resampling in the BOSS CMASS DR12 galaxy sample adopting four different configurations summarised in Table 10.1. We divide the survey footprint into 200, 100, 50 and 20 RA×DEC cells approximately containing the same number of galaxies (randoms). The CMASS covariance matrix for $N_{\rm JK}$ jackknife resamplings is [e.g. 209, 90]:

$$C_{\rm ij} = \frac{N_{\rm JK}-1}{N_{\rm JK}} \sum_{\rm a=1}^{N_{\rm JK}} (\xi^a_{\rm i} - \bar{\xi}_{\rm i})(\xi^a_{\rm j} - \bar{\xi}_{\rm j}), \tag{10.7}$$

$N_{\rm JK}$	$A_{\rm JK}$ [deg^2]	$S_{\rm JK}$ [h^{-1}Mpc]
200	46.6	110.7
100	93.3	156.6
50	186.6	221.4
20	932.9	495.1

Table 10.1: Jackknife configurations adopted in our analysis. For each of the four cases implemented, we indicate the number of jackknife resamplings ($N_{\rm JK}$), the area ($A_{\rm JK}$) and comoving size ($S_{\rm JK}$) of the individual cell computed in Planck et al. [202] cosmology at the mean redshift of CMASS, $z = 0.56$.

where $\bar{\xi}_i$ is the mean jackknife correlation function in the $i^{\rm th}$ bin,

$$\bar{\xi}_{\rm i} = \sum_{\rm a=1}^{N_{\rm JK}} \xi_{\rm i}^{a}/N_{\rm JK}. \tag{10.8}$$

The overall factor in Eq. 10.7 corrects from the lack of independence between the $N_{\rm JK}$ jackknife copies, which is the main limitation of the jackknife method. In fact, from one configuration to the next, $N_{\rm JK} - 2$ cells are the same [182].

10.4 Models

10.4.1 Log-normal mock galaxy catalogues and light-cones

We generated 200 log-normal mock galaxy catalogues for the BOSS CMASS sample at mean redshift $z \sim 0.56$. The target power spectrum was computed with the code *CLASS*[4] [46]. We applied a linear bias and the Halofit [246] prescription to model the non-linear galaxy power spectrum: $P(k) = b^2 P_m(k)$ with the value $b = 2.1$.

We present the *Synmock* code used to produce the log-normal catalogues in a public repository[5]. The implementation follows the standard approach for generating log-normal simulations [see also 38, 194]. For each realization, first we generated a Gaussian density field $\delta_G(\vec{x})$ on a cubic grid with dimension $L = 4096h^{-1}$Mpc and step size $h = 8h^{-1}$Mpc and transformed it to derive the target log-normal field:

$$\delta(\vec{x}) = \exp\big(\delta_G(\vec{x}) - \sigma^2/2\big) - 1\,, \tag{10.9}$$

where σ^2 is the variance of the Gaussian field. In order to match the target power spectrum we made a Fourier transform of the power spectrum to compute the correlation function and using the relationship [58]

$$\xi_G(|\vec{x} - \vec{x}'|) = \log\left(1 + \xi(|\vec{x} - \vec{x}'|)\right)\,. \tag{10.10}$$

The log-normal density field was used to build a discrete galaxy field by Poisson sampling the number density $n(\vec{x}) = \bar{n}(1 + \delta(\vec{x}))$. We applied a uniform random offset to move the mock galaxies away from the grid points.

---

[4] `https://github.com/lesgourg/class_public`

[5] `https://github.com/bengranett/synmock`

The velocity field was computed on the same $\delta(x)$ grid using the linear continuity equation in Fourier space:

$$\vec{v}(\vec{k}) = i\frac{faH}{b}\delta(\vec{k})\frac{\vec{k}}{k^2}\,, \tag{10.11}$$

where f is the logarithmic growth rate. After a final Fourier transform, the velocity of each galaxy was assigned using the value at the nearest grid point.

We built 200 BOSS CMASS light-cones (LCs) with $0.43 < z < 0.7$ by cutting the BOSS survey geometry in the log-normal simulation box above. The Cartesian galaxy coordinates were transformed to the spherical coordinates right ascension, declination and radial distance with the origin at the center of the simulation box. In order to transform to the redshift-space coordinates, the line-of-sight peculiar velocity component was computed and applied to the radial comoving distance: $r_s = r + \vec{r}\cdot\vec{v}/\left(aH|\vec{r}|\right)$. We constructed a coarse angular mask using the Healpix [107] scheme at resolution $n_{\rm side} = 64$ and discarded galaxies outside the mask. The catalog was further downsampled along the radial direction to match the target redshift distribution. We generated an unclustered random catalogue with 10 times the number density of the CMASS data that precisely corresponds to the mock construction using the same angular mask and radial selection function.

After computing the correlation functions of the $N_{\rm LC} = 200$ log-normal light-cones, we derive their covariance matrix as:

$$C_{\rm ij} = \frac{1}{N_{\rm LC}-1}\sum_{\rm a=1}^{N_{\rm LC}}(\xi_{\rm i}^{a} - \bar{\xi}_{\rm i})(\xi_{\rm j}^{a} - \bar{\xi}_{\rm j}), \tag{10.12}$$

where $\bar{\xi}_{\rm i}$ is their mean 2PCF in the $i^{\rm th}$ bin. The pre-factor properly accounts for the fact that the mock realisations are independent.

10.4.2 Analytic models

Besides the log-normal mocks, we also model the multipoles of the BOSS CMASS two-point correlation function using an analytic approach, which is required to run the Monte Carlo analysis (see Sec. 10.5). The 2PCF can be obtained from the Fourier transform of the matter power spectrum, $P(k)$, for which we assume the template from [185]:

$$P(k) = \left[P_{\rm lin}(k) - P_{\rm dw}(k)\right]{\rm e}^{-k^2\Sigma_{\rm nl}^2/2} + P_{\rm dw}(k)\,. \tag{10.13}$$

In the equation above, $P_{\rm lin}(k)$ is the linear matter power spectrum computed using the Boltzmann code CLASS [151] assuming the Planck 2015 [6] fiducial cosmology. The $P_{dw}(k)$ term is the de-wiggled power spectrum [83], while the Σ_{nl} parameter encodes the smoothing of the BAO peak due to non-linear effects [65]. The multipoles of the analytic 2PCF are defined as:

$$\xi_l(s) = \frac{i^l}{2\pi^2}\int_0^\infty P_l(k)j_l(ks)k^2{\rm d}k\,, \tag{10.14}$$

from which we recover the monopole ($l = 0$) and the quadrupole ($l = 2$). In Eq. 10.14, $j_l(x)$ represents the spherical Bessel function of first kind and order l, while $P_l(k)$ are the multipoles of the power spectrum defined as:

$$P_l(k) = \frac{2l+1}{2}\int_{-1}^{1}\left(1 + f\mu^2\right)^2 P(k)L_l(\mu){\rm d}\mu\,, \tag{10.15}$$

where $L_l(x)$ is the Legendre polynomial of order l and $P(k)$ is the template given in Eq. 10.13. By replacing Eq. 10.15 in Eq. 10.14, the analytic expressions for monopole ($l = 0$) and quadrupole ($l = 2$) are respectively [266]:

$$\xi^{(0)}_{\rm model}(s) = B_0\xi_0(\alpha s) + a_0^{(0)} + \frac{a_1^{(0)}}{s} + \frac{a_2^{(0)}}{s^2}\,, \tag{10.16}$$

$$\xi^{(2)}_{\rm model}(s) = B_2\xi_2(\alpha s) + a_0^{(2)} + \frac{a_1^{(2)}}{s} + \frac{a_2^{(2)}}{s^2}\,, \tag{10.17}$$

where α is the shift parameter, while $(a_1^{(i)}, a_2^{(i)}, a_3^{(i)})$ are linear nuisance parameters.

The shift parameter α in Eqs. 10.16 and 10.17 is usually defined as [185]:

$$\alpha = \frac{D_V}{r_s}\frac{r_s^{\rm fid}}{D_V^{\rm fid}}\,, \tag{10.18}$$

where r_s represents the sound horizon [128], and D_V the volume-averaged distance given by [86]:

$$D_{\rm V}(z) = \left[cz(1+z)^2 D_{\rm A}^2(z) H^{-1}(z)\right]^{1/3}, \tag{10.19}$$

with $D_{\rm A}(z)$ being the angular diameter distance, and $H(z)$ the Hubble parameter at redshift z. The α shift parameter accounts for the observed distortion between distances due to the chosen fiducial cosmology, while the nuisance parameters $(a_1^{(i)}, a_2^{(i)}, a_3^{(i)})$ and B_1, B_2 incorporate those effects that are responsible of modulating the clustering amplitude, such as redshift-space distortions [266], linear bias, and the power spectrum normalisation, σ_8.

10.5 Shift parameter estimation

Following the methodology presented in Favole et al. [92], we analyse the BOSS CMASS covariances, both computed from jackknife and 200 light-cones, using a Monte Carlo Markov Chain based on a Metropolis-Hastings algorithm[6]. Our MCMC code is publicly available on GitHub[7].

In order to find the optimal parameter values, we assume a likelihood function $\mathcal{L} \propto \exp(-\chi^2/2)$, with

$$\chi^2 = \left(\vec{\xi}_{\rm model} - \vec{\xi}_{\rm obs}\right)^T \hat{\Psi} \left(\vec{\xi}_{\rm model} - \vec{\xi}_{\rm obs}\right), \tag{10.20}$$

where

$$\vec{\xi}_{\rm model} \equiv \left(\vec{\xi}^{(0)}_{\rm model}, \vec{\xi}^{(2)}_{\rm model}\right) \tag{10.21}$$

represents the theoretical correlation function whose components are given in Eqs. 10.16-10.17, while $\vec{\xi}_{\rm obs}$ corresponds to the observed monopole and quadrupole 2PCFs, both grouped in a vector depending on the comoving distance. The Ψ term above is the precision matrix defined as:

$$\hat{\Psi} = \left(1 - \frac{n_{\rm b}+1}{N_{\rm res}-1}\right)\left(\hat{C} \circ T\right)^{-1} \circ T\,, \tag{10.22}$$

---

[6] `https://emcee.readthedocs.io/en/stable/`

[7] `https://github.com/javiersilvalafaurie/BTCosmo`

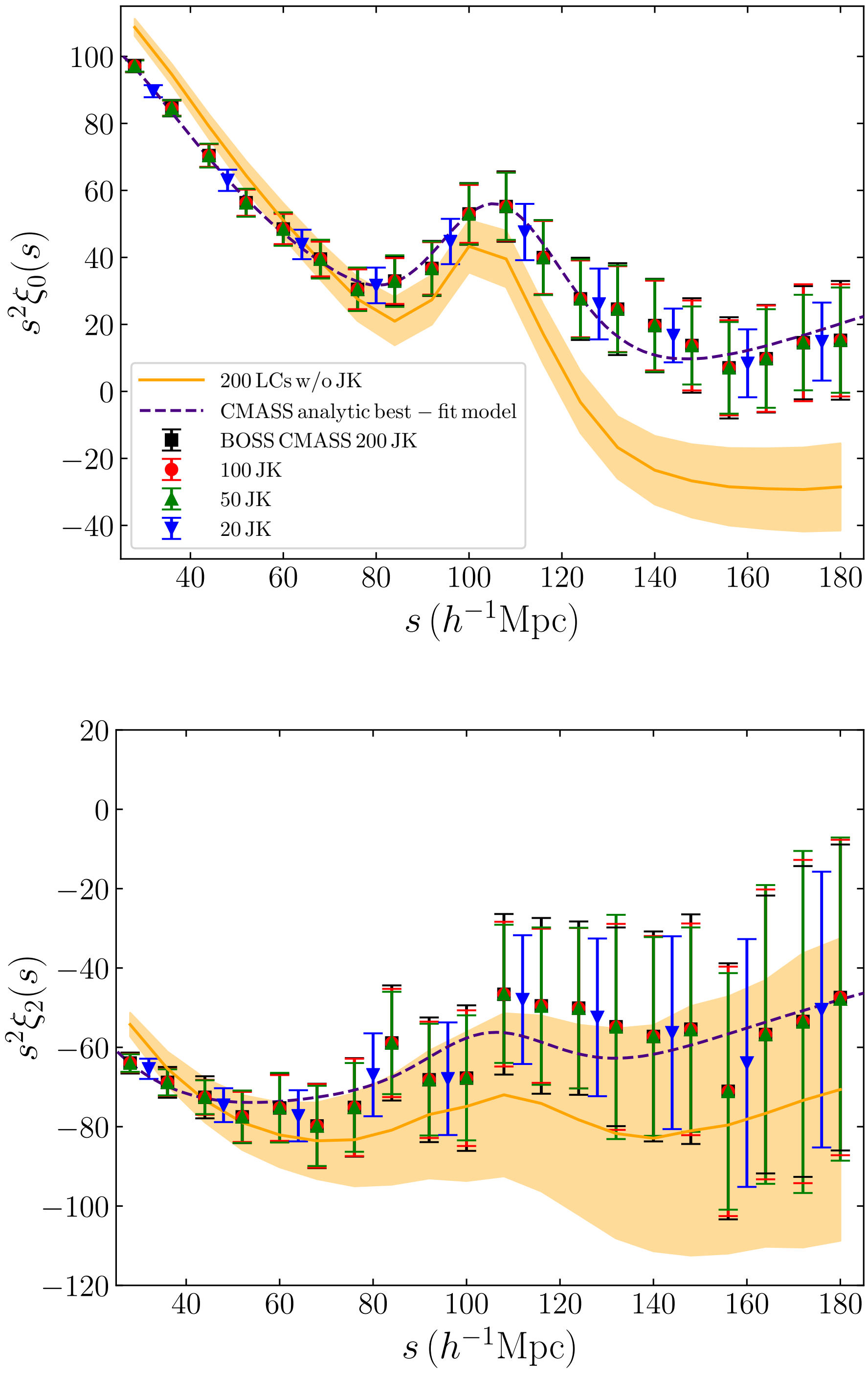


Figure 10.1: Monopole (top) and quadrupole (bottom) auto-correlation functions of the BOSS CMASS galaxies (markers) computed using two different binning schemes (20 and 10 linear bins in s) coupled with the jackknife configurations given in Table 10.1 for the error estimation (200, 100, 50, 20 resamplings). We overplot the mean $\pm\sigma$ values from the 200 log-normal light-cones (Sec. 10.4.1) as orange lines with the $1\,\sigma$ uncertainty as shaded area. The analytic best-fit models to the CMASS measurements that we use to estimate the α shift parameter (see Sec. 10.4.2) are shown as dashed purple curves.

where $\hat{C}$ is the total assembled covariance matrix:

$$\hat{C} = \begin{pmatrix} \hat{C}_{\xi_0\xi_0} & \hat{C}_{\xi_0\xi_2} \\ \\ \hat{C}^T_{\xi_0\xi_2} & \hat{C}_{\xi_2\xi_2} \end{pmatrix} . \tag{10.23}$$

The first term in parenthesis in Eq. 10.22 is the Hartlap factor [119], which corrects from the bias introduced in the covariance matrix by the limited number of jackknife resamplings and 2PCF bins. In Tab. 10.2, we report the values of the Hartlap factor as a function of the number of jackknife resamplings and bins used in our analysis.

The quantity T in Eq. 10.22 is the tapering correction [138] that minimises the noise in the off-diagonal terms of the covariance matrix; for further details see also [92]. In this work, we assume a tapering parameter $T_{\rm p} = 500\,{\rm h}^{-1}{\rm Mpc}$ to ensure that the entire covariance matrix is positive semi-definite and the noise in the off-diagonal terms is minimised. In Sec. 10.6, we test how a variation in the tapering parameter affects the results for α and its uncertainty. Further details on the dependence of α on $T_{\rm p}$ are addressed also in Paz and Sánchez [192].

10.6 Results

In Fig. 10.1, we present the BOSS CMASS monopole and quadrupole two-point auto- correlation functions compared to the mean predictions from the 200 log-normal light-cones. The CMASS error bars are inferred from the jackknife covariances based on the four configurations shown in Tab. 10.1 coupled with two different binning schemes (see Sec. 10.3.1). For the LCs, we show the dispersion obtained from the 200 realisations without jackknife resampling. The LCs reproduce the BAO peak well, but differ from the CMASS measurements in the broadband shape. The LC monopole prediction tends to overestimate the observed clustering amplitude at $s \lesssim 60\,h^{-1}{\rm Mpc}$ and to underestimate it beyond BAO scales. The systematic difference in shape will be accounted for by the nusiance parameters in the model and so will not influence the analysis of the α shift parameter. We also overplot the analytic 2PCF model used in our MCMC algorithm to estimate the α BAO parameter (Sec. 10.5). The best-fit analytic model is in good agreement with the CMASS multipole measurements on all scales.

In Fig. 10.2 we display the normalised covariance of the monopole and quadrupole from the 200 log-normal light-cones, in 20 s bins, built in Sec. 10.4.1, without jackknife resamplings. The matrix is normalised as $C^{\rm norm}_{\rm ij} = C_{\rm ij}/\sqrt{C_{\rm ii}\,C_{\rm jj}}$, with $C_{\rm ij}$ given in Eq. 10.12.

$n_{\rm b}$	$N_{\rm JK}$	Hartlap factor
10	20	0.42105
20	50	0.57143
20	100	0.78788
20	200	0.89447

Table 10.2: Values of the Hartlap factor [119] as a function of the number of bins $n_{\rm b}$ and jackknife resamplings $N_{\rm JK}$ used in our analysis.

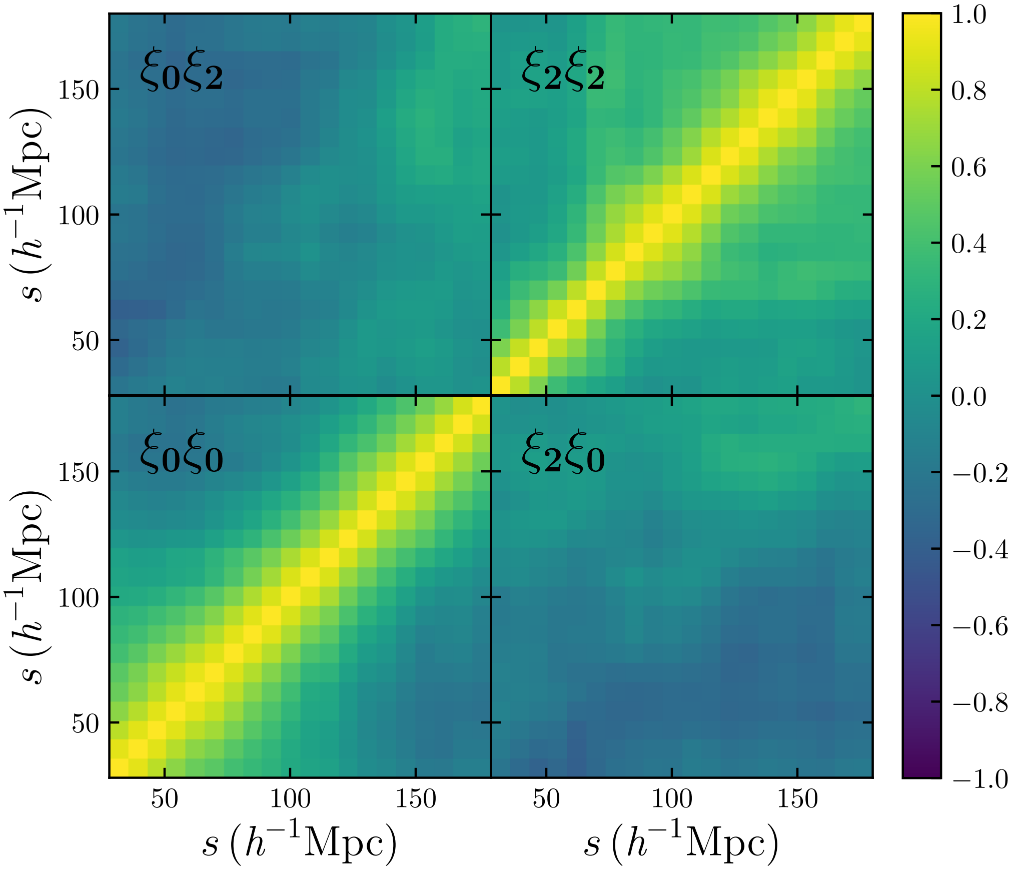


Figure 10.2: Normalised monopole and quadrupole auto- and cross-covariances obtained from the 200 log-normal light-cones without jackknife resampling. The normalisation is computed as $C_{\mathrm{ij}}^{\mathrm{norm}} = C_{\mathrm{ij}}/\sqrt{C_{\mathrm{ii}}\,C_{\mathrm{jj}}}$, where C_{ij} is given in Eq. 10.12. The mean value and $1\,\sigma$ dispersion of these mocks are shown in Fig. 10.1 as a solid line with the corresponding shaded region.

Fig. 10.3 compares the normalised covariances obtained by performing jackknife resampling on a light-cone (upper triangles) versus BOSS CMASS data (lower triangles). From top to bottom, we present the 20, 50, 100 and 200 jackknife configurations listed in Tab. 10.1, respectively coupled with 10, 20, 20, 20 linear bins in s (see Sec. 10.3). The normalisation is calculated as described in Fig. 10.2, with C_{ij} given by Eq. 10.7 for CMASS data and by Eq. 10.12 for the light-cone covariance estimate without jackknife resampling. It is evident that the noise in the covariance estimate is reduced as the number of resamplings is increased. The covariances from 200 jackknife resamplings with 20 bins on a light-cone and on CMASS observations are consistent with the result from the 200 LCs without jackknife shown in Fig. 10.2. These covariances lead to consistent error bars on the galaxy clustering multipoles, as shown in Fig. 10.1.

The top panel of Fig. 10.4 shows, as a function of the scale, the ratio of the uncertainties obtained from CMASS jackknife covariances and from 200 light-cones without jackknife resampling. The results for the monopole and the quadrupole are shown as solid and dashed lines, respectively. We remind the reader that the 200, 100 and 50 JK configurations are

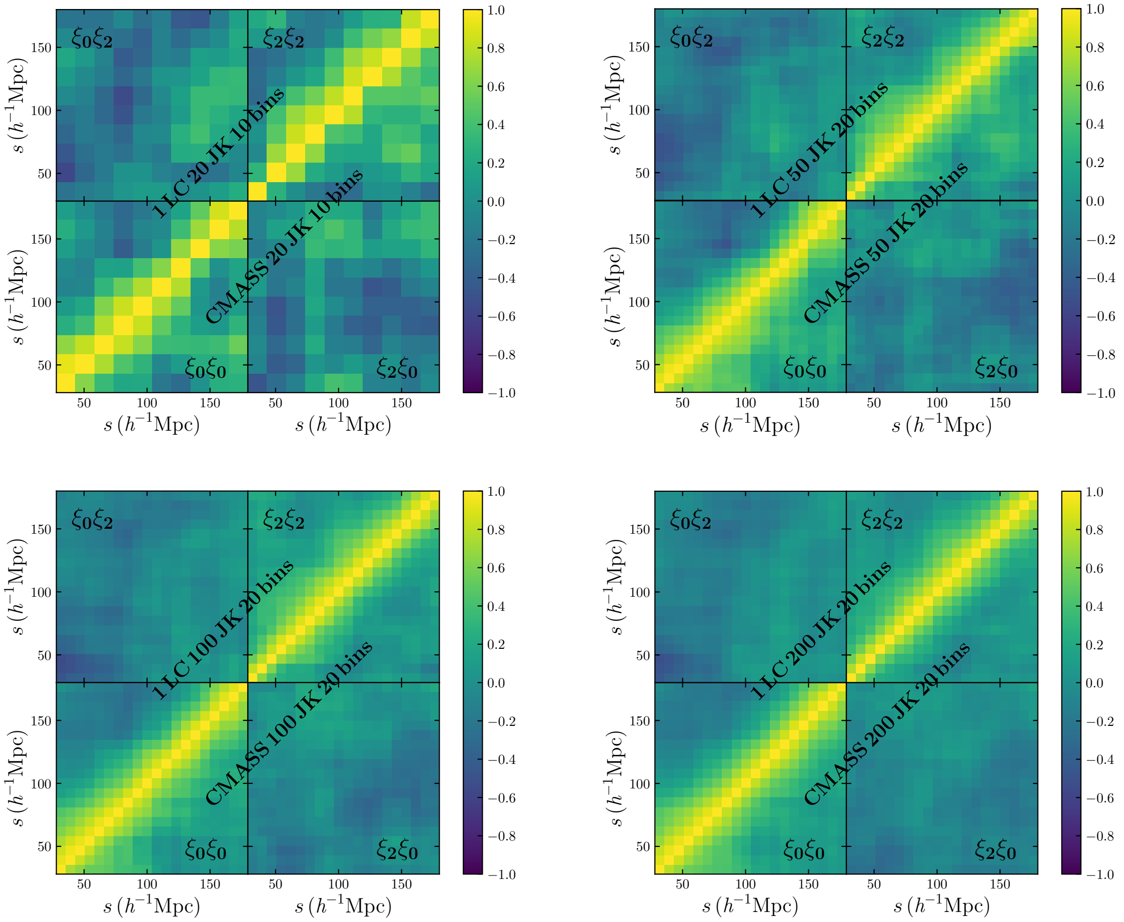


Figure 10.3: Normalised covariances obtained from jackknife resampling performed on a light-cone (upper triangles) and on BOSS CMASS data (lower triangles). We display the 20, 50, 100 and 200 jackknife configurations, respectively, coupled with two binning schemes.

measured in 20 s bins, while the 20 JK case in 10 bins. The combined action of the jackknife size, number and binning is what determines the level of noise in the covariances. The 20 JK bins scheme leads to the largest fluctuations in the σ estimate due to having relatively few jackknife resamplings available. However, doing only 10 bins instead of 20 helps to partially mitigate these fluctuations.

On small scales, the errors from CMASS resamplings shown in the top panel of Fig. 10.4 are underestimated with respect to the LCs by up to $\sim 40\%$ for both monopole and quadrupole. Around $130\,h^{-1}$Mpc, the monopole errors from CMASS covariances are $\sim 10-50\%$ larger than those from 200 light-cones, and the discrepancy increases with the number of resamplings. Compared to the monopole, the quadrupole shows smaller fluctuations in the 1σ ratio shown in the top panel of Fig. 10.4.

The amplitude of the quadrupole 20 JK result is $\sim 10-20\%$ lower than the others on scales below $150\,h^{-1}$Mpc, while the monopole is lower than the rest only beyond $100\,h^{-1}$Mpc. As expected, the 20 and 50 jackknife schemes return the largest fluctuations. Although larger jackknife regions with greater independence may give a more accurate covariance estimate,

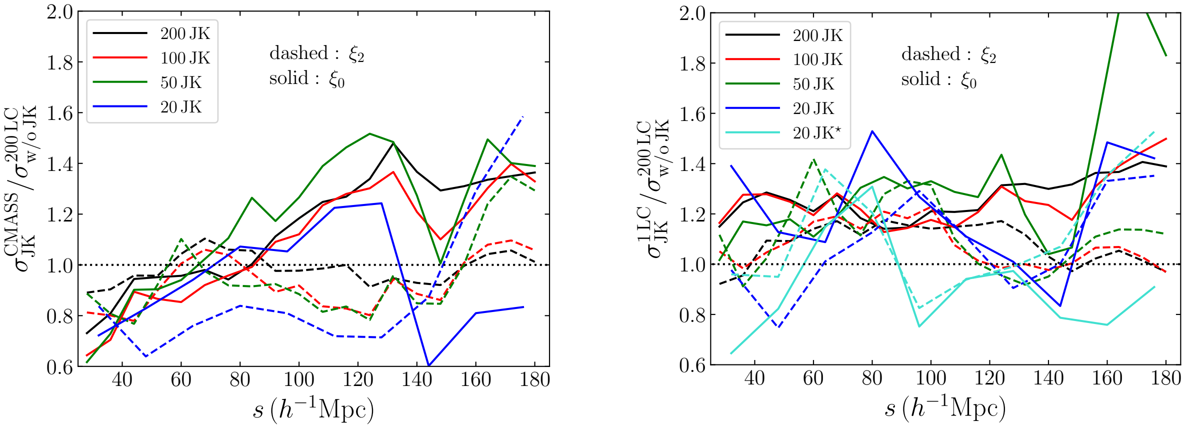


Figure 10.4: *Left:* Ratios of the 1σ uncertainties obtained from the CMASS jackknife covariances and the 200 LCs without JK. The solid (dashed) lines correspond to the monopole (quadrupole) measurements. *Right:* Ratios of the $1\,\sigma$ errors obtained by applying jackknife to one of the light-cones and those from 200 LCs without jackknife. For the 20 JK LC scheme we show two different light-cone realisations (blue and turquoise lines). We remind the reader that the 200, 100 and 50 JK configurations are coupled with 20 s bins, while the 20 JK case with 10 bins. The horizontal dotted lines are shown to help the comparison.

the uncertainty on the covariance is large due to having few resamplings available. In the 20 JK scheme coupled with 10 s bins, the large fluctuations due to a limited number of resamplings are partially mitigated by the smaller number of bins compared to the other cases.

In the bottom panel of Fig. 10.4, we display the ratio of the uncertainties obtained by performing jackknife resampling on one of the light-cones and those from 200 LCs without jackknife. For the 20 JK case, we show the results from two different light-cone realisations (blue and turquoise lines) to highlight the significant uncertainty in the covariance estimate, which we expect to be high in this case. Compared to the ratio of CMASS jackknife to 200 LCs shown in the upper panel, overall here we find smaller fluctuations and a flatter trend. Again, the 20 and 50 JK configurations are the ones exhibiting the largest fluctuations due to the limited number of resamplings.

Out to BAO scales, the errors from jackknife resampling on the LC are overestimated by up to $\sim 30\%$ for both the monopole and quadrupole. Beyond $110\,h^{-1}$Mpc, the error estimates fluctuate significantly, but tend to indicate overestimation by $\sim 50\%$. Those from 50 and 20 JK also maintain a growing trend, but with even larger fluctuations due to the smaller number of resamplings. Beyond BAO scales, the quadrupole errors tend to decrease such that the JK estimates agree with the covariance computed from 200 independent realizations.

Overall, we find strong consistency between the uncertainties based on covariances computed either performing jackknife re-sampling on CMASS data, or on a light-cone with the same volume of CMASS, or from 200 LC realisations without jackknife. In general, with respect to the 200 LCs without jackknife, which represent the "ideal" case, the monopole errors from CMASS JK are underestimated on small scales and overestimated beyond $\sim 80\,h^{-1}$Mpc, while those from jackknife on a light-cone are overestimated on all scales. For the quadrupole, the CMASS JK errors are overall underestimated compared to 200 LCs, while those from jackknife resampling on a LC are overestimated. All these discrepancies remain within 30% in most cases.

We next consider how the covariance of the correlation function propagates to the error on the α shift parameter. Fig. 10.5 compares the values of the α shift parameter and corresponding uncertainties inferred from (i) the BOSS CMASS jackkife covariances, (ii) the log-normal light-cone jackknife covariances, and (iii) the covariance from 200 LCs without jackknife. The specific values are reported in Tab. 10.3. All these results assume a tapering parameter $T_{\rm p} = 500$ which we found to be optimal. For the 20 JK/10 bins scheme applied to the light-cone we show two different LC realisations to highlight the fluctuations that the α uncertainty can suffer due to the small number of resamplings. Overall, we find good agreement between the results based on covariances from jackknife, both applied to CMASS observations and LCs. For the LC results, we use the same realisation for all the jackknife configurations except the second 20 JK case, indicated with a star symbol, in which we test a different realisation. The 200 LC result without jackknife is consistent with the JK outcomes, despite the difference in the pre-factor of their covariances (Eqs. 10.7,10.12) and the lower level of noise in their off-diagonal terms (see Fig. 10.2).

The uncertainties on α are all in agreement with each other, independently from the number/size of jackknife resamplings adopted. The average errors obtained from 200, 100 and 50 JK resamplings (i.e. the most robust ones) performed on CMASS data and a light-cone are $\sim 1.6\%$ and $\sim 1.2\%$, respectively. That from the 200 LC covariances without jackknife

	α	
	BOSS CMASS	**log-normal LC**
200 JK, 20 bins:	$0.974^{+0.022}_{-0.020}$	$0.974^{+0.016}_{-0.014}$
100 JK, 20 bins:	$0.978^{+0.017}_{-0.017}$	$0.973^{+0.011}_{-0.011}$
50 JK, 20 bins:	$0.973^{+0.011}_{-0.011}$	$0.966^{+0.010}_{-0.010}$
20 JK, 10 bins:	$0.980^{+0.019}_{-0.018}$	$0.996^{+0.021}_{-0.016}$
20 JK, 10 bins*:		$0.985^{+0.013}_{-0.013}$
200 LCs w/o JK, 20 bins:		$0.992^{+0.010}_{-0.009}$

Table 10.3: Estimates of the α shift parameter and its uncertainty obtained from the four jackknife configurations coupled with two binning schemes applied to both CMASS data and a log-normal lightcone. The last row shows the result obtained from the covariances of the 200 LCs without performing jackknife resampling. All these results here assume an optimal tapering parameter of $T_{\rm p} = 500$. These results are shown in Fig. 10.5.

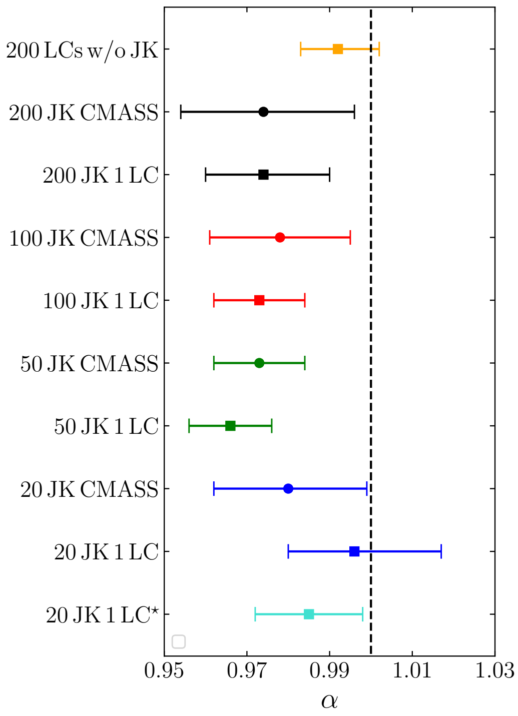


Figure 10.5: Summary of the α shift parameters obtained from the covariances calculated using the jackknife configurations and binning schemes reported in Table 10.3. The points are color-coded as in Fig. 10.1, where each colour corresponds to a different jackknife/binning scheme. The results from CMASS are represented by dots, those from LCs by squares. The vertical line shows the value $\alpha = 1$ to help the comparison. For the 20 JK case applied to a LC, we show two different LC realisations, one of them indicated with a star symbol (turquoise), to highlight how the 1σ error can fluctuate due to the small number of resamplings. All these results are calculated assuming a tapering parameter $T_{\rm p} = 500$.

is $\sim 0.95\%$. This latter case is "ideal" since the 200 LCs are all independent (but we do not expect the log-normal catalogs to capture the full covariance of the CMASS galaxy sample). The 50 and 20 JK schemes are the ones returning the largest fluctuations in the covariances, which can result in errors on α as large as $\sim 2.1\%$. In order to precisely estimate the effect of the fluctuations in the 20 JK case, one should repeat the resamplings on many different LC realisations, but this goes beyond the scope of our analysis.

Despite the large fluctuations observed in the 20 JK configurations, the constraints on α in some cases show that the error bars tend to reduce when the jackknife size increases. However, the trend depends on the tapering scheme and further trials on mocks are required to determine if this trend is real or not.

Fig. 10.6 shows the dependence of α and its uncertainty on the tapering parameter, $T_{\rm p}$. In the top panel, we show the results from the 200 light-cones and from jackknife applied to one of the LCs; in the bottom panel we show the CMASS jackknife outcomes. In both cases, we have run the MCMC chains assuming $T_{\rm p} = [50,\ 100,\ 300,\ 500,\ 700]$. In the plots we offset the tapering values by a multiplicative factor to avoid crowding (see caption of Fig. 10.6).

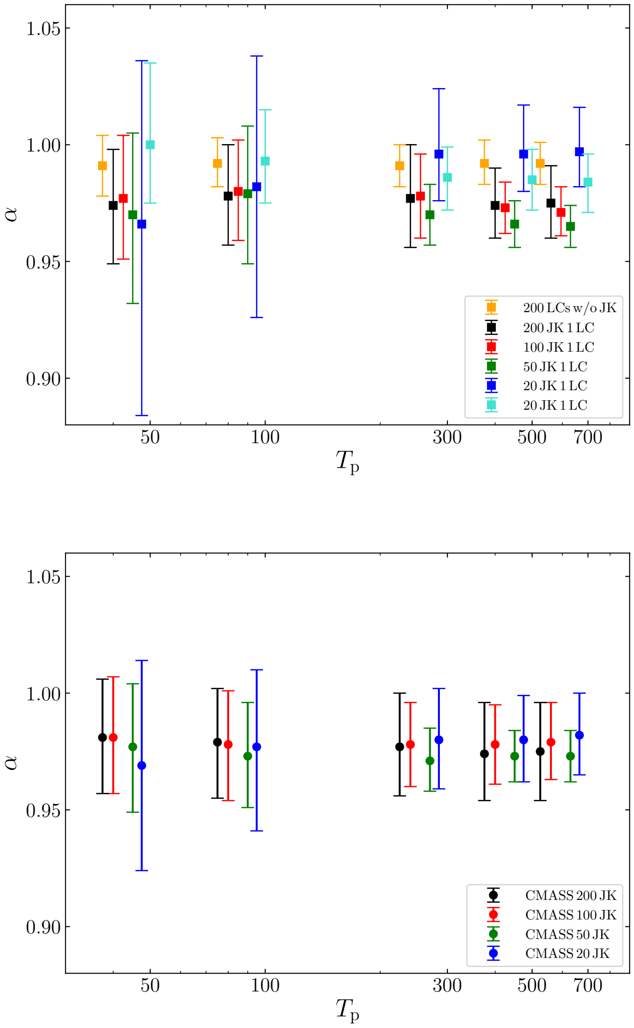


Figure 10.6: Shift parameter α and its uncertainty as a function of the tapering parameter T_p. *Top:* results from covariances computed from 200 light-cones without jackknife and from jackknife performed on a LC. We have offset the T_p values on the x-axis by multiplying them, from left to right, by $[0.75, 0.80, 0.85, 0.90, 0.95, 1.0]$. *Bottom:* results from CMASS jackknife resampling. The T_p values have been offset by multiplying them, from left to right, by $[0.80, 0.85, 0.90, 0.95, 1.0]$.

The optimal value, which provides errors on α of $\sim 1-2\%$, turns out to be $T_\mathrm{p} = 500$. In this way, our BOSS CMASS α estimates are comparable with previous results in the literature [pre-reconstruction, e.g. 211, 193].

The 20 JK configuration shows the largest fluctuations due to the limited number of resamplings. We have further tested our MCMC code without including any tapering correction and leaving only the Hartlap factor. In this case we find that the covariances from 20 JK resamplings are no longer semi-positive definite, meaning that they are not invertible, hence not useful for assembling the precision matrix needed to estimate α. Such a result confirms that jackknife configurations with few cells tend to provide non robust covariance estimates.

10.7 Discussion and summary

We have studied the impact of choosing different sizes and numbers of jackknife resamplings on the accuracy of the covariance estimates and the α shift parameter. To this purpose, we have measured the first two even multipoles of the BOSS CMASS DR12 galaxy sample at $0.43 < z < 0.7$ and we have modelled the results both using a set of 200 log-normal light-cones (Sec. 10.4.1) and an analytic approach (Sec. 10.4.2). We have computed their covariances using 200, 100, 50 and 20 jackknife resamplings coupled with two binning schemes: 20 or 10 linear bins in $24 < s < 184\,h^{-1}$Mpc, with 120 linear bins in $0 < \mu < 1$ (see Sec. 10.3). We have compared the results with the covariances obtained from the 200 log-normal light-cones without jackknife. We have then applied the same jackknife configurations above on one of the light-cones to derive LC JK covariances directly comparable with the CMASS ones.

From these different covariance matrices we have derived corresponding precision matrices (Sec. 10.5), which we have used as inputs for our Monte Carlo Markov Chain to estimate the baryon acoustic scale through the α shift parameter and its uncertainty. Our main findings are summarised in what follows:

- We find good consistency between the covariances obtained from CMASS and LC jackknife resamplings, and from 200 LCs without jackknife resampling. This leads to consistent error bars in both the galaxy clustering measurements and the α shift parameter.
- We find no evidence for a bias in the inferred value of α or its error when the jackknife cell size is smaller than the maximum 2PCF scale measured. However, with few resamplings available the error estimate becomes unreliable.
- We have demonstrated that it is useful to apply the Hartlap factor and the tapering scheme to estimate the precision matrix with jackknife resampling. The α shift parameter estimated either from CMASS or LC jackknife covariances, or from 200 light-cones without jackknife, are all consistent between each other and with previous BOSS CMASS DR12 results from galaxy clustering pre-reconstruction analysis [211]. We find uncertainties on α of 1-2%, depending on the jackknife size/ 2PCF binning scheme adopted. This confirms that the jackknife methodology applied to both observations and mocks produces a comparable level of noise in the covariance estimates. This noise is then reduced in the precision matrix by applying the tapering correction (see Sec. 10.5).
- We have tested different values for the tapering parameter, in the range $50 \leq T_{\rm p} \leq 700$, to maximise the accuracy in the α shift parameter estimation. We find that the optimal value is $T_{\rm p} = 500$. By lowering it, the noise in the precision matrix estimate is suppressed but the error on α grows.

To summarise, performing jackknife resamplings either on BOSS CMASS DR12 data or on a log-normal light-cone with the same CMASS volume provides covariances that are consistent with those obtained from a set of 200 independent log-normal LCs and with previous results in the literature [211]. These covariances lead to α estimates with 1-2% uncertainties, depending on the jackknife size/2PCF binning scheme assumed.

The largest differences between covariance estimates from jackknife resampling and 200 log-normal light-cones without JK are visible in the off-diagonal terms. Here the jackknife results exhibit a higher level of noise. This difference is key for determining the accuracy of the α shift parameter. The action of the tapering correction (Sec. 10.5) is to considerably reduce this noise returning comparable uncertainties on α from all of the different covariance estimates tested.

Although previous works limit the jackknife scale to larger than the measured 2PCF scale [e.g. 38, 124], we find that this is not essential. In fact, when using jackknife to estimate covariances, one should prioritise building a large number of resamplings rather than choosing a jackknife size larger than the maximum galaxy clustering scale measured. In fact, especially when studying BAO scales, by requiring $S_{\rm JK} \geq \max(s)$, we are able to build only few wide jackknife regions, which leads to large uncertainties in the error estimates. In our results we do see a trend that the α error bars tend to reduce as the jackknife size increases, but it is not seen in all tapering configurations and we do not have sufficient statistics to confirm whether it is real or not.

The new generation of cosmological surveys, such as DESI, Euclid or LSST, will span larger volumes compared to SDSS-III/BOSS. The precision in their covariance estimates based on jackknife will be determined by the number of resamplings. We find that it is not essential to use jackknife sizes larger than the BAO scale, and so it will be possible to achieve $N > 10^3$ resamplings to reach percent level precision on the error of cosmological parameters using the jackknife approach. In a followup work, we will address the feasibility of inferring accurate covariance estimates for a survey such as Euclid using a large number of jackknife resamplings.

Acknowledgements

GF is funded through a Dennis Sciama fellowship at the Institute of Cosmology and Gravitation (ICG), at Portsmouth University.

DS acknowledges financial support from the Fondecyt Regular project number 1200171.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is `http://www.sdss3.org/`.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

Conclusions

In the first chapters of this thesis, we made an extensive review of the principal equations needed at the background and first-order level in the ΛCDM scenario. The idea was to thoroughly understand the final result from eq. (4.50) without introducing a lot of theory beyond general relativity or anything unknown for any physicist at the undergraduate level. That result, together with those obtained in chapter 4, allows us to compare the framework constructed with direct observation of galaxies.

We also review the technical part to recover the two-point signal in the New York University Value-Added Galaxy Catalog and some ways to obtain its uncertainty. Although the main objective of this thesis was not to recover the covariance matrix, we regret not delving further into it because we underestimated it, which is propagated into the inference of cosmological information as we saw. Nonetheless, for the works in chapters 9 and 10, we take the right approach for the covariance matrix in their respective catalogs, which results are in concordance with the associated literature.

We want to remark again on the importance of the techniques used in the analysis from the construction of the signal until the information inference. For example, it was an enormous effort to pass from approximately $560,000$ galaxies to just one number α. As we saw, this requires a lot of computation time in data mining. Which motivates us to understand the algorithms involved, but also look for others more efficient and faster. Therefore, we can obtain our observables with a fair inference in the least time, and also we can find new ones.

Without counting the errors in the estimation of the covariance matrix for the New York University Value-Added Galaxy Catalog, detailed and analyzed in section 7.4. We can agree that the inference with a deep learning-based algorithm can give us competent results, which is advantageous, especially for models that require a lot of computation power as simulations or numerical integration, This opens a window for those who could not be tested entirely with traditional inference algorithms. Nonetheless, the likelihood free inference is an open research field, which needs to go deeper to recover many of the physical properties that give us the traditional methods.

We take this as an invitation to research more techniques in the field that can help us with the actual and new models, and that we do not only focus on the theory and how to test it with data because it is a joint effort from several fields as we saw. Thus, we could obtain new information that we ignored before, but also we can save a lot of time using efficient methods, which is becoming important in the Big Data era.

Also, it is essential to highlight the importance to combine data sets to make a better inference. As we saw in chapter 8, the usage of the same geometrical observable $f\sigma_8$ from different surveys, combined with the evolution one H, can constrain and compare several models of dark energy. Together with the right metrics of likeliness, like those used in that chapter, it can help us to discard models or, in our case, to position the ΛCDM model as the preferred. However, it is always important to test more metrics, especially in data where production assumes some reference cosmology. And also, test with other combinations of data.

We also note that we can split the information based on physical properties for the same survey, as we did in chapter 9. Then, we could use those tracers to obtain a better constrain of cosmological information. This is remarkable because it follows the line of making a significant impact without spending new resources. In this case, it is complemented by the physical background that carries the targets. Nonetheless, we need to test in other catalogs and try with more tracers since their bias is fundamental to constrain some theories of structure evolution. What a challenge because if the splitting is disjoint, we could decrease the signal to noise ratio in our estimations. Besides, we need to take into account the astrophysical effects in the model to better understand the constraints.

We also want to remark that better constrains in galaxies clustering require a reliable estimation of the covariance matrix, which, as we highlighted in section 7.4 and chapter 10, is not trivial. As observers inside our Universe, we can not compute precisely an ensemble of Universes, but we can at least try the best possible. And it requires the consideration of all the systematics, astrophysical effects, and cosmological properties to construct a feasible estimation from simulations. On the other hand, we can at least try to build that estimator accessing only to a volume average. Still, we need to go deeper in models that contemplate the cosmic bias not only by galaxies as dark matter tracers but also the cosmic bias due to our average election, as we argued in chapter 4. Therefore, it is crucial to pay attention to the research in these fields as in new galactic formation theories and evolution.

Finally, although the methods and analysis of this thesis were applied to Galaxy Clustering, we can expand the same ideas to other observables and problems that are not related to Cosmology or Astronomy. The techniques presented were mostly developed to use in different fields, and they represent an insight that multidisciplinary is the key to this kind of analysis.

Bibliography

[1] T M C Abbott, F B Abdalla, J Annis, K Bechtol, J Blazek, B A Benson, R A Bernstein, G M Bernstein, E Bertin, D Brooks, and et al. Dark energy survey year 1 results: A precise h0 estimate from des y1, bao, and d/h data. *Monthly Notices of the Royal Astronomical Society*, 480(3):3879–3888, Jul 2018. ISSN 1365-2966. doi: 10.1093/mnras/sty1939. URL `http://dx.doi.org/10.1093/mnras/sty1939`.

[2] T. M. C. Abbott et al. Dark Energy Survey Year 1 Results: Cosmological Constraints from Galaxy Clustering and Weak Lensing. 2017.

[3] T. M. C. Abbott et al. Dark Energy Survey Year 1 Results: Measurement of the Baryon Acoustic Oscillation scale in the distribution of galaxies to redshift 1. *Mon. Not. Roy. Astron. Soc.*, 483(4):4866–4883, 2019. doi: 10.1093/mnras/sty3351.

[4] Paul A. Abell et al. LSST Science Book, Version 2.0. 2009.

[5] L. Raul Abramo and Katie E. Leonard. Why multi-tracer surveys beat cosmic variance. *Mon. Not. Roy. Astron. Soc.*, 432:318, 2013. doi: 10.1093/mnras/stt465.

[6] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 594:A13, 2016. doi: 10.1051/0004-6361/201525830.

[7] Jennifer K. Adelman-McCarthy, Marcel A. Agüeros, Sahar S. Allam, Carlos Allende Prieto, Kurt S.J. Anderson, Scott F. Anderson, James Annis, Neta A. Bahcall, C. A. L. Bailer-Jones, Ivan K. Baldry, and et al. The sixth data release of the sloan digital sky survey. *The Astrophysical Journal Supplement Series*, 175(2):297–313, Apr 2008. ISSN 1538-4365. doi: 10.1086/524984. URL `http://dx.doi.org/10.1086/524984`.

[8] Amir Aghamousa et al. The DESI Experiment Part II: Instrument Design. 2016.

[9] Amir Aghamousa et al. The DESI Experiment Part I: Science,Targeting, and Survey Design. 2016.

[10] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. 2018.

[11] Aniket Agrawal, Ryu Makiya, Chi-Ting Chiang, Donghui Jeong, Shun Saito, and Eiichiro Komatsu. Generating log-normal mock catalog of galaxies in redshift space. J. Cosmology Astropart. Phys., 2017(10):003, October 2017. doi: 10.1088/1475-7516/2017/10/003.

[12] Christopher P. Ahn, Rachael Alexandroff, Carlos Allende Prieto, Scott F. Anderson, Timothy Anderton, Brett H. Andrews, Éric Aubourg, Stephen Bailey, Eduardo Balbinot, Rory Barnes, Julian Bautista, Timothy C. Beers, Alessandra Beifiori, Andreas A. Berlind, Vaishali Bhardwaj, Dmitry Bizyaev, Cullen H. Blake, Michael R. Blanton, Michael Blomqvist, John J. Bochanski, Adam S. Bolton, Arnaud Borde, Jo Bovy, W. N. Brandt, J. Brinkmann, Peter J. Brown, Joel R. Brownstein, Kevin Bundy, N. G. Busca, William Carithers, Aurelio R. Carnero, Michael A. Carr, Dana I. Casetti-Dinescu, Yanmei Chen, Cristina Chiappini, Johan Comparat, Natalia Connolly, Justin R. Crepp, Stefano Cristiani, Rupert A. C. Croft, Antonio J. Cuesta, Luiz N. da Costa, James R. A. Davenport, Kyle S. Dawson, Roland de Putter, Nathan De Lee, Timothée Delubac, Saurav Dhital, Anne Ealet, Garrett L. Ebelke, Edward M. Edmondson, Daniel J. Eisenstein, S. Escoffier, Massimiliano Esposito, Michael L. Evans, Xiaohui Fan, Bruno Femenía Castellá, Emma Fernández Alvar, Leticia D. Ferreira, N. Filiz Ak, Hayley Finley, Scott W. Fleming, Andreu Font-Ribera, Peter M. Frinchaboy, D. A. García-Hernández, A. E. García Pérez, Jian Ge, R. Génova-Santos, Bruce A. Gillespie, Léo Girardi, Jonay I. González Hernández, Eva K. Grebel, James E. Gunn, Hong Guo, Daryl Haggard, Jean-Christophe Hamilton, David W. Harris, Suzanne L. Hawley, Frederick R. Hearty, Shirley Ho, David W. Hogg, Jon A. Holtzman, Klaus Honscheid, J. Huehnerhoff, Inese I. Ivans, Željko Ivezić, Heather R. Jacobson, Linhua Jiang, Jonas Johansson, Jennifer A. Johnson, Guinevere Kauffmann, David Kirkby, Jessica A. Kirkpatrick, Mark A. Klaene, Gillian R. Knapp, Jean-Paul Kneib, Jean-Marc Le Goff, Alexie Leauthaud, Khee-Gan Lee, Young Sun Lee, Daniel C. Long, Craig P. Loomis, Sara Lucatello, Britt Lundgren, Robert H. Lupton, Bo Ma, Zhibo Ma, Nicholas MacDonald, Claude E. Mack, Suvrath Mahadevan, Marcio A. G. Maia, Steven R. Majewski, Martin Makler, Elena Malanushenko, Viktor Malanushenko, A. Manchado, Rachel Mandelbaum, Marc Manera, Claudia Maraston, Daniel Margala, Sarah L. Martell, Cameron K. McBride, Ian D. McGreer, Richard G. McMahon, Brice Ménard, Sz. Meszaros, Jordi Miralda-Escudé, Antonio D. Montero-Dorta, Francesco Montesano, Heather L. Morrison, Demitri Muna, Jeffrey A. Munn, Hitoshi Murayama, Adam D. Myers, A. F. Neto, Duy Cuong Nguyen, Robert C. Nichol, David L. Nidever, Pasquier Noterdaeme, Sebastián E. Nuza, Ricardo L. C. Ogand o, Matthew D. Olmstead, Daniel J. Oravetz, Russell Owen, Nikhil Padmanabhan, Nathalie Palanque-Delabrouille, Kaike Pan, John K. Parejko, Prachi Parihar, Isabelle Pâris, Petchara Pattarakijwanich, Joshua Pepper, Will J. Percival, Ismael Pérez-Fournon, Ignasi Pérez-Ràfols, Patrick Petitjean, Janine Pforr, Matthew M. Pieri, Marc H. Pinsonneault, G. F. Porto de Mello, Francisco Prada, Adrian M. Price-Whelan, M. Jordan Raddick, Rafael Rebolo, James Rich, Gordon T. Richards, Annie C. Robin, Helio J. Rocha-Pinto, Constance M. Rockosi, Natalie A. Roe, Ashley J. Ross, Nicholas P. Ross, Graziano Rossi, J. A. Rubiño-Martin, Lado Samushia, J. Sanchez Almeida, Ariel G. Sánchez, Basílio Santiago, Conor Sayres, David J. Schlegel, Katharine J. Schlesinger, Sarah J. Schmidt, Donald P. Schneider, Mathias Schultheis, Axel D. Schwope, C. G. Scóccola, Uros Seljak, Erin Sheldon, Yue Shen, Yiping Shu, Jennifer Simmerer, Audrey E. Simmons, Ramin A. Skibba, M. F. Skrutskie, A. Slosar, Flavia Sobreira, Jennifer S. Sobeck, Keivan G. Stassun, Oliver Steele, Matthias Steinmetz, Michael A. Strauss, Alina Streblyanska, Nao Suzuki, Molly E. C. Swanson, Tomer Tal, Aniruddha R. Thakar, Daniel Thomas, Benjamin A. Thompson, Jeremy L. Tinker, Rita Tojeiro, Christy A. Tremonti, M. Vargas Magaña, Licia Verde, Matteo Viel, Shailendra K. Vikas, Nicole P. Vogt,

David A. Wake, Ji Wang, Benjamin A. Weaver, David H. Weinberg, Benjamin J. Weiner, Andrew A. West, Martin White, John C. Wilson, John P. Wisniewski, W. M. Wood-Vasey, Brian Yanny, Christophe Yèche, Donald G. York, O. Zamora, Gail Zasowski, Idit Zehavi, Gong-Bo Zhao, Zheng Zheng, Guangtun Zhu, and Joel C. Zinn. The Ninth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Baryon Oscillation Spectroscopic Survey. ApJS, 203(2):21, Dec 2012. doi: 10.1088/0067-0049/203/2/21.

[13] H. Aihara, C. Allende Prieto, D. An, S. F. Anderson, É. Aubourg, E. Balbinot, T. C. Beers, A. A. Berlind, S. J. Bickerton, D. Bizyaev, M. R. Blanton, J. J. Bochanski, A. S. Bolton, J. Bovy, W. N. Brandt, J. Brinkmann, P. J. Brown, J. R. Brownstein, N. G. Busca, H. Campbell, M. A. Carr, Y. Chen, C. Chiappini, J. Comparat, N. Connolly, M. Cortes, R. A. C. Croft, A. J. Cuesta, L. N. da Costa, J. R. A. Davenport, K. Dawson, S. Dhital, A. Ealet, G. L. Ebelke, E. M. Edmondson, D. J. Eisenstein, S. Escoffier, M. Esposito, M. L. Evans, X. Fan, B. Femenía Castellá, A. Font-Ribera, P. M. Frinchaboy, J. Ge, B. A. Gillespie, G. Gilmore, J. I. González Hernández, J. R. Gott, A. Gould, E. K. Grebel, J. E. Gunn, J.-C. Hamilton, P. Harding, D. W. Harris, S. L. Hawley, F. R. Hearty, S. Ho, D. W. Hogg, J. A. Holtzman, K. Honscheid, N. Inada, I. I. Ivans, L. Jiang, J. A. Johnson, C. Jordan, W. P. Jordan, E. A. Kazin, D. Kirkby, M. A. Klaene, G. R. Knapp, J.-P. Kneib, C. S. Kochanek, L. Koesterke, J. A. Kollmeier, R. G. Kron, H. Lampeitl, D. Lang, J.-M. Le Goff, Y. S. Lee, Y.-T. Lin, D. C. Long, C. P. Loomis, S. Lucatello, B. Lundgren, R. H. Lupton, Z. Ma, N. MacDonald, S. Mahadevan, M. A. G. Maia, M. Makler, E. Malanushenko, V. Malanushenko, R. Mandelbaum, C. Maraston, D. Margala, K. L. Masters, C. K. McBride, P. M. McGehee, I. D. McGreer, B. Ménard, J. Miralda-Escudé, H. L. Morrison, F. Mullally, D. Muna, J. A. Munn, H. Murayama, A. D. Myers, T. Naugle, A. Fausti Neto, D. Cuong Nguyen, R. C. Nichol, R. W. O'Connell, R. L. C. Ogando, M. D. Olmstead, D. J. Oravetz, N. Padmanabhan, N. Palanque-Delabrouille, K. Pan, P. Pandey, I. Pâris, W. J. Percival, P. Petitjean, R. Pfaffenberger, J. Pforr, S. Phleps, C. Pichon, M. M. Pieri, F. Prada, A. M. Price-Whelan, M. J. Raddick, B. H. F. Ramos, C. Reylé, J. Rich, G. T. Richards, H.-W. Rix, A. C. Robin, H. J. Rocha-Pinto, C. M. Rockosi, N. A. Roe, E. Rollinde, A. J. Ross, N. P. Ross, B. M. Rossetto, A. G. Sánchez, C. Sayres, D. J. Schlegel, K. J. Schlesinger, S. J. Schmidt, D. P. Schneider, E. Sheldon, Y. Shu, J. Simmerer, A. E. Simmons, T. Sivarani, S. A. Snedden, J. S. Sobeck, M. Steinmetz, M. A. Strauss, A. S. Szalay, M. Tanaka, A. R. Thakar, D. Thomas, J. L. Tinker, B. M. Tofflemire, R. Tojeiro, C. A. Tremonti, J. Vandenberg, M. Vargas Magaña, L. Verde, N. P. Vogt, D. A. Wake, J. Wang, B. A. Weaver, D. H. Weinberg, M. White, S. D. M. White, B. Yanny, N. Yasuda, C. Yeche, and I. Zehavi. The Eighth Data Release of the Sloan Digital Sky Survey: First Data from SDSS-III. ApJS, 193:29–+, April 2011. doi: 10.1088/0067-0049/193/2/29.

[14] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, Dec 1974. ISSN 0018-9286. doi: 10.1109/TAC.1974.1100705.

[15] Shadab Alam, Franco D. Albareti, Carlos Allende Prieto, F. Anders, Scott F. Anderson, Timothy Anderton, Brett H. Andrews, Eric Armengaud, Éric Aubourg, Stephen Bai-

ley, Sarbani Basu, Julian E. Bautista, Rachael L. Beaton, Timothy C. Beers, Chad F. Bender, Andreas A. Berlind, Florian Beutler, Vaishali Bhardwaj, Jonathan C. Bird, Dmitry Bizyaev, Cullen H. Blake, Michael R. Blanton, Michael Blomqvist, John J. Bochanski, Adam S. Bolton, Jo Bovy, A. Shelden Bradley, W. N. Brandt, D. E. Brauer, J. Brinkmann, Peter J. Brown, Joel R. Brownstein, Angela Burden, Etienne Burtin, Nicolás G. Busca, Zheng Cai, Diego Capozzi, Aurelio Carnero Rosell, Michael A. Carr, Ricardo Carrera, K. C. Chambers, William James Chaplin, Yen-Chi Chen, Cristina Chiappini, S. Drew Chojnowski, Chia-Hsun Chuang, Nicolas Clerc, Johan Comparat, Kevin Covey, Rupert A. C. Croft, Antonio J. Cuesta, Katia Cunha, Luiz N. da Costa, Nicola Da Rio, James R. A. Davenport, Kyle S. Dawson, Nathan De Lee, Timothée Delubac, Rohit Deshpande, Saurav Dhital, Letícia Dutra-Ferreira, Tom Dwelly, Anne Ealet, Garrett L. Ebelke, Edward M. Edmondson, Daniel J. Eisenstein, Tristan Ellsworth, Yvonne Elsworth, Courtney R. Epstein, Michael Eracleous, Stephanie Escoffier, Massimiliano Esposito, Michael L. Evans, Xiaohui Fan, Emma Fernández-Alvar, Diane Feuillet, Nurten Filiz Ak, Hayley Finley, Alexis Finoguenov, Kevin Flaherty, Scott W. Fleming, Andreu Font-Ribera, Jonathan Foster, Peter M. Frinchaboy, J. G. Galbraith-Frew, Rafael A. García, D. A. García-Hernández, Ana E. García Pérez, Patrick Gaulme, Jian Ge, R. Génova-Santos, A. Georgakakis, Luan Ghezzi, Bruce A. Gillespie, Léo Girardi, Daniel Goddard, Satya Gontcho A. Gontcho, Jonay I. González Hernández, Eva K. Grebel, Paul J. Green, Jan Niklas Grieb, Nolan Grieves, James E. Gunn, Hong Guo, Paul Harding, Sten Hasselquist, Suzanne L. Hawley, Michael Hayden, Fred R. Hearty, Saskia Hekker, Shirley Ho, David W. Hogg, Kelly Holley-Bockelmann, Jon A. Holtzman, Klaus Honscheid, Daniel Huber, Joseph Huehnerhoff, Inese I. Ivans, Linhua Jiang, Jennifer A. Johnson, Karen Kinemuchi, David Kirkby, Francisco Kitaura, Mark A. Klaene, Gillian R. Knapp, Jean-Paul Kneib, Xavier P. Koenig, Charles R. Lam, Ting-Wen Lan, Dustin Lang, Pierre Laurent, Jean-Marc Le Goff, Alexie Leauthaud, Khee-Gan Lee, Young Sun Lee, Timothy C. Licquia, Jian Liu, Daniel C. Long, Martín López-Corredoira, Diego Lorenzo-Oliveira, Sara Lucatello, Britt Lundgren, Robert H. Lupton, III Mack, Claude E., Suvrath Mahadevan, Marcio A. G. Maia, Steven R. Majewski, Elena Malanushenko, Viktor Malanushenko, A. Manchado, Marc Manera, Qingqing Mao, Claudia Maraston, Robert C. Marchwinski, Daniel Margala, Sarah L. Martell, Marie Martig, Karen L. Masters, Savita Mathur, Cameron K. McBride, Peregrine M. McGehee, Ian D. McGreer, Richard G. McMahon, Brice Ménard, Marie-Luise Menzel, Andrea Merloni, Szabolcs Mészáros, Adam A. Miller, Jordi Miralda-Escudé, Hironao Miyatake, Antonio D. Montero-Dorta, Surhud More, Eric Morganson, Xan Morice-Atkinson, Heather L. Morrison, Benôit Mosser, Demitri Muna, Adam D. Myers, Kirpal Nand ra, Jeffrey A. Newman, Mark Neyrinck, Duy Cuong Nguyen, Robert C. Nichol, David L. Nidever, Pasquier Noterdaeme, Sebastián E. Nuza, Julia E. O'Connell, Robert W. O'Connell, Ross O'Connell, Ricardo L. C. Ogando, Matthew D. Olmstead, Audrey E. Oravetz, Daniel J. Oravetz, Keisuke Osumi, Russell Owen, Deborah L. Padgett, Nikhil Padmanabhan, Martin Paegert, Nathalie Palanque-Delabrouille, Kaike Pan, John K. Parejko, Isabelle Pâris, Changbom Park, Petchara Pattarakijwanich, M. Pellejero-Ibanez, Joshua Pepper, Will J. Percival, Ismael Pérez-Fournon, Ignasi Pérez-Ràfols, Patrick Petitjean, Matthew M. Pieri, Marc H. Pinsonneault, Gustavo F. Porto de Mello, Francisco Prada, Abhishek Prakash, Adrian M. Price-Whelan, Pavlos Protopapas, M. Jordan Raddick, Mubdi Rahman, Beth A. Reid, James Rich, Hans-Walter Rix, Annie C. Robin, Constance M. Rockosi, Thaíse S. Rodrigues, Sergio Rodríguez-Torres, Natalie A. Roe, Ashley J. Ross, Nicholas P. Ross, Graziano Rossi, John J. Ruan, J. A. Rubiño-Martín, Eli S. Rykoff, Salvador Salazar-Albornoz, Mara Salvato, Lado Samushia, Ariel G. Sánchez, Basílio Santiago, Conor Sayres, Ricardo P. Schiavon, David J. Schlegel, Sarah J. Schmidt, Donald P. Schneider, Mathias Schultheis, Axel D. Schwope, C. G. Scóccola, Caroline Scott, Kris Sellgren, Hee-Jong Seo, Aldo Serenelli, Neville Shane, Yue Shen, Matthew Shetrone, Yiping Shu, V. Silva Aguirre, Thirupathi Sivarani, M. F. Skrutskie, Anže Slosar, Verne V. Smith, Flávia Sobreira, Diogo Souto, Keivan G. Stassun, Matthias Steinmetz, Dennis Stello, Michael A. Strauss, Alina Streblyanska, Nao Suzuki, Molly E. C. Swanson, Jonathan C. Tan, Jamie Tayar, Ryan C. Terrien, Aniruddha R. Thakar, Daniel Thomas, Neil Thomas, Benjamin A. Thompson, Jeremy L. Tinker, Rita Tojeiro, Nicholas W. Troup, Mariana Vargas-Magaña, Jose A. Vazquez, Licia Verde, Matteo Viel, Nicole P. Vogt, David A. Wake, Ji Wang, Benjamin A. Weaver, David H. Weinberg, Benjamin J. Weiner, Martin White, John C. Wilson, John P. Wisniewski, W. M. Wood-Vasey, Christophe Ye‘che, Donald G. York, Nadia L. Zakamska, O. Zamora, Gail Zasowski, Idit Zehavi, Gong-Bo Zhao, Zheng Zheng, Xu Zhou, Zhimin Zhou, Hu Zou, and Guangtun Zhu. The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III. ApJS, 219(1):12, Jul 2015. doi: 10.1088/0067-0049/219/1/12.

[16] Shadab Alam, Shirley Ho, and Alessandra Silvestri. Testing deviations from ΛCDM with growth rate measurements from six large-scale structure surveys at z= 0.06–1. *Monthly Notices of the Royal Astronomical Society*, 456(4):3743–3756, 01 2016. ISSN 0035-8711. doi: 10.1093/mnras/stv2935. URL `https://doi.org/10.1093/mnras/stv2935`.

[17] Shadab Alam, Metin Ata, Stephen Bailey, Florian Beutler, Dmitry Bizyaev, Jonathan A. Blazek, Adam S. Bolton, Joel R. Brownstein, Angela Burden, Chia-Hsun Chuang, Johan Comparat, Antonio J. Cuesta, Kyle S. Dawson, Daniel J. Eisenstein, Stephanie Escoffier, Héctor Gil-Marín, Jan Niklas Grieb, Nick Hand, Shirley Ho, Karen Kinemuchi, David Kirkby, Francisco Kitaura, Elena Malanushenko, Viktor Malanushenko, Claudia Maraston, Cameron K. McBride, Robert C. Nichol, Matthew D. Olmstead, Daniel Oravetz, Nikhil Padmanabhan, Nathalie Palanque-Delabrouille, Kaike Pan, Marcos Pellejero-Ibanez, Will J. Percival, Patrick Petitjean, Francisco Prada, Adrian M. Price-Whelan, Beth A. Reid, Sergio A. Rodríguez-Torres, Natalie A. Roe, Ashley J. Ross, Nicholas P. Ross, Graziano Rossi, Jose Alberto Rubiño-Martín, Shun Saito, Salvador Salazar-Albornoz, Lado Samushia, Ariel G. Sánchez, Siddharth Satpathy, David J. Schlegel, Donald P. Schneider, Claudia G. Scóccola, Hee-Jong Seo, Erin S. Sheldon, Audrey Simmons, Anže Slosar, Michael A. Strauss, Molly E. C. Swanson, Daniel Thomas, Jeremy L. Tinker, Rita Tojeiro, Mariana Vargas Magaña, Jose Alberto Vazquez, Licia Verde, David A. Wake, Yuting Wang, David H. Weinberg, Martin White, W. Michael Wood-Vasey, Christophe Yèche, Idit Zehavi, Zhongxu Zhai, and Gong-Bo Zhao. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. MNRAS, 470(3):2617–2652, Sep 2017. doi: 10.1093/mnras/stx721.

[18] C. Alcock and B. Paczynski. An evolution free test for non-zero cosmological constant. *Nature*, 281:358–359, 1979. doi: 10.1038/281358a0.

[19] Yacine Ali-Haïmoud and Simeon Bird. An efficient implementation of massive neutrinos in non-linear structure formation simulations. MNRAS, 428(4):3375–3389, February 2013. doi: 10.1093/mnras/sts286.

[20] Justin Alsing and Benjamin Wandelt. Generalized massive optimal data compression. *Monthly Notices of the Royal Astronomical Society: Letters*, 476(1):L60–L64, Feb 2018. ISSN 1745-3933. doi: 10.1093/mnrasl/sly029. URL `http://dx.doi.org/10.1093/mnrasl/sly029`.

[21] Justin Alsing, Benjamin Wandelt, and Stephen Feeney. Massive optimal data compression and density estimation for scalable, likelihood-free inference in cosmology. *Monthly Notices of the Royal Astronomical Society*, 477(3):2874–2885, Mar 2018. ISSN 1365-2966. doi: 10.1093/mnras/sty819. URL `http://dx.doi.org/10.1093/mnras/sty819`.

[22] Justin Alsing, Tom Charnock, Stephen Feeney, and Benjamin Wandelt. Fast likelihood-free cosmology with neural density estimators and active learning. *Monthly Notices of the Royal Astronomical Society*, Jul 2019. ISSN 1365-2966. doi: 10.1093/mnras/stz1960. URL `http://dx.doi.org/10.1093/mnras/stz1960`.

[23] Luca Amendola and Shinji Tsujikawa. *Dark Energy: Theory and Observations*. Cambridge University Press, 2010. doi: 10.1017/CBO9780511750823.

[24] Luca Amendola et al. Cosmology and fundamental physics with the Euclid satellite. *Living Rev. Rel.*, 21(1):2, 2018. doi: 10.1007/s41114-017-0010-3.

[25] Lauren Anderson, Eric Aubourg, Stephen Bailey, Dmitry Bizyaev, Michael Blanton, Adam S. Bolton, J. Brinkmann, Joel R. Brownstein, Angela Burden, Antonio J. Cuesta, Luiz A. N. da Costa, Kyle S. Dawson, Roland de Putter, Daniel J. Eisenstein, James E. Gunn, Hong Guo, Jean-Christophe Hamilton, Paul Harding, Shirley Ho, Klaus Honscheid, Eyal Kazin, David Kirkby, Jean-Paul Kneib, Antoine Labatie, Craig Loomis, Robert H. Lupton, Elena Malanushenko, Viktor Malanushenko, Rachel Mandelbaum, Marc Manera, Claudia Maraston, Cameron K. McBride, Kushal T. Mehta, Olga Mena, Francesco Montesano, Demetri Muna, Robert C. Nichol, Sebastián E. Nuza, Matthew D. Olmstead, Daniel Oravetz, Nikhil Padmanabhan, Nathalie Palanque-Delabrouille, Kaike Pan, John Parejko, Isabelle Pâris, Will J. Percival, Patrick Petitjean, Francisco Prada, Beth Reid, Natalie A. Roe, Ashley J. Ross, Nicholas P. Ross, Lado Samushia, Ariel G. Sánchez, David J. Schlegel, Donald P. Schneider, Claudia G. Scóccola, Hee-Jong Seo, Erin S. Sheldon, Audrey Simmons, Ramin A. Skibba, Michael A. Strauss, Molly E. C. Swanson, Daniel Thomas, Jeremy L. Tinker, Rita Tojeiro, Mariana Vargas Magaña, Licia Verde, Christian Wagner, David A. Wake, Benjamin A. Weaver, David H. Weinberg, Martin White, Xiaoying Xu, Christophe Yèche, Idit Zehavi, and Gong-Bo Zhao. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Release 9 spectroscopic galaxy sample. MNRAS, 427(4):3435–3467, Dec 2012. doi: 10.1111/j.1365-2966.2012.22066.x.

[26] Santiago Avila, Steven G. Murray, Alexander Knebe, Chris Power, Aaron S. G. Robotham, and Juan Garcia-Bellido. HALOGEN: a tool for fast generation of mock

halo catalogues. MNRAS, 450(2):1856–1867, June 2015. doi: 10.1093/mnras/stv711.

[27] Kyle Barbary. Nestle. `https://github.com/kbarbary/nestle`, 2015.

[28] James M. Bardeen. Gauge-invariant cosmological perturbations. *Phys. Rev. D*, 22: 1882–1905, Oct 1980. doi: 10.1103/PhysRevD.22.1882. URL `https://link.aps.org/doi/10.1103/PhysRevD.22.1882`.

[29] Spyros Basilakos and Savvas Nesseris. Conjoined constraints on modified gravity from the expansion history and cosmic growth. *Phys. Rev.*, D96(6):063517, 2017. doi: 10.1103/PhysRevD.96.063517.

[30] C. M. Baugh. A primer on hierarchical galaxy formation: the semi-analytical approach. *Reports on Progress in Physics*, 69:3101–3156, December 2006. doi: 10.1088/0034-4885/69/12/R02.

[31] Daniel Baumann. Cosmology, part III mathematical tripos.

[32] Daniel Baumann. Tasi lectures on inflation, 2009.

[33] Robert H. Becker, Richard L. White, and David J. Helfand. The FIRST Survey: Faint Images of the Radio Sky at Twenty Centimeters. ApJ, 450:559, September 1995. doi: 10.1086/176166.

[34] A. J. Benson. G ALACTICUS: A semi-analytic model of galaxy formation. 17:175–197, February 2012. doi: 10.1016/j.newast.2011.07.004.

[35] Andrew J. Benson. Galaxy formation theory. Phys. Rep., 495(2-3):33–86, October 2010. doi: 10.1016/j.physrep.2010.06.001.

[36] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro. Large-scale structure of the universe and cosmological perturbation theory. *Physics Reports*, 367(1-3):1–248, Sep 2002. ISSN 0370-1573. doi: 10.1016/s0370-1573(02)00135-7. URL `http://dx.doi.org/10.1016/S0370-1573(02)00135-7`.

[37] M. Betoule et al. Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples. *Astron. Astrophys.*, 568:A22, 2014. doi: 10.1051/0004-6361/201423413.

[38] Florian Beutler, Chris Blake, Matthew Colless, D. Heath Jones, Lister Staveley-Smith, Lachlan Campbell, Quentin Parker, Will Saunders, and Fred Watson. The 6dF Galaxy Survey: baryon acoustic oscillations and the local Hubble constant. MNRAS, 416(4): 3017–3032, October 2011. doi: 10.1111/j.1365-2966.2011.19250.x.

[39] Christopher M. Bishop. Mixture density networks. Technical report, 1994.

[40] J.D. Bjorken and S.D. Drell. *Relativistic quantum fields*. International series in pure and applied physics. McGraw-Hill, 1965. URL `https://books.google.cl/books?id=ZczvAAAAMAAJ`.

[41] Chris Blake et al. The WiggleZ Dark Energy Survey: testing the cosmological model with baryon acoustic oscillations at z=0.6. *Mon. Not. Roy. Astron. Soc.*, 415:2892–2909, 2011. doi: 10.1111/j.1365-2966.2011.19077.x.

[42] Chris Blake et al. The WiggleZ Dark Energy Survey: Joint measurements of the expansion and growth history at z $<$ 1. *Mon. Not. Roy. Astron. Soc.*, 425:405–414, 2012. doi: 10.1111/j.1365-2966.2012.21473.x.

[43] Chris Blake et al. Galaxy And Mass Assembly (GAMA): improved cosmic growth measurements using multiple tracers of large-scale structure. *Mon. Not. Roy. Astron. Soc.*, 436:3089, 2013. doi: 10.1093/mnras/stt1791.

[44] A. Blanchard et al. Euclid preparation: VII. Forecast validation for Euclid cosmological probes. 2019.

[45] M. R. Blanton, D. J. Schlegel, M. A. Strauss, J. Brinkmann, D. Finkbeiner, M. Fukugita, J. E. Gunn, D. W. Hogg, Ž. Ivezić, G. R. Knapp, R. H. Lupton, J. A. Munn, D. P. Schneider, M. Tegmark, and I. Zehavi. New York University Value-Added Galaxy Catalog: A Galaxy Catalog Based on New Public Surveys. AJ, 129:2562–2578, June 2005. doi: 10.1086/429803.

[46] Diego Blas, Julien Lesgourgues, and Thomas Tram. The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes. J. Cosmology Astropart. Phys., 2011(7):034, July 2011. doi: 10.1088/1475-7516/2011/07/034.

[47] Adam S. Bolton, David J. Schlegel, Éric Aubourg, Stephen Bailey, Vaishali Bhardwaj, Joel R. Brownstein, Scott Burles, Yan-Mei Chen, Kyle Dawson, Daniel J. Eisenstein, James E. Gunn, G. R. Knapp, Craig P. Loomis, Robert H. Lupton, Claudia Maraston, Demitri Muna, Adam D. Myers, Matthew D. Olmstead, Nikhil Padmanabhan, Isabelle Pâris, Will J. Percival, Patrick Petitjean, Constance M. Rockosi, Nicholas P. Ross, Donald P. Schneider, Yiping Shu, Michael A. Strauss, Daniel Thomas, Christy A. Tremonti, David A. Wake, Benjamin A. Weaver, and W. Michael Wood-Vasey. Spectral Classification and Redshift Measurement for the SDSS-III Baryon Oscillation Spectroscopic Survey. AJ, 144(5):144, Nov 2012. doi: 10.1088/0004-6256/144/5/144.

[48] Fernando V. Bonassi, Lingchong You, and Mike West. Bayesian learning from marginal data in bionetwork models. *Statistical Applications in Genetics and Molecular Biology*, 10(1), 2011. doi: https://doi.org/10.2202/1544-6115.1684. URL `https://www.degruyter.com/view/journals/sagmb/10/1/article-1544-6115.1684.xml.xml`.

[49] Camille Bonvin, Ruth Durrer, and Martin Kunz. The dipole of the luminosity distance: a direct measure of H(z). *Phys. Rev. Lett.*, 96:191302, 2006. doi: 10.1103/PhysRevLett.96.191302.

[50] R. G. Bower, A. J. Benson, R. Malbon, J. C. Helly, C. S. Frenk, C. M. Baugh, S. Cole, and C. G. Lacey. Breaking the hierarchy of galaxy formation. MNRAS, 370:645–655, August 2006. doi: 10.1111/j.1365-2966.2006.10519.x.

[51] Alicia Bueno Belloso, Juan Garcia-Bellido, and Domenico Sapone. A parametriza-

tion of the growth index of matter perturbations in various Dark Energy models and observational prospects using a Euclid-like survey. *JCAP*, 1110:010, 2011. doi: 10.1088/1475-7516/2011/10/010.

[52] E. Cameron and A. N. Pettitt. Approximate Bayesian Computation for astronomical model analysis: a case study in galaxy demographics and morphological transformation at high redshift. *Monthly Notices of the Royal Astronomical Society*, 425(1):44–65, 09 2012. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2012.21371.x. URL `https://doi.org/10.1111/j.1365-2966.2012.21371.x`.

[53] Michel Chevallier and David Polarski. Accelerating universes with scaling dark matter. *Int. J. Mod. Phys.*, D10:213–224, 2001. doi: 10.1142/S0218271801000822.

[54] Chia-Hsun Chuang and Yun Wang. Measurements of H(z) and DA(z) from the two-dimensional two-point correlation function of Sloan Digital Sky Survey luminous red galaxies. *Monthly Notices of the Royal Astronomical Society*, 426(1):226–236, 10 2012. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2012.21565.x.

[55] Chia-Hsun Chuang, Francisco-Shu Kitaura, Francisco Prada, Cheng Zhao, and Gustavo Yepes. EZmocks: extending the Zel'dovich approximation to generate mock galaxy catalogues with accurate clustering statistics. MNRAS, 446(3):2621–2628, January 2015. doi: 10.1093/mnras/stu2301.

[56] Chia-Hsun Chuang et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: single-probe measurements from CMASS anisotropic galaxy clustering. *Mon. Not. Roy. Astron. Soc.*, 461(4):3781–3793, 2016. doi: 10.1093/mnras/stw1535.

[57] Shaun Cole et al. The 2dF Galaxy Redshift Survey: Power-spectrum analysis of the final dataset and cosmological implications. *Mon. Not. Roy. Astron. Soc.*, 362:505–534, 2005. doi: 10.1111/j.1365-2966.2005.09318.x.

[58] Peter Coles and Bernard Jones. A lognormal model for the cosmological mass distribution. MNRAS, 248:1–13, January 1991. doi: 10.1093/mnras/248.1.1.

[59] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, J. Chluba, L. P. L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, J. Delabrouille, J. M. Delouis, E. Di Valentino, J. M. Diego, O. Doré, M. Douspis, A. Ducout, X. Dupac, S. Dusini, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, Y. Fantaye, M. Farhang, J. Fergusson, R. Fernandez-Cobos, F. Finelli, F. Forastieri, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gruppuso, J. E. Gudmundsson, J. Hamann, W. Handley, F. K. Hansen, D. Herranz, S. R. Hildebrandt, E. Hivon, Z. Huang, A. H. Jaffe, W. C. Jones, A. Karakci, E. Keihä-

nen, R. Keskitalo, K. Kiiveri, J. Kim, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, J. M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, P. Lemos, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Lilley, V. Lindholm, M. López-Caniego, P. M. Lubin, Y. Z. Ma, J. F. Macías-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marcos-Caballero, M. Maris, P. G. Martin, M. Martinelli, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M. A. Miville-Deschênes, D. Molinari, L. Montier, G. Morgante, A. Moss, P. Natoli, H. U. Nørgaard-Nielsen, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, H. V. Peiris, F. Perrotta, V. Pettorino, F. Piacentini, L. Polastri, G. Polenta, J. L. Puget, J. P. Rachen, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, C. Rosset, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, E. P. S. Shellard, C. Sirignano, G. Sirri, L. D. Spencer, R. Sunyaev, A. S. Suur-Uski, J. A. Tauber, D. Tavagnacco, M. Tenti, L. Toffolatti, M. Tomasi, T. Trombetti, L. Valenziano, J. Valiviita, B. Van Tent, L. Vibert, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K. Wehus, M. White, S. D. M. White, A. Zacchei, and A. Zonca. Planck 2018 results. vi. cosmological parameters, 2018.

[60] Florencia Collacchioni, Sofía A. Cora, Claudia D. P. Lagos, and Cristian A. Vega-Martínez. Semi-analytic galaxies - III. The impact of supernova feedback on the mass-metallicity relation. MNRAS, 481(1):954–969, Nov 2018. doi: 10.1093/mnras/sty2347.

[61] Matthew Colless, Gavin Dalton, Steve Maddox, Will Sutherland , Peder Norberg, Shaun Cole, Joss Bland -Hawthorn, Terry Bridges, Russell Cannon, Chris Collins, Warrick Couch, Nicholas Cross, Kathryn Deeley, Roberto De Propris, Simon P. Driver, George Efstathiou, Richard S. Ellis, Carlos S. Frenk, Karl Glazebrook, Carole Jackson, Ofer Lahav, Ian Lewis, Stuart Lumsden, Darren Madgwick, John A. Peacock, Bruce A. Peterson, Ian Price, Mark Seaborne, and Keith Taylor. The 2dF Galaxy Redshift Survey: spectra and redshifts. MNRAS, 328(4):1039–1063, December 2001. doi: 10.1046/j.1365-8711.2001.04902.x.

[62] S. A. Cora. Metal enrichment of the intracluster medium: a three-dimensional picture of chemical and dynamical properties. MNRAS, 368:1540–1560, June 2006. doi: 10.1111/j.1365-2966.2006.10271.x.

[63] S. A. Cora, C. A. Vega-Martínez, T. Hough, A. N. Ruiz, Á. A. Orsi, A. M. Muñoz Arancibia, I. D. Gargiulo, F. Collacchioni, N. D. Padilla, S. Gottlöber, and G. Yepes. Semi-analytic galaxies - I. Synthesis of environmental and star-forming regulation mechanisms. MNRAS, 479:2–24, September 2018. doi: 10.1093/mnras/sty1131.

[64] Robert A. Crain, Joop Schaye, Richard G. Bower, Michelle Furlong, Matthieu Schaller, Tom Theuns, Claudio Dalla Vecchia, Carlos S. Frenk, Ian G. McCarthy, John C. Helly, Adrian Jenkins, Yetli M. Rosas-Guevara, Simon D. M. White, and James W. Trayford. The EAGLE simulations of galaxy formation: calibration of subgrid physics and model variations. MNRAS, 450(2):1937–1961, June 2015. doi: 10.1093/mnras/stv725.

[65] Martin Crocce and Roman Scoccimarro. Renormalized cosmological perturbation theory. *Phys. Rev.*, D73:063519, 2006. doi: 10.1103/PhysRevD.73.063519.

[66] Martín Crocce and Román Scoccimarro. Nonlinear evolution of baryon acoustic oscillations. *Physical Review D*, 77(2), Jan 2008. ISSN 1550-2368. doi: 10.1103/physrevd.77.023533. URL `http://dx.doi.org/10.1103/PhysRevD.77.023533`.

[67] D. J. Croton, V. Springel, S. D. M. White, G. De Lucia, C. S. Frenk, L. Gao, A. Jenkins, G. Kauffmann, J. F. Navarro, and N. Yoshida. The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. MNRAS, 365: 11–28, January 2006. doi: 10.1111/j.1365-2966.2005.09675.x.

[68] D. J. Croton, A. R. H. Stevens, C. Tonini, T. Garel, M. Bernyk, A. Bibiano, L. Hodkinson, S. J. Mutch, G. B. Poole, and G. M. Shattow. Semi-Analytic Galaxy Evolution (SAGE): Model Calibration and Basic Results. ApJS, 222:22, February 2016. doi: 10.3847/0067-0049/222/2/22.

[69] M. Davis and P. J. E. Peebles. A survey of galaxy redshifts. V. The two-point position and velocity correlations. ApJ, 267:465–482, April 1983. doi: 10.1086/160884.

[70] Marc Davis, Adi Nusser, Karen Masters, Christopher Springob, John P. Huchra, and Gerard Lemson. Local Gravity versus Local Velocity: Solutions for β and nonlinear bias. *Mon. Not. Roy. Astron. Soc.*, 413:2906, 2011. doi: 10.1111/j.1365-2966.2011.18362.x.

[71] A. C. Davison and D. V. Hinkley. *Frontmatter*, pages i–iv. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997.

[72] Kyle S. Dawson et al. The Baryon Oscillation Spectroscopic Survey of SDSS-III. *Astron. J.*, 145:10, 2013. doi: 10.1088/0004-6256/145/1/10.

[73] Kyle S. Dawson et al. The SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data. *Astron. J.*, 151:44, 2016. doi: 10.3847/0004-6256/151/2/44.

[74] W De Boer. Grand unified theories and supersymmetry in particle physics and cosmology. *Progress in Particle and Nuclear Physics*, 33:201–301, Jan 1994. ISSN 0146-6410. doi: 10.1016/0146-6410(94)90045-0. URL `http://dx.doi.org/10.1016/0146-6410(94)90045-0`.

[75] S. de la Torre et al. The VIMOS Public Extragalactic Redshift Survey (VIPERS). Galaxy clustering and redshift-space distortions at z=0.8 in the first data release. *Astron. Astrophys.*, 557:A54, 2013. doi: 10.1051/0004-6361/201321463.

[76] Sylvain de la Torre and Luigi Guzzo. Modelling non-linear redshift-space distortions in the galaxy clustering pattern: systematic errors on the growth rate parameter. *Monthly Notices of the Royal Astronomical Society*, 427(1):327–342, Oct 2012. ISSN 1365-2966. doi: 10.1111/j.1365-2966.2012.21824.x. URL `http://dx.doi.org/10.1111/j.1365-2966.2012.21824.x`.

[77] Gabriella De Lucia, Michael Boylan-Kolchin, Andrew J. Benson, Fabio Fontanot, and Pierluigi Monaco. A semi-analytic model comparison – gas cooling and galaxy mergers. *Monthly Notices of the Royal Astronomical Society*, 406(3):1533–1552, 08 2010. ISSN

0035-8711. doi: 10.1111/j.1365-2966.2010.16806.x. URL `https://doi.org/10.1111/j.1365-2966.2010.16806.x`.

[78] Gerard de Vaucouleurs, Antoinette de Vaucouleurs, Jr. Corwin, Herold G., Ronald J. Buta, Georges Paturel, and Pascal Fouque. *Third Reference Catalogue of Bright Galaxies.* 1991.

[79] Timothée Delubac et al. Baryon acoustic oscillations in the Lyα forest of BOSS DR11 quasars. *Astron. Astrophys.*, 574:A59, 2015. doi: 10.1051/0004-6361/201423969.

[80] Scott Dodelson. *Modern cosmology.* Academic Press, San Diego, CA, 2003. URL `https://cds.cern.ch/record/1282338`.

[81] B. Efron. Bootstrap methods: Another look at the jackknife. *Ann. Statist.*, 7(1): 1–26, 01 1979. doi: 10.1214/aos/1176344552. URL `https://doi.org/10.1214/aos/1176344552`.

[82] Bradley Efron. *The Jackknife, the Bootstrap and other resampling plans.* 1982.

[83] Daniel J. Eisenstein and Wayne Hu. Baryonic features in the matter transfer function. *Astrophys. J.*, 496:605, 1998. doi: 10.1086/305424.

[84] Daniel J. Eisenstein, Hee-jong Seo, Edwin Sirko, and David Spergel. Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak. *Astrophys. J.*, 664:675–679, 2007. doi: 10.1086/518712.

[85] Daniel J. Eisenstein et al. Spectroscopic target selection for the Sloan Digital Sky Survey: The Luminous red galaxy sample. *Astron. J.*, 122:2267, 2001. doi: 10.1086/323717.

[86] Daniel J. Eisenstein et al. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *Astrophys. J.*, 633:560–574, 2005. doi: 10.1086/466512.

[87] Y. Fan, D. J. Nott, and S. A. Sisson. Approximate bayesian computation via regression density estimation, 2012.

[88] G. Favole, S. A. Rodríguez-Torres, J. Comparat, F. Prada, H. Guo, A. Klypin, and A. D. Montero-Dorta. Galaxy clustering dependence on the [O II] emission line luminosity in the local Universe. MNRAS, 472:550–558, November 2017. doi: 10.1093/mnras/stx1980.

[89] Ginevra Favole. The clustering of Hα emitters in the nearby Universe. *in preparation*, 2019.

[90] Ginevra Favole, Cameron K. McBride, Daniel J. Eisenstein, Francisco Prada, Molly E. Swanson, Chia-Hsun Chuang, and Donald P. Schneider. Building a better understanding of the massive high-redshift BOSS CMASS galaxies as tools for cosmology. MNRAS, 462(2):2218–2236, Oct 2016. doi: 10.1093/mnras/stw1801.

[91] Ginevra Favole, Sergio A. Rodríguez-Torres, Johan Comparat, Francisco Prada, Hong Guo, Anatoly Klypin, and Antonio D. Montero-Dorta. Galaxy clustering dependence on the [O II] emission line luminosity in the local universe. *Monthly Notices of the Royal Astronomical Society*, 472(1):550?558, Aug 2017. ISSN 1365-2966. doi: 10.1093/mnras/stx1980.

[92] Ginevra Favole, Domenico Sapone, and Javier Silva Lafaurie. Cosmological constraints from galaxy multi-tracers in the nearby Universe. *arXiv e-prints*, art. arXiv:1912.06155, December 2019.

[93] Ginevra Favole, Benjamin R. Granett, Javier Silva Lafaurie, and Domenico Sapone. Does jackknife scale really matter for accurate large-scale structure covariances?, 2020.

[94] Eric D. Feigelson and G. Jogesh Babu. *Modern Statistical Methods for Astronomy: With R Applications*. Cambridge University Press, 2012. doi: 10.1017/CBO9781139015653.

[95] Martin Feix, Adi Nusser, and Enzo Branchini. Growth Rate of Cosmological Perturbations at z 0.1 from a New Observational Test. *Phys. Rev. Lett.*, 115(1):011301, 2015. doi: 10.1103/PhysRevLett.115.011301.

[96] Hume A. Feldman, Nick Kaiser, and John A. Peacock. Power-Spectrum Analysis of Three-dimensional Redshift Surveys. ApJ, 426:23, May 1994. doi: 10.1086/174036.

[97] Yu Feng, Man-Yat Chu, Uroš Seljak, and Patrick McDonald. FASTPM: a new scheme for fast simulations of dark matter and haloes. MNRAS, 463(3):2273–2286, December 2016. doi: 10.1093/mnras/stw2123.

[98] F. Feroz, M. P. Hobson, and M. Bridges. MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Mon. Not. Roy. Astron. Soc.*, 398: 1601–1614, 2009. doi: 10.1111/j.1365-2966.2009.14548.x.

[99] A. S. Font, K. V. Johnston, A. M. N. Ferguson, J. S. Bullock, B. E. Robertson, J. Tumlinson, and P. Guhathakurta. The Stellar Content of Galaxy Halos: A Comparison between ΛCDM Models and Observations of M31. ApJ, 673:215-225, January 2008. doi: 10.1086/524102.

[100] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. emcee: The MCMC Hammer. PASP, 125(925):306, March 2013. doi: 10.1086/670067.

[101] M. Fukugita, T. Ichikawa, J. E. Gunn, M. Doi, K. Shimasaku, and D. P. Schneider. The Sloan Digital Sky Survey Photometric System. AJ, 111:1748, Apr 1996. doi: 10.1086/117915.

[102] I. D. Gargiulo, S. A. Cora, N. D. Padilla, A. M. Muñoz Arancibia, A. N. Ruiz, A. A. Orsi, T. E. Tecce, C. Weidner, and G. Bruzual. Chemoarchaeological downsizing in a hierarchical universe: impact of a top-heavy IGIMF. MNRAS, 446:3820–3841, February 2015. doi: 10.1093/mnras/stu2272.

[103] Shy Genel, Mark Vogelsberger, Volker Springel, Debora Sijacki, Dylan Nelson, Greg

Snyder, Vicente Rodriguez-Gomez, Paul Torrey, and Lars Hernquist. Introducing the Illustris project: the evolution of galaxy populations across cosmic time. MNRAS, 445 (1):175–200, November 2014. doi: 10.1093/mnras/stu1654.

[104] Mathieu Germain, Karol Gregor, Iain Murray, and Hugo Larochelle. Made: Masked autoencoder for distribution estimation. In Francis Bach and David Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 881–889, Lille, France, 07–09 Jul 2015. PMLR. URL `http://proceedings.mlr.press/v37/germain15.html`.

[105] Héctor Gil-Marín et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: BAO measurement from the LOS-dependent power spectrum of DR12 BOSS galaxies. *Mon. Not. Roy. Astron. Soc.*, 460(4):4210–4219, 2016. doi: 10.1093/mnras/stw1264.

[106] V. Gonzalez-Perez, C. G. Lacey, C. M. Baugh, C. D. P. Lagos, J. Helly, D. J. R. Campbell, and P. D. Mitchell. How sensitive are predicted galaxy luminosities to the choice of stellar population synthesis model? MNRAS, 439:264–283, March 2014. doi: 10.1093/mnras/stt2410.

[107] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wand elt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. ApJ, 622(2):759–771, April 2005. doi: 10.1086/427976.

[108] Benjamin R. Granett, Ginevra Favole, Antonio D. Montero-Dorta, Enzo Branchini, Luigi Guzzo, and Sylvain de la Torre. Measuring the growth of structure by matching dark matter haloes to galaxies with VIPERS and SDSS. MNRAS, 489(1):653–662, October 2019. doi: 10.1093/mnras/stz2152.

[109] J. Green et al. Wide-Field InfraRed Survey Telescope (WFIRST) Final Report. 2012.

[110] J. E. Gunn, M. Carr, C. Rockosi, M. Sekiguchi, K. Berry, B. Elms, E. de Haas, Ž. Ivezić, G. Knapp, R. Lupton, G. Pauls, R. Simcoe, R. Hirsch, D. Sanford, S. Wang, D. York, F. Harris, J. Annis, L. Bartozek, W. Boroski, J. Bakken, M. Haldeman, S. Kent, S. Holm, D. Holmgren, D. Petravick, A. Prosapio, R. Rechenmacher, M. Doi, M. Fukugita, K. Shimasaku, N. Okada, C. Hull, W. Siegmund, E. Mannery, M. Blouke, D. Heidtman, D. Schneider, R. Lucinio, and J. Brinkman. The Sloan Digital Sky Survey Photometric Camera. AJ, 116:3040–3081, December 1998. doi: 10.1086/300645.

[111] J. E. Gunn, W. A. Siegmund, E. J. Mannery, R. E. Owen, C. L. Hull, R. F. Leger, L. N. Carey, G. R. Knapp, D. G. York, W. N. Boroski, S. M. Kent, R. H. Lupton, C. M. Rockosi, M. L. Evans, P. Waddell, J. E. Anderson, J. Annis, J. C. Barentine, L. M. Bartoszek, S. Bastian, S. B. Bracker, H. J. Brewington, C. I. Briegel, J. Brinkmann, Y. J. Brown, M. A. Carr, P. C. Czarapata, C. C. Drennan, T. Dombeck, G. R. Federwitz, B. A. Gillespie, C. Gonzales, S. U. Hansen, M. Harvanek, J. Hayes, W. Jordan, E. Kinney, M. Klaene, S. J. Kleinman, R. G. Kron, J. Kresinski, G. Lee, S. Limmongkol, C. W. Lindenmeyer, D. C. Long, C. L. Loomis, P. M. McGehee, P. M. Mantsch, E. H.

Neilsen, Jr., R. M. Neswold, P. R. Newman, A. Nitta, J. Peoples, Jr., J. R. Pier, P. S. Prieto, A. Prosapio, C. Rivetta, D. P. Schneider, S. Snedden, and S.-i. Wang. The 2.5 m Telescope of the Sloan Digital Sky Survey. AJ, 131:2332–2359, April 2006. doi: 10.1086/500975.

[112] H. Guo, I. Zehavi, Z. Zheng, D. H. Weinberg, A. A. Berlind, M. Blanton, Y. Chen, D. J. Eisenstein, S. Ho, E. Kazin, M. Manera, C. Maraston, C. K. McBride, S. E. Nuza, N. Padmanabhan, J. K. Parejko, W. J. Percival, A. J. Ross, N. P. Ross, L. Samushia, A. G. Sanchez, D. J. Schlegel, D. P. Schneider, R. A. Skibba, M. E. C. Swanson, J. L. Tinker, R. Tojeiro, D. A. Wake, M. White, N. A. Bahcall, D. Bizyaev, H. Brewington, K. Bundy, L. N. A. da Costa, G. Ebelke, V. Malanushenko, E. Malanushenko, D. Oravetz, G. Rossi, A. Simmons, S. Snedden, A. Streblyanska, and D. Thomas. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Luminosity and Color Dependence and Redshift Evolution. *ArXiv e-prints: 1212.1211*, December 2012.

[113] H. Guo, Z. Zheng, Y. P. Jing, I. Zehavi, C. Li, D. H. Weinberg, R. A. Skibba, R. C. Nichol, G. Rossi, C. G. Sabiu, D. P. Schneider, and C. K. McBride. Modelling the redshift-space three-point correlation function in SDSS-III. MNRAS, 449:L95–L99, April 2015. doi: 10.1093/mnrasl/slv020.

[114] H. Guo, Z. Zheng, I. Zehavi, P. S. Behroozi, C.-H. Chuang, J. Comparat, G. Favole, S. Gottloeber, A. Klypin, F. Prada, D. H. Weinberg, and G. Yepes. Redshift-space clustering of SDSS galaxies - luminosity dependence, halo occupation distribution, and velocity bias. MNRAS, 453:4368–4383, November 2015. doi: 10.1093/mnras/stv1966.

[115] Hong Guo et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Luminosity and Color Dependence and Redshift Evolution. *Astrophys. J.*, 767:122, 2013. doi: 10.1088/0004-637X/767/2/122.

[116] Q. Guo, S. White, M. Boylan-Kolchin, G. De Lucia, G. Kauffmann, G. Lemson, C. Li, V. Springel, and S. Weinmann. From dwarf spheroidals to cD galaxies: simulating the galaxy population in a ΛCDM cosmology. MNRAS, 413:101–131, May 2011. doi: 10.1111/j.1365-2966.2010.18114.x.

[117] A. J. S. Hamilton. Toward Better Ways to Measure the Galaxy Correlation Function. ApJ, 417:19, November 1993. doi: 10.1086/173288.

[118] A. J. S. Hamilton. Linear redshift distortions: A review. *The Evolving Universe*, page 185–275, 1998. ISSN 2214-7985. doi: 10.1007/978-94-011-4960-0_17. URL `http://dx.doi.org/10.1007/978-94-011-4960-0_17`.

[119] 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.*, 2006. doi: 10.1051/0004-6361:20066170. [Astron. Astrophys.464,399(2007)].

[120] B. Henriques, C. Maraston, P. Monaco, F. Fontanot, N. Menci, G. De Lucia, and

C. Tonini. The effect of thermally pulsating asymptotic giant branch stars on the evolution of the rest-frame near-infrared galaxy luminosity function. MNRAS, 415: 3571–3579, August 2011. doi: 10.1111/j.1365-2966.2011.18972.x.

[121] B. M. B. Henriques, S. D. M. White, P. A. Thomas, R. E. Angulo, Q. Guo, G. Lemson, and V. Springel. Simulations of the galaxy population constrained by observations from z = 3 to the present day: implications for galactic winds and the fate of their ejecta. MNRAS, 431:3373–3395, June 2013. doi: 10.1093/mnras/stt415.

[122] Samuel R Hinton, Cullan Howlett, and Tamara M Davis. barry and the bao model comparison. *Monthly Notices of the Royal Astronomical Society*, 493(3):4078–4093, Feb 2020. ISSN 1365-2966. doi: 10.1093/mnras/staa361. URL `http://dx.doi.org/10.1093/mnras/staa361`.

[123] M. Hirschmann, G. De Lucia, and F. Fontanot. Galaxy assembly, stellar feedback and metal enrichment: the view from the GAEA model. MNRAS, 461:1760–1785, September 2016. doi: 10.1093/mnras/stw1318.

[124] Tao Hong, J. L. Han, and Z. L. Wen. A Detection of Baryon Acoustic Oscillations from the Distribution of Galaxy Clusters. ApJ, 826(2):154, August 2016. doi: 10.3847/0004-637X/826/2/154.

[125] J. Hou, C. G. Lacey, and C. S. Frenk. A new gas cooling model for semi-analytical galaxy formation models. *ArXiv e-prints*, August 2017.

[126] C. Howlett, M. Manera, and W. J. Percival. L-PICOLA: A parallel code for fast dark matter simulation. *Astronomy and Computing*, 12:109–126, September 2015. doi: 10.1016/j.ascom.2015.07.003.

[127] Cullan Howlett, Ashley Ross, Lado Samushia, Will Percival, and Marc Manera. The clustering of the SDSS main galaxy sample. Mock galaxy catalogues and a measurement of the growth of structure from redshift space distortions at $z = 0.15$. *Mon. Not. Roy. Astron. Soc.*, 449(1):848–866, 2015. doi: 10.1093/mnras/stu2693.

[128] Wayne Hu and Naoshi Sugiyama. Small scale cosmological perturbations: An Analytic approach. *Astrophys. J.*, 471:542–570, 1996. doi: 10.1086/177989.

[129] 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. ISSN 0027-8424. doi: 10.1073/pnas.15.3.168. URL `https://www.pnas.org/content/15/3/168`.

[130] Michael J. Hudson and Stephen J. Turnbull. The growth rate of cosmic structure from peculiar velocities at low and high redshifts. *Astrophys. J.*, 751:L30, 2013. doi: 10.1088/2041-8205/751/2/L30.

[131] Dragan Huterer, Daniel Shafer, Daniel Scolnic, and Fabian Schmidt. Testing LCDM at the lowest redshifts with SN Ia and galaxy velocities. 2016.

[132] Ž. Ivezić, S. M. Kahn, J. A. Tyson, B. Abel, E. Acosta, R. Allsman, D. Alonso, Y. AlSayyad, S. F. Anderson, J. Andrew, and et al. LSST: From Science Drivers to Reference Design and Anticipated Data Products. ApJ, 873:111, March 2019. doi: 10.3847/1538-4357/ab042c.

[133] Zeljko Ivezic, Andrew J. Connolly, Jacob T. VanderPlas, and Alexander Gray. *Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data.* Princeton University Press, USA, 2014. ISBN 0691151687.

[134] Albert Izard, Martin Crocce, and Pablo Fosalba. ICE-COLA: towards fast and accurate synthetic galaxy catalogues optimizing a quasi-N-body method. MNRAS, 459(3):2327–2341, July 2016. doi: 10.1093/mnras/stw797.

[135] M. Jarvis, G. Bernstein, and B. Jain. The skewness of the aperture mass statistic. MNRAS, 352(1):338–352, July 2004. doi: 10.1111/j.1365-2966.2004.07926.x.

[136] Nick Kaiser. Clustering in real space and in redshift space. *Monthly Notices of the Royal Astronomical Society*, 227(1):1–21, 07 1987. ISSN 0035-8711. doi: 10.1093/mnras/227.1.1. URL `https://doi.org/10.1093/mnras/227.1.1`.

[137] G. Kauffmann, S. D. M. White, and B. Guiderdoni. The Formation and Evolution of Galaxies Within Merging Dark Matter Haloes. MNRAS, 264:201, September 1993. doi: 10.1093/mnras/264.1.201.

[138] Cari G. Kaufman, Mark J. Schervish, and Douglas W. Nychka. Covariance tapering for likelihood-based estimation in large spatial data sets. *Journal of the American Statistical Association*, 103(484):1545–1555, 2008. doi: 10.1198/016214508000000959.

[139] Martin Kerscher. The geometry of second-order statistics - biases in common estimators, 1998.

[140] Francisco-Shu Kitaura, Sergio Rodríguez-Torres, Chia-Hsun Chuang, Cheng Zhao, Francisco Prada, Héctor Gil-Marín, Hong Guo, Gustavo Yepes, Anatoly Klypin, Claudia G. Scóccola, Jeremy Tinker, Cameron McBride, Beth Reid, Ariel G. Sánchez, Salvador Salazar-Albornoz, Jan Niklas Grieb, Mariana Vargas-Magana, Antonio J. Cuesta, Mark Neyrinck, Florian Beutler, Johan Comparat, Will J. Percival, and Ashley Ross. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: mock galaxy catalogues for the BOSS Final Data Release. MNRAS, 456(4):4156–4173, March 2016. doi: 10.1093/mnras/stv2826.

[141] A. Klypin and F. Prada. Dark matter statistics for large galaxy catalogues: power spectra and covariance matrices. MNRAS, 478:4602–4621, August 2018. doi: 10.1093/mnras/sty1340.

[142] Jun Koda, Chris Blake, Florian Beutler, Eyal Kazin, and Felipe Marin. Fast and accurate mock catalogue generation for low-mass galaxies. MNRAS, 459(2):2118–2129, June 2016. doi: 10.1093/mnras/stw763.

[143] Martin Kunz and Domenico Sapone. Crossing the Phantom Divide. *Phys. Rev.*, D74:

123503, 2006. doi: 10.1103/PhysRevD.74.123503.

[144] Martin Kunz and Domenico Sapone. Dark Energy versus Modified Gravity. *Phys. Rev. Lett.*, 98:121301, 2007. doi: 10.1103/PhysRevLett.98.121301.

[145] C. D. P. Lagos, C. M. Baugh, C. G. Lacey, A. J. Benson, H.-S. Kim, and C. Power. Cosmic evolution of the atomic and molecular gas contents of galaxies. MNRAS, 418: 1649–1667, December 2011. doi: 10.1111/j.1365-2966.2011.19583.x.

[146] C. d. P. Lagos, C. G. Lacey, and C. M. Baugh. A dynamical model of supernova feedback: gas outflows from the interstellar medium. MNRAS, 436:1787–1817, December 2013. doi: 10.1093/mnras/stt1696.

[147] Stephen D. Landy and Alexander S. Szalay. Bias and Variance of Angular Correlation Functions. ApJ, 412:64, July 1993. doi: 10.1086/172900.

[148] R. Laureijs, J. Amiaux, S. Arduini, J. L. Auguères, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet, B. Garilli, P. Gondoin, L. Guzzo, J. Hoar, H. Hoekstra, R. Holmes, T. Kitching, T. Maciaszek, Y. Mellier, F. Pasian, W. Percival, J. Rhodes, G. Saavedra Criado, M. Sauvage, R. Scaramella, L. Valenziano, S. Warren, R. Bender, F. Castander, A. Cimatti, O. Le Fèvre, H. Kurki-Suonio, M. Levi, P. Lilje, G. Meylan, R. Nichol, K. Pedersen, V. Popa, R. Rebolo Lopez, H. W. Rix, H. Rottgering, W. Zeilinger, F. Grupp, P. Hudelot, R. Massey, M. Meneghetti, L. Miller, S. Paltani, S. Paulin-Henriksson, S. Pires, C. Saxton, T. Schrabback, G. Seidel, J. Walsh, N. Aghanim, L. Amendola, J. Bartlett, C. Baccigalupi, J. P. Beaulieu, K. Benabed, J. G. Cuby, D. Elbaz, P. Fosalba, G. Gavazzi, A. Helmi, I. Hook, M. Irwin, J. P. Kneib, M. Kunz, F. Mannucci, L. Moscardini, C. Tao, R. Teyssier, J. Weller, G. Zamorani, M. R. Zapatero Osorio, O. Boulade, J. J. Foumond, A. Di Giorgio, P. Guttridge, A. James, M. Kemp, J. Martignac, A. Spencer, D. Walton, T. Blümchen, C. Bonoli, F. Bortoletto, C. Cerna, L. Corcione, C. Fabron, K. Jahnke, S. Ligori, F. Madrid, L. Martin, G. Morgante, T. Pamplona, E. Prieto, M. Riva, R. Toledo, M. Trifoglio, F. Zerbi, F. Abdalla, M. Douspis, C. Grenet, S. Borgani, R. Bouwens, F. Courbin, J. M. Delouis, P. Dubath, A. Fontana, M. Frailis, A. Grazian, J. Koppenhöfer, O. Mansutti, M. Melchior, M. Mignoli, J. Mohr, C. Neissner, K. Noddle, M. Poncet, M. Scodeggio, S. Serrano, N. Shane, J. L. Starck, C. Surace, A. Taylor, G. Verdoes-Kleijn, C. Vuerli, O. R. Williams, A. Zacchei, B. Altieri, I. Escudero Sanz, R. Kohley, T. Oosterbroek, P. Astier, D. Bacon, S. Bardelli, C. Baugh, F. Bellagamba, C. Benoist, D. Bianchi, A. Biviano, E. Branchini, C. Carbone, V. Cardone, D. Clements, S. Colombi, C. Conselice, G. Cresci, N. Deacon, J. Dunlop, C. Fedeli, F. Fontanot, P. Franzetti, C. Giocoli, J. Garcia-Bellido, J. Gow, A. Heavens, P. Hewett, C. Heymans, A. Holland, Z. Huang, O. Ilbert, B. Joachimi, E. Jennins, E. Kerins, A. Kiessling, D. Kirk, R. Kotak, O. Krause, O. Lahav, F. van Leeuwen, J. Lesgourgues, M. Lombardi, M. Magliocchetti, K. Maguire, E. Majerotto, R. Maoli, F. Marulli, S. Maurogordato, H. McCracken, R. McLure, A. Melchiorri, A. Merson, M. Moresco, M. Nonino, P. Norberg, J. Peacock, R. Pello, M. Penny, V. Pettorino, C. Di Porto, L. Pozzetti, C. Quercellini, M. Radovich, A. Rassat, N. Roche, S. Ronayette, E. Rossetti, B. Sartoris, P. Schneider, E. Semboloni, S. Serjeant, F. Simpson, C. Skordis, G. Smadja, S. Smartt, P. Spano, S. Spiro, M. Sullivan, A. Tilquin, R. Trotta, L. Verde,

Y. Wang, G. Williger, G. Zhao, J. Zoubian, and E. Zucca. Euclid Definition Study Report. *arXiv e-prints*, art. arXiv:1110.3193, October 2011.

[149] R. Laureijs et al. Euclid Definition Study Report. 2011.

[150] G. Lemaître. Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Annales de la Société Scientifique de Bruxelles*, 47:49–59, January 1927.

[151] Julien Lesgourgues. The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview. 2011.

[152] Andrew R. Liddle. How many cosmological parameters? *Mon. Not. Roy. Astron. Soc.*, 351:L49–L53, 2004. doi: 10.1111/j.1365-2966.2004.08033.x.

[153] Andrew R. Liddle and David H. Lyth. The cold dark matter density perturbation. *Physics Reports*, 231(1-2):1–105, Aug 1993. ISSN 0370-1573. doi: 10.1016/0370-1573(93)90114-s. URL `http://dx.doi.org/10.1016/0370-1573(93)90114-S`.

[154] Eric V. Linder. Exploring the expansion history of the universe. *Phys. Rev. Lett.*, 90: 091301, 2003. doi: 10.1103/PhysRevLett.90.091301.

[155] Eric V. Linder. Cosmic Growth and Expansion Conjoined. *Astropart. Phys.*, 86:41–45, 2017. doi: 10.1016/j.astropartphys.2016.11.002.

[156] Jarno Lintusaari, Michael U. Gutmann, Ritabrata Dutta, Samuel Kaski, and Jukka Corander. Fundamentals and Recent Developments in Approximate Bayesian Computation. *Systematic Biology*, 66(1):e66–e82, 09 2016. ISSN 1063-5157. doi: 10.1093/sysbio/syw077. URL `https://doi.org/10.1093/sysbio/syw077`.

[157] Martha Lippich, Ariel G. Sánchez, Manuel Colavincenzo, Emiliano Sefusatti, Pierluigi Monaco, Linda Blot, Martin Crocce, Marcelo A. Alvarez, Aniket Agrawal, Santiago Avila, Andrés Balaguera-Antolínez, Richard Bond, Sandrine Codis, Claudio Dalla Vecchia, Antonio Dorta, Pablo Fosalba, Albert Izard, Francisco-Shu Kitaura, Marcos Pellejero-Ibanez, George Stein, Mohammadjavad Vakili, and Gustavo Yepes. Comparing approximate methods for mock catalogues and covariance matrices - I. Correlation function. MNRAS, 482(2):1786–1806, January 2019. doi: 10.1093/mnras/sty2757.

[158] Jan-Matthis Lueckmann, Pedro J Goncalves, Giacomo Bassetto, Kaan Öcal, Marcel Nonnenmacher, and Jakob H Macke. Flexible statistical inference for mechanistic models of neural dynamics. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 1289–1299. Curran Associates, Inc., 2017.

[159] Jan-Matthis Lueckmann, Giacomo Bassetto, Theofanis Karaletsos, and Jakob H. Macke. Likelihood-free inference with emulator networks, 2018.

[160] Chung-Pei Ma and Edmund Bertschinger. Cosmological perturbation theory in the synchronous and conformal newtonian gauges. *The Astrophysical Journal*, 455:7, Dec

1995. ISSN 1538-4357. doi: 10.1086/176550. URL `http://dx.doi.org/10.1086/176550`.

[161] E. Macaulay, I. K. Wehus, and H. K. Eriksen. Lower growth rate from recent redshift space distortion measurements than expected from planck. *Physical Review Letters*, 111(16), Oct 2013. ISSN 1079-7114. doi: 10.1103/physrevlett.111.161301. URL `http://dx.doi.org/10.1103/PhysRevLett.111.161301`.

[162] Marc Manera, Roman Scoccimarro, Will J. Percival, Lado Samushia, Cameron K. McBride, Ashley J. Ross, Ravi K. Sheth, Martin White, Beth A. Reid, Ariel G. Sánchez, Roland de Putter, Xiaoying Xu, Andreas A. Berlind, Jonathan Brinkmann, Claudia Maraston, Bob Nichol, Francesco Montesano, Nikhil Padmanabhan, Ramin A. Skibba, Rita Tojeiro, and Benjamin A. Weaver. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: a large sample of mock galaxy catalogues. MNRAS, 428(2):1036–1054, January 2013. doi: 10.1093/mnras/sts084.

[163] Valerio Marra and Domenico Sapone. Null tests of the standard model using the linear model formalism. *Phys. Rev.*, D97(8):083510, 2018. doi: 10.1103/PhysRevD.97.083510.

[164] Vicent Martínez, Enrique Martinez-Gonzalez, Enn Saar, and M. Pons-Borderia. *Data Analysis in Cosmology*, volume 665. 01 2009. doi: 10.1007/978-3-540-44767-2.

[165] Takahiko Matsubara. The correlation function in redshift space: General formula with wide-angle effects and cosmological distortions. *The Astrophysical Journal*, 535(1): 1–23, may 2000. doi: 10.1086/308827. URL `https://doi.org/10.1086%2F308827`.

[166] Alexander Merson, Yun Wang, Andrew Benson, Andreas Faisst, Daniel Masters, Alina Kiessling, and Jason Rhodes. Predicting Hα emission-line galaxy counts for future galaxy redshift surveys. *Mon. Not. Roy. Astron. Soc.*, 474(1):177–196, 2018. doi: 10.1093/mnras/stx2649.

[167] Lisa Dale Miller, Melvin E. Miller, and Melissa Sivvy. The jackknife-a review. *Biometrika*, pages 1–15, 1974.

[168] R. G. Miller. The jackknife-a review. *Biometrika*, 61:1–15, April 1974. doi: 10.1093/biomet/61.1.1.

[169] H. Mo, F. C. van den Bosch, and S. White. *Galaxy Formation and Evolution.* May 2010.

[170] P. Monaco, E. Sefusatti, S. Borgani, M. Crocce, P. Fosalba, R. K. Sheth, and T. Theuns. An accurate tool for the fast generation of dark matter halo catalogues. MNRAS, 433 (3):2389–2402, August 2013. doi: 10.1093/mnras/stt907.

[171] P. Monaco, A. J. Benson, G. De Lucia, F. Fontanot, S. Borgani, and M. Boylan-Kolchin. A semi-analytic model comparison: testing cooling models against hydrodynamical simulations. MNRAS, 441:2058–2077, July 2014. doi: 10.1093/mnras/stu655.

[172] Pierluigi Monaco, Tom Theuns, and Giuliano Taffoni. The pinocchio algorithm: pin-

pointing orbit-crossing collapsed hierarchical objects in a linear density field. MNRAS, 331(3):587–608, April 2002. doi: 10.1046/j.1365-8711.2002.05162.x.

[173] Antonio D. Montero-Dorta, L. Raul Abramo, Benjamin R. Granett, Sylvain de la Torre, and Luigi Guzzo. The Multi-Tracer Optimal Estimator applied to VIPERS. 2019.

[174] M. Moresco et al. Improved constraints on the expansion rate of the Universe up to z 1.1 from the spectroscopic evolution of cosmic chronometers. *JCAP*, 1208:006, 2012. doi: 10.1088/1475-7516/2012/08/006.

[175] Michele Moresco. Raising the bar: new constraints on the Hubble parameter with cosmic chronometers at z = 2. *Mon. Not. Roy. Astron. Soc.*, 450(1):L16–L20, 2015. doi: 10.1093/mnrasl/slv037.

[176] Michele Moresco and Federico Marulli. Cosmological constraints from a joint analysis of cosmic growth and expansion. *Mon. Not. Roy. Astron. Soc.*, 471(1):L82–L86, 2017. doi: 10.1093/mnrasl/slx112.

[177] Michele Moresco, Lucia Pozzetti, Andrea Cimatti, Raul Jimenez, Claudia Maraston, Licia Verde, Daniel Thomas, Annalisa Citro, Rita Tojeiro, and David Wilkinson. A 6% measurement of the Hubble parameter at $z \sim 0.45$: direct evidence of the epoch of cosmic re-acceleration. *JCAP*, 1605(05):014, 2016. doi: 10.1088/1475-7516/2016/05/014.

[178] Thorsten Naab and Jeremiah P. Ostriker. Theoretical challenges in galaxy formation. *Annual Review of Astronomy and Astrophysics*, 55(1):59–109, 2017. doi: 10.1146/annurev-astro-081913-040019. URL `https://doi.org/10.1146/annurev-astro-081913-040019`.

[179] Savvas Nesseris and Juan Garcia-Bellido. Is the Jeffreys' scale a reliable tool for Bayesian model comparison in cosmology? *JCAP*, 1308:036, 2013. doi: 10.1088/1475-7516/2013/08/036.

[180] Savvas Nesseris and Domenico Sapone. Accuracy of the growth index in the presence of dark energy perturbations. *Phys. Rev.*, D92(2):023013, 2015. doi: 10.1103/PhysRevD.92.023013.

[181] Savvas Nesseris, George Pantazis, and Leandros Perivolaropoulos. Tension and constraints on modified gravity parametrizations of $G_{\mathrm{eff}}(z)$ from growth rate and Planck data. *Phys. Rev.*, D96(2):023542, 2017. doi: 10.1103/PhysRevD.96.023542.

[182] P. Norberg, E. Gaztañaga, C. M. Baugh, and D. J. Croton. Statistical analysis of galaxy surveys - IV. An objective way to quantify the impact of superstructures on galaxy clustering statistics. MNRAS, 418:2435–2450, December 2011. doi: 10.1111/j.1365-2966.2011.19636.x.

[183] Peder Norberg, Carlton M. Baugh, Enrique Gaztanaga, and Darren J. Croton. Statistical Analysis of Galaxy Surveys - I. Robust error estimation for 2-point clustering statistics. *Mon. Not. Roy. Astron. Soc.*, 396:19, 2009. doi: 10.1111/j.1365-2966.2009.14389.x.

[184] Teppei Okumura et al. The Subaru FMOS galaxy redshift survey (FastSound). IV. New constraint on gravity theory from redshift space distortions at $z \sim 1.4$. *Publ. Astron. Soc. Jap.*, 68(3, id. 38):24, 2016. doi: 10.1093/pasj/psw029.

[185] Nikhil Padmanabhan and Martin J. White. Constraining Anisotropic Baryon Oscillations. *Phys. Rev.*, D77:123540, 2008. doi: 10.1103/PhysRevD.77.123540.

[186] Nikhil Padmanabhan, David J. Schlegel, Douglas P. Finkbeiner, J. C. Barentine, Michael R. Blanton, Howard J. Brewington, James E. Gunn, Michael Harvanek, David W. Hogg, Željko Ivezić, and et al. An improved photometric calibration of the sloan digital sky survey imaging data. *The Astrophysical Journal*, 674(2):1217–1233, Feb 2008. ISSN 1538-4357. doi: 10.1086/524677. URL `http://dx.doi.org/10.1086/524677`.

[187] Nikhil Padmanabhan, Xiaoying Xu, Daniel J. Eisenstein, Richard Scalzo, Antonio J. Cuesta, Kushal T. Mehta, and Eyal Kazin. A 2 per cent distance to z=0.35 by reconstructing baryon acoustic oscillations - I. Methods and application to the Sloan Digital Sky Survey. *Mon. Not. Roy. Astron. Soc.*, 427(3):2132–2145, 2012. doi: 10.1111/j.1365-2966.2012.21888.x.

[188] George Papamakarios and Iain Murray. Fast ϵ-free inference of simulation models with bayesian conditional density estimation. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems 29*, pages 1028–1036. Curran Associates, Inc., 2016.

[189] George Papamakarios, Theo Pavlakou, and Iain Murray. Masked autoregressive flow for density estimation. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 2338–2347. Curran Associates, Inc., 2017. URL `http://papers.nips.cc/paper/6828-masked-autoregressive-flow-for-density-estimation.pdf`.

[190] George Papamakarios, David C. Sterratt, and Iain Murray. Sequential neural likelihood: Fast likelihood-free inference with autoregressive flows, 2018.

[191] Gabriele Parimbelli, Matteo Viel, and Emiliano Sefusatti. On the degeneracy between baryon feedback and massive neutrinos as probed by matter clustering and weak lensing. J. Cosmology Astropart. Phys., 2019(1):010, January 2019. doi: 10.1088/1475-7516/2019/01/010.

[192] Dante J. Paz and Ariel G. Sánchez. Improving the precision matrix for precision cosmology. MNRAS, 454(4):4326–4334, December 2015. doi: 10.1093/mnras/stv2259.

[193] David W Pearson and Lado Samushia. A Detection of the Baryon Acoustic Oscillation features in the SDSS BOSS DR12 Galaxy Bispectrum. *Monthly Notices of the Royal Astronomical Society*, 478(4):4500–4512, 05 2018. ISSN 0035-8711. doi: 10.1093/mnras/sty1266. URL `https://doi.org/10.1093/mnras/sty1266`.

[194] David W. Pearson, Lado Samushia, and Praful Gagrani. Optimal weights for measuring

redshift space distortions in multitracer galaxy catalogues. MNRAS, 463(3):2708–2715, December 2016. doi: 10.1093/mnras/stw2177.

[195] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[196] P. J. E. Peebles and M. G. Hauser. Statistical Analysis of Catalogs of Extragalactic Objects. III. The Shane-Wirtanen and Zwicky Catalogs. ApJS, 28:19, November 1974. doi: 10.1086/190308.

[197] S. Perlmutter et al. Measurements of Omega and Lambda from 42 high redshift supernovae. *Astrophys. J.*, 517:565–586, 1999. doi: 10.1086/307221.

[198] A. Pezzotta et al. The VIMOS Public Extragalactic Redshift Survey (VIPERS): The growth of structures at $0.5 < z < 1.2$ from redshift-space distortions in the clustering of the PDR-2 final sample. 2016.

[199] Oliver H. E. Philcox, Daniel J. Eisenstein, Ross O'Connell, and Alexander Wiegand. RASCALC: a jackknife approach to estimating single- and multitracer galaxy covariance matrices. MNRAS, 491(3):3290–3317, January 2020. doi: 10.1093/mnras/stz3218.

[200] Oliver Piattella. Lecture notes in cosmology. *UNITEXT for Physics*, 2018. ISSN 2198-7890. doi: 10.1007/978-3-319-95570-4. URL `http://dx.doi.org/10.1007/978-3-319-95570-4`.

[201] Annalisa Pillepich, Volker Springel, Dylan Nelson, Shy Genel, Jill Naiman, Rüdiger Pakmor, Lars Hernquist, Paul Torrey, Mark Vogelsberger, Rainer Weinberger, and Federico Marinacci. Simulating galaxy formation with the IllustrisTNG model. MNRAS, 473(3):4077–4106, January 2018. doi: 10.1093/mnras/stx2656.

[202] Planck, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XVI. Cosmological parameters. A&A, 571:A16, November 2014. doi: 10.1051/0004-6361/201321591.

[203] Maria-Jesus Pons-Borderia, Vicent J. Martinez, Dietrich Stoyan, Helga Stoyan, and Enn Saar. Comparing estimators of the galaxy correlation function. *The Astrophysical Journal*, 523(2):480–491, Oct 1999. ISSN 1538-4357. doi: 10.1086/307754. URL `http://dx.doi.org/10.1086/307754`.

[204] Judit Pérez-Romero and Savvas Nesseris. Cosmological constraints and comparison of viable $f(R)$ models. *Phys. Rev.*, D97(2):023525, 2018. doi: 10.1103/PhysRevD.97.023525.

[205] M. H. Quenouille. Notes on bias in estimation. *Biometrika*, 43:353 – 360, 1956.

[206] 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. doi: 10.1093/mnras/stx301.

[207] Beth Reid, Shirley Ho, Nikhil Padmanabhan, Will J. Percival, Jeremy Tinker, Rita Tojeiro, Martin White, Daniel J. Eisenstein, Claudia Maraston, Ashley J. Ross, Ariel G. Sánchez, David Schlegel, Erin Sheldon, Michael A. Strauss, Daniel Thomas, David Wake, Florian Beutler, Dmitry Bizyaev, Adam S. Bolton, Joel R. Brownstein, Chia-Hsun Chuang, Kyle Dawson, Paul Harding, Francisco-Shu Kitaura, Alexie Leauthaud, Karen Masters, Cameron K. McBride, Surhud More, Matthew D. Olmstead, Daniel Oravetz, Sebastián E. Nuza, Kaike Pan, John Parejko, Janine Pforr, Francisco Prada, Sergio Rodríguez-Torres, Salvador Salazar-Albornoz, Lado Samushia, Donald P. Schneider, Claudia G. Scóccola, Audrey Simmons, and Mariana Vargas-Magana. SDSS-III Baryon Oscillation Spectroscopic Survey Data Release 12: galaxy target selection and large-scale structure catalogues. MNRAS, 455(2):1553–1573, Jan 2016. doi: 10.1093/mnras/stv2382.

[208] Adam G. Riess et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.*, 116:1009–1038, 1998. doi: 10.1086/300499.

[209] A. J. Ross, W. J. Percival, A. G. Sánchez, L. Samushia, S. Ho, E. Kazin, M. Manera, B. Reid, M. White, R. Tojeiro, C. K. McBride, X. Xu, D. A. Wake, M. A. Strauss, F. Montesano, M. E. C. Swanson, S. Bailey, A. S. Bolton, A. M. Dorta, D. J. Eisenstein, H. Guo, J.-C. Hamilton, R. C. Nichol, N. Padmanabhan, F. Prada, D. J. Schlegel, M. V. Magaña, I. Zehavi, M. Blanton, D. Bizyaev, H. Brewington, A. J. Cuesta, E. Malanushenko, V. Malanushenko, D. Oravetz, J. Parejko, K. Pan, D. P. Schneider, A. Shelden, A. Simmons, S. Snedden, and G.-b. Zhao. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: analysis of potential systematics. MNRAS, 424:564–590, July 2012. doi: 10.1111/j.1365-2966.2012.21235.x.

[210] Ashley J. Ross, Lado Samushia, Cullan Howlett, Will J. Percival, Angela Burden, and Marc Manera. The clustering of the SDSS DR7 main Galaxy sample - I. A 4 per cent distance measure at $z = 0.15$. *Mon. Not. Roy. Astron. Soc.*, 449(1):835–847, 2015. doi: 10.1093/mnras/stv154.

[211] Ashley J. Ross, Florian Beutler, Chia-Hsun Chuang, Marcos Pellejero-Ibanez, Hee-Jong Seo, Mariana Vargas-Magaña, Antonio J. Cuesta, Will J. Percival, Angela Burden, Ariel G. Sánchez, Jan Niklas Grieb, Beth Reid, Joel R. Brownstein, Kyle S. Dawson, Daniel J. Eisenstein, Shirley Ho, Francisco-Shu Kitaura, Robert C. Nichol, Matthew D. Olmstead, Francisco Prada, Sergio A. Rodríguez-Torres, Shun Saito, Salvador Salazar-Albornoz, Donald P. Schneider, Daniel Thomas, Jeremy Tinker, Rita Tojeiro, Yuting Wang, Martin White, and Gong-bo Zhao. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: observational systematics and baryon acoustic oscillations in the correlation function. *Monthly Notices of the Royal Astronomical Society*, 464(1):1168–1191, 09 2016. ISSN 0035-8711. doi: 10.1093/mnras/stw2372. URL `https://doi.org/10.1093/mnras/stw2372`.

[212] B.S. Ryden. *Introduction to Cosmology*. Addison-Wesley, 2003. ISBN 9780805389128. URL `https://books.google.cl/books?id=z27vAAAAMAAJ`.

[213] Bryan Sagredo, Javier Silva Lafaurie, and Domenico Sapone. Comparing dark energy models with hubble versus growth rate data, 2018.

[214] Bryan Sagredo, Savvas Nesseris, and Domenico Sapone. The Internal Robustness of Growth Rate data. 2018.

[215] Lado Samushia, Will J. Percival, and Alvise Raccanelli. Interpreting large-scale redshift-space distortion measurements. *Mon. Not. Roy. Astron. Soc.*, 420:2102–2119, 2012. doi: 10.1111/j.1365-2966.2011.20169.x.

[216] Ariel G. Sánchez, C. G. Scóccola, A. J. Ross, W. Percival, M. Manera, F. Montesano, X. Mazzalay, A. J. Cuesta, D. J. Eisenstein, E. Kazin, C. K. McBride, K. Mehta, A. D. Montero-Dorta, N. Padmanabhan, F. Prada, J. A. Rubiño-Martín, R. Tojeiro, X. Xu, M. Vargas Magaña, E. Aubourg, N. A. Bahcall, S. Bailey, D. Bizyaev, A. S. Bolton, H. Brewington, J. Brinkmann, J. R. Brownstein, J. Richard Gott, J. C. Hamilton, S. Ho, K. Honscheid, A. Labatie, E. Malanushenko, V. Malanushenko, C. Maraston, D. Muna, R. C. Nichol, D. Oravetz, K. Pan, N. P. Ross, N. A. Roe, B. A. Reid, D. J. Schlegel, A. Shelden, D. P. Schneider, A. Simmons, R. Skibba, S. Snedden, D. Thomas, J. Tinker, D. A. Wake, B. A. Weaver, David H. Weinberg, Martin White, I. Zehavi, and G. Zhao. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological implications of the large-scale two-point correlation function. MNRAS, 425(1):415–437, Sep 2012. doi: 10.1111/j.1365-2966.2012.21502.x.

[217] Ariel G. Sanchez et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological implications of the full shape of the clustering wedges in the data release 10 and 11 galaxy samples. *Mon. Not. Roy. Astron. Soc.*, 440(3):2692–2713, 2014. doi: 10.1093/mnras/stu342.

[218] Domenico Sapone. Dark Energy in Practice. *Int. J. Mod. Phys.*, A25:5253–5331, 2010. doi: 10.1142/S0217751X10050743.

[219] B. Sartoris, A. Biviano, C. Fedeli, J. G. Bartlett, S. Borgani, M. Costanzi, C. Giocoli, L. Moscardini, J. Weller, B. Ascaso, S. Bardelli, S. Maurogordato, and P. T. P. Viana. Next generation cosmology: constraints from the Euclid galaxy cluster survey. MNRAS, 459(2):1764–1780, June 2016. doi: 10.1093/mnras/stw630.

[220] W. Saunders, W. J. Sutherland, S. J. Maddox, O. Keeble, S. J. Oliver, M. Rowan-Robinson, R. G. McMahon, G. P. Efstathiou, H. Tadros, S. D. M. White, and et al. The pscz catalogue. *Monthly Notices of the Royal Astronomical Society*, 317(1):55–63, Sep 2000. ISSN 1365-2966. doi: 10.1046/j.1365-8711.2000.03528.x. URL `http://dx.doi.org/10.1046/j.1365-8711.2000.03528.x`.

[221] J. Schaye, R. A. Crain, R. G. Bower, M. Furlong, M. Schaller, T. Theuns, C. Dalla Vecchia, C. S. Frenk, I. G. McCarthy, J. C. Helly, A. Jenkins, Y. M. Rosas-Guevara, S. D. M. White, M. Baes, C. M. Booth, P. Camps, J. F. Navarro, Y. Qu, A. Rahmati,

T. Sawala, P. A. Thomas, and J. Trayford. The EAGLE project: simulating the evolution and assembly of galaxies and their environments. MNRAS, 446:521–554, January 2015. doi: 10.1093/mnras/stu2058.

[222] David J. Schlegel, Robert D. Blum, Francisco Javier Castand er, Arjun Dey, Douglas P. Finkbeiner, Sebastien Foucaud, Klaus Honscheid, David James, Dustin Lang, Michael Levi, John Moustakas, Adam D. Myers, Jeffrey Newman, Brian Nord, Peter E. Nugent, Anna Patej, Kevin Reil, Gregory Rudnick, Eli S. Rykoff, Eddie Ford Schlafly, Casey Stark, Francisco Valdes, Alistair R. Walker, Benjamin Weaver, and DECam Legacy Survey Collaboration. The Dark Energy Spectroscopic Instrument (DESI): The NOAO DECam Legacy Imaging Survey and DESI Target Selection. In *American Astronomical Society Meeting Abstracts #225*, volume 225 of *American Astronomical Society Meeting Abstracts*, page 336.07, January 2015.

[223] Gideon Schwarz. Estimating the Dimension of a Model. *Annals Statist.*, 6:461–464, 1978.

[224] Román Scoccimarro. Redshift-space distortions, pairwise velocities, and nonlinearities. *Phys. Rev. D*, 70:083007, Oct 2004. doi: 10.1103/PhysRevD.70.083007. URL `https://link.aps.org/doi/10.1103/PhysRevD.70.083007`.

[225] Román Scoccimarro and Ravi K. Sheth. PTHALOS: a fast method for generating mock galaxy distributions. MNRAS, 329(3):629–640, January 2002. doi: 10.1046/j.1365-8711.2002.04999.x.

[226] Hee-Jong Seo and Daniel J. Eisenstein. Probing dark energy with baryonic acoustic oscillations from future large galaxy redshift surveys. *The Astrophysical Journal*, 598 (2):720–740, Dec 2003. ISSN 1538-4357. doi: 10.1086/379122. URL `http://dx.doi.org/10.1086/379122`.

[227] R. Shaw, Michael Bridges, and M. P. Hobson. Clustered nested sampling: Efficient Bayesian inference for cosmology. *Mon. Not. Roy. Astron. Soc.*, 378:1365–1370, 2007. doi: 10.1111/j.1365-2966.2007.11871.x.

[228] Masatoshi Shoji, Donghui Jeong, and Eiichiro Komatsu. EXTRACTING ANGULAR DIAMETER DISTANCE AND EXPANSION RATE OF THE UNIVERSE FROM TWO-DIMENSIONAL GALAXY POWER SPECTRUM AT HIGH REDSHIFTS: BARYON ACOUSTIC OSCILLATION FITTING VERSUS FULL MODELING. *The Astrophysical Journal*, 693(2):1404–1416, mar 2009. doi: 10.1088/0004-637x/693/2/1404. URL `https://doi.org/10.1088%2F0004-637x%2F693%2F2%2F1404`.

[229] Joan Simon, Licia Verde, and Raul Jimenez. Constraints on the redshift dependence of the dark energy potential. *Phys. Rev.*, D71:123001, 2005. doi: 10.1103/PhysRevD.71.123001.

[230] M. F. Skrutskie, R. M. Cutri, R. Stiening, M. D. Weinberg, S. Schneider, J. M. Carpenter, C. Beichman, R. Capps, T. Chester, J. Elias, J. Huchra, J. Liebert, C. Lonsdale, D. G. Monet, S. Price, P. Seitzer, T. Jarrett, J. D. Kirkpatrick, J. E. Gizis, E. Howard,

T. Evans, J. Fowler, L. Fullmer, R. Hurt, R. Light, E. L. Kopan, K. A. Marsh, H. L. McCallon, R. Tam, S. Van Dyk, and S. Wheelock. The two micron all sky survey (2mass). *The Astronomical Journal*, 131(2):1163–1183, feb 2006. doi: 10.1086/498708. URL `https://doi.org/10.1086%2F498708`.

[231] S. A. Smee, J. E. Gunn, A. Uomoto, N. Roe, D. Schlegel, C. M. Rockosi, M. A. Carr, F. Leger, K. S. Dawson, M. D. Olmstead, J. Brinkmann, R. Owen, R. H. Barkhouser, K. Honscheid, P. Harding, D. Long, R. H. Lupton, C. Loomis, L. Anderson, J. Annis, M. Bernardi, V. Bhardwaj, D. Bizyaev, A. S. Bolton, H. Brewington, J. W. Briggs, S. Burles, J. G. Burns, F. J. Castander, A. Connolly, J. R. A. Davenport, G. Ebelke, H. Epps, P. D. Feldman, S. D. Friedman, J. Frieman, T. Heckman, C. L. Hull, G. R. Knapp, D. M. Lawrence, J. Loveday, E. J. Mannery, E. Malanushenko, V. Malanushenko, A. J. Merrelli, D. Muna, P. R. Newman, R. C. Nichol, D. Oravetz, K. Pan, A. C. Pope, P. G. Ricketts, A. Shelden, D. Sandford, W. Siegmund, A. Simmons, D. S. Smith, S. Snedden, D. P. Schneider, M. SubbaRao, C. Tremonti, P. Waddell, and D. G. York. The Multi-object, Fiber-fed Spectrographs for the Sloan Digital Sky Survey and the Baryon Oscillation Spectroscopic Survey. AJ, 146:32, August 2013. doi: 10.1088/0004-6256/146/2/32.

[232] Stephen Smee et al. The Multi-Object, Fiber-Fed Spectrographs for SDSS and the Baryon Oscillation Spectroscopic Survey. *Astron. J.*, 146:32, 2013. doi: 10.1088/0004-6256/146/2/32.

[233] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson. Structure in the COBE Differential Microwave Radiometer First-Year Maps. ApJ, 396:L1, September 1992. doi: 10.1086/186504.

[234] Rachel S. Somerville and Romeel Davé. Physical models of galaxy formation in a cosmological framework. *Annual Review of Astronomy and Astrophysics*, 53(1):51–113, 2015. doi: 10.1146/annurev-astro-082812-140951. URL `https://doi.org/10.1146/annurev-astro-082812-140951`.

[235] Yong-Seon Song and Will J. Percival. Reconstructing the history of structure formation using Redshift Distortions. *JCAP*, 0910:004, 2009. doi: 10.1088/1475-7516/2009/10/004.

[236] D. Spergel et al. Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report. 2015.

[237] Volker Springel. E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh. MNRAS, 401(2):791–851, January 2010. doi: 10.1111/j.1365-2966.2009.15715.x.

[238] Volker Springel and Lars Hernquist. The history of star formation in a Λ cold dark matter universe. MNRAS, 339(2):312–334, February 2003. doi: 10.1046/j.1365-8711.

2003.06207.x.

[239] Daniel Stern, Raul Jimenez, Licia Verde, Marc Kamionkowski, and S. Adam Stanford. Cosmic Chronometers: Constraining the Equation of State of Dark Energy. I: H(z) Measurements. *JCAP*, 1002:008, 2010. doi: 10.1088/1475-7516/2010/02/008.

[240] Adam R. H. Stevens and Toby Brown. Physical drivers of galaxies' cold-gas content: exploring environmental and evolutionary effects with Dark Sage. MNRAS, 471(1): 447–462, Oct 2017. doi: 10.1093/mnras/stx1596.

[241] Michael A. Strauss et al. Spectroscopic Target Selection in the Sloan Digital Sky Survey: The Main Galaxy Sample. *Astron. J.*, 124:1810, 2002. doi: 10.1086/342343.

[242] Hajime Sugai et al. Prime focus spectrograph: Subaru's future. *Proc. SPIE Int. Soc. Opt. Eng.*, 8446:84460Y, 2012. doi: 10.1117/12.926954.

[243] Nariaki Sugiura. Further analysts of the data by akaike' s information criterion and the finite corrections. *Communications in Statistics - Theory and Methods*, 7(1):13–26, 1978. doi: 10.1080/03610927808827599.

[244] Ariel G. Sánchez, C. M. Baugh, and R. Angulo. What is the best way to measure baryonic acoustic oscillations? *Monthly Notices of the Royal Astronomical Society*, Oct 2008. ISSN 1365-2966. doi: 10.1111/j.1365-2966.2008.13769.x. URL `http://dx.doi.org/10.1111/j.1365-2966.2008.13769.x`.

[245] E. Sánchez, A. Carnero, J. García-Bellido, E. Gaztañaga, F. de Simoni, M. Crocce, A. Cabré, P. Fosalba, and D. Alonso. Tracing the sound horizon scale with photometric redshift surveys. *Monthly Notices of the Royal Astronomical Society*, 411(1):277–288, 01 2011. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2010.17679.x. URL `https://doi.org/10.1111/j.1365-2966.2010.17679.x`.

[246] Ryuichi Takahashi, Masanori Sato, Takahiro Nishimichi, Atsushi Taruya, and Masamune Oguri. Revising the Halofit Model for the Nonlinear Matter Power Spectrum. ApJ, 761(2):152, December 2012. doi: 10.1088/0004-637X/761/2/152.

[247] Svetlin Tassev, Matias Zaldarriaga, and Daniel J. Eisenstein. Solving large scale structure in ten easy steps with COLA. J. Cosmology Astropart. Phys., 2013(6):036, June 2013. doi: 10.1088/1475-7516/2013/06/036.

[248] C. Tonini, C. Maraston, J. Devriendt, D. Thomas, and J. Silk. The impact of thermally pulsing asymptotic giant branch stars on hierarchical galaxy formation models. MNRAS, 396:L36–L40, June 2009. doi: 10.1111/j.1745-3933.2009.00657.x.

[249] Roberto Trotta. Applications of Bayesian model selection to cosmological parameters. *Mon. Not. Roy. Astron. Soc.*, 378:72–82, 2007. doi: 10.1111/j.1365-2966.2007.11738.x.

[250] Roberto Trotta. Bayes in the sky: Bayesian inference and model selection in cosmology. *Contemp. Phys.*, 49:71–104, 2008. doi: 10.1080/00107510802066753.

[251] Shinji Tsujikawa. Modified gravity models of dark energy. *Lect. Notes Phys.*, 800: 99–145, 2010. doi: 10.1007/978-3-642-10598-2_3.

[252] J. Turkey. Bias and confidence in not quite large samples. *The Annals of Mathematical Statistics*, 29:1–614, 1958.

[253] Stephen J. Turnbull, Michael J. Hudson, Hume A. Feldman, Malcolm Hicken, Robert P. Kirshner, and Richard Watkins. Cosmic flows in the nearby universe from Type Ia Supernovae. *Mon. Not. Roy. Astron. Soc.*, 420:447–454, 2012. doi: 10.1111/j.1365-2966.2011.20050.x.

[254] Benigno Uria, Marc-Alexandre Côté, Karol Gregor, Iain Murray, and Hugo Larochelle. Neural autoregressive distribution estimation. *Journal of Machine Learning Research*, 17(205):1–37, 2016. URL `http://jmlr.org/papers/v17/16-272.html`.

[255] A. Veropalumbo, F. Marulli, L. Moscardini, M. Moresco, and A. Cimatti. An improved measurement of baryon acoustic oscillations from the correlation function of galaxy clusters at z $\sim$ 0.3. *Monthly Notices of the Royal Astronomical Society*, 442(4):3275–3283, 06 2014. ISSN 0035-8711. doi: 10.1093/mnras/stu1050. URL `https://doi.org/10.1093/mnras/stu1050`.

[256] M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, S. Bird, D. Nelson, and L. Hernquist. Properties of galaxies reproduced by a hydrodynamic simulation. Nature, 509:177–182, May 2014. doi: 10.1038/nature13316.

[257] M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, D. Nelson, and L. Hernquist. Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. MNRAS, 444:1518–1547, October 2014. doi: 10.1093/mnras/stu1536.

[258] Michael J. Way, Jeffrey D. Scargle, Kamal M. Ali, and Ashok N. Srivastava. *Advances in Machine Learning and Data Mining for Astronomy.* Chapman & Hall/CRC, 1st edition, 2012. ISBN 143984173X.

[259] S. M. Weinmann, F. C. van den Bosch, X. Yang, and H. J. Mo. Properties of galaxy groups in the Sloan Digital Sky Survey - I. The dependence of colour, star formation and morphology on halo mass. MNRAS, 366:2–28, February 2006. doi: 10.1111/j.1365-2966.2005.09865.x.

[260] Holger Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in Computational Mathematics*, 4(1): 389–396, Dec 1995. ISSN 1572-9044. doi: 10.1007/BF02123482.

[261] Holger Wendland. Error estimates for interpolation by compactly supported radial basis functions of minimal degree. *Journal of Approximation Theory*, 93(2):258 – 272, 1998. ISSN 0021-9045. doi: https://doi.org/10.1006/jath.1997.3137.

[262] Anja Weyant, Chad Schafer, and W. Michael Wood-Vasey. LIKELIHOOD-FREE COSMOLOGICAL INFERENCE WITH TYPE ia SUPERNOVAE: APPROXIMATE

BAYESIAN COMPUTATION FOR a COMPLETE TREATMENT OF UNCERTAINTY. *The Astrophysical Journal*, 764(2):116, jan 2013. doi: 10.1088/0004-637x/764/2/116. URL `https://doi.org/10.1088%2F0004-637x%2F764%2F2%2F116`.

[263] Martin White, Jeremy L. Tinker, and Cameron K. McBride. Mock galaxy catalogues using the quick particle mesh method. MNRAS, 437(3):2594–2606, January 2014. doi: 10.1093/mnras/stt2071.

[264] S. D. M. White and C. S. Frenk. Galaxy formation through hierarchical clustering. ApJ, 379:52–79, September 1991. doi: 10.1086/170483.

[265] Bill S. Wright, Hans A. Winther, and Kazuya Koyama. COLA with massive neutrinos. *Journal of Cosmology and Astroparticle Physics*, 2017(10):054–054, oct 2017. doi: 10.1088/1475-7516/2017/10/054. URL `https://doi.org/10.1088%2F1475-7516%2F2017%2F10%2F054`.

[266] Xiaoying Xu, Nikhil Padmanabhan, Daniel J. Eisenstein, Kushal T. Mehta, and Antonio J. Cuesta. A 2% Distance to z=0.35 by Reconstructing Baryon Acoustic Oscillations - II: Fitting Techniques. *Mon. Not. Roy. Astron. Soc.*, 427:2146, 2012. doi: 10.1111/j.1365-2966.2012.21573.x.

[267] G. Yepes, R. Kates, A. Khokhlov, and A. Klypin. Hydrodynamical simulations of galaxy formation: effects of supernova feedback. MNRAS, 284(1):235–256, January 1997. doi: 10.1093/mnras/284.1.235.

[268] Donald G. York, J. Adelman, John E. Anderson, Jr., Scott F. Anderson, James Annis, Neta A. Bahcall, J. A. Bakken, Robert Barkhouser, Steven Bastian, Eileen Berman, and et al. The sloan digital sky survey: Technical summary. *The Astronomical Journal*, 120(3):1579–1587, Sep 2000. ISSN 0004-6256. doi: 10.1086/301513. URL `http://dx.doi.org/10.1086/301513`.

[269] Naoki Yoshida, Aaron Sokasian, Lars Hernquist, and Volker Springel. Early Structure Formation and Reionization in a Cosmological Model with a Running Primordial Power Spectrum. ApJ, 598(1):73–85, November 2003. doi: 10.1086/378852.

[270] I. Zehavi, M. R. Blanton, J. A. Frieman, D. H. Weinberg, H. J. Mo, M. A. Strauss, S. F. Anderson, J. Annis, N. A. Bahcall, M. Bernardi, J. W. Briggs, J. Brinkmann, S. Burles, L. Carey, F. J. Castander, A. J. Connolly, I. Csabai, J. J. Dalcanton, S. Dodelson, M. Doi, D. Eisenstein, M. L. Evans, D. P. Finkbeiner, S. Friedman, M. Fukugita, J. E. Gunn, G. S. Hennessy, R. B. Hindsley, Ž. Ivezić, S. Kent, G. R. Knapp, R. Kron, P. Kunszt, D. Q. Lamb, R. F. Leger, D. C. Long, J. Loveday, R. H. Lupton, T. McKay, A. Meiksin, A. Merrelli, J. A. Munn, V. Narayanan, M. Newcomb, R. C. Nichol, R. Owen, J. Peoples, A. Pope, C. M. Rockosi, D. Schlegel, D. P. Schneider, R. Scoccimarro, R. K. Sheth, W. Siegmund, S. Smee, Y. Snir, A. Stebbins, C. Stoughton, M. SubbaRao, A. S. Szalay, I. Szapudi, M. Tegmark, D. L. Tucker, A. Uomoto, D. Vanden Berk, M. S. Vogeley, P. Waddell, B. Yanny, and D. G. York. Galaxy Clustering in Early Sloan Digital Sky Survey Redshift Data. ApJ, 571:172–190, May 2002. doi: 10.1086/339893.

[271] I. Zehavi, Z. Zheng, D. H. Weinberg, J. A. Frieman, A. A. Berlind, M. R. Blanton, R. Scoccimarro, R. K. Sheth, M. A. Strauss, I. Kayo, Y. Suto, M. Fukugita, O. Nakamura, N. A. Bahcall, J. Brinkmann, J. E. Gunn, G. S. Hennessy, Ž. Ivezić, G. R. Knapp, J. Loveday, A. Meiksin, D. J. Schlegel, D. P. Schneider, I. Szapudi, M. Tegmark, M. S. Vogeley, D. G. York, and SDSS Collaboration. The Luminosity and Color Dependence of the Galaxy Correlation Function. ApJ, 630:1–27, September 2005. doi: 10.1086/431891.

[272] I. Zehavi, Z. Zheng, D. H. Weinberg, M. R. Blanton, N. A. Bahcall, A. A. Berlind, J. Brinkmann, J. A. Frieman, J. E. Gunn, R. H. Lupton, R. C. Nichol, W. J. Percival, D. P. Schneider, R. A. Skibba, M. A. Strauss, M. Tegmark, and D. G. York. Galaxy Clustering in the Completed SDSS Redshift Survey: The Dependence on Color and Luminosity. ApJ, 736:59–+, July 2011. doi: 10.1088/0004-637X/736/1/59.

[273] Cong Zhang, Han Zhang, Shuo Yuan, Tong-Jie Zhang, and Yan-Chun Sun. Four new observational $H(z)$ data from luminous red galaxies in the Sloan Digital Sky Survey data release seven. *Res. Astron. Astrophys.*, 14(10):1221–1233, 2014. doi: 10.1088/1674-4527/14/10/002.

[274] Gong-Bo Zhao et al. The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: a tomographic measurement of cosmic structure growth and expansion rate based on optimal redshift weights. 2018.