Metal Forests in Large Scale Structure: Insights from IGM Metallicity Using First-Year DESI Survey Data

Andrea Muñoz Gutiérrez

Supervised by:
Dr. Axel Ricardo de la Macorra Pettersson Moriel
Dr. Alma Xóchitl González Morales

Spring 2024

Abstract

Falta

& Resumen (Español)

2? 10? pg long

Contents

Chapter 1

Cosmological Background

Cosmology is the science that investigates the origin, structure, evolution, composition and fate of the Universe. It is studied from different angles: from the first imprints in temperature in background radiation after recombination to the pattern in which galaxies are distributed nowadays, it seeks to understand the nature of the Universe through theoretical models and observational data. The Lambda Cold Dark Matter (ΛCDM) model, currently the most accepted cosmological paradigm, provides a robust framework for describing the Large Scale Structure (LSS, see 1.5.2) of the Universe, dark matter (responsible for the gravitational attraction of different cosmic structures), and dark energy (to which the accelerated expansion of the Universe is attributed). However, alternative models and theories, including modifications to general relativity and the standard model of particle physics, offer different perspectives on cosmic phenomena. Complementary to these theoretical models, a variety of observational probes, such as the Cosmic Microwave Background (CMB), galaxy redshift surveys (explained in 2), Baryon Acoustic Oscillations (BAO, defined in 2.3.1), and the Lyman-α forest (see 3.2.1), offer critical insights into the past and present of the Universe. In this context, the Dark Energy Spectroscopic Instrument (DESI; see 2) survey, with its unprecedented precision, plays a fundamental role.

This chapter provides a foundational overview of cosmology, including a discussion of the ΛCDM model and alternative approaches that will be introduced later, setting the stage for a detailed explanation of the role that metal forests, the intergalactic medium (IGM, see 3.2.1), and Lyman-α forest contamination play in our understanding of cosmic structures and the evolution of the Universe.

Observations indicate that the Universe is conformed by four main components:

- ⋆ Baryonic matter refers to all known forms of what we call ordinary matter, made by elementary particles. It interacts via the strong nuclear force (which binds atomic nuclei together), the weak nuclear force (which governs certain types of particle decay and nuclear reactions), gravitationally, and electromagnetically. The latter interaction allows it to be studied by diverse instruments, like optic and radio telescopes, interferometers, spectrographs, X-ray and gamma-ray detectors, etc. It accounts for $\sim 5\%$ of the mass-energy of the Universe.

- ⋆ Radiation refers to all forms of energy that decrease with the expansion of the Universe. This includes photons and relativistic particles, such as neutrinos in the early Universe. Radiation was the dominant component in the earliest stages of cosmic history and played a major role

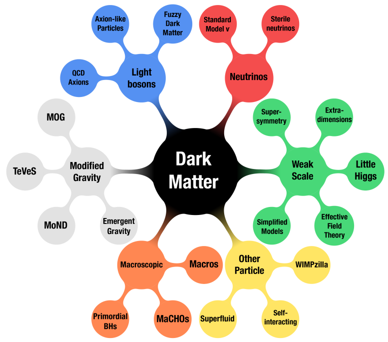


Figure 1.1: Possible explanations for the nature of dark matter (Credit: G. Bertone and T. M. P. Tait) [2]

in shaping its evolution.

- Dark matter is thought to be a form of matter that does not interact with electromagnetic radiation —meaning it neither emits, absorbs, nor scatters light—, making it detectable only through its gravitational effects on visible matter, the bending of light (gravitational lensing), and the dynamics of cosmic structures such as galaxies, galaxy clusters, and the LSS (see 1.5.2). It makes up $\sim 27\%$ of the total mass-energy content of the Universe. Despite of thorough studies of diverse characteristics, the true nature of dark matter remains to be one of the most profound questions in modern physics. Although various theoretical models have been proposed to explain its existence, a dark matter particle has not yet been detected. Depending on the energy of the hypothetical particles that make up dark matter, different models describe their behavior. These include cold dark matter (CDM), the most favored model, consisting of non-relativistic particles that lead to hierarchical structure formation—where small structures form first and later merge into larger ones. Another possibility is warm dark matter (WDM), a constrained but not entirely ruled-out model, composed of particles that were relativistic in the early universe but are slow-moving today, suppressing small-scale structures compared to CDM. Lastly, hot dark matter (HDM) consists of highly relativistic particles that smooth out structure on small scales, preventing early galaxy formation. Since this contradicts observed structure formation, HDM is considered ruled out as the primary form of dark matter. There are also several theoretical models developed to explain its nature, like Weakly Interacting Massive Particles (WIMPs), which are stable, weakly interacting particles predicted by supersymmetry; axions, light neutral particles proposed to solve the strong CP problem in QCD and potentially detectable via their interaction with photons; and primordial black holes (PBHs), which could have formed in the early Universe and contribute to dark matter under specific constraints [1]. Additionally, more exotic candidates have been proposed. Many of these are shown in 1.1 [2].

- Dark Energy, a type of energy responsible for the accelerated expansion of the Universe that acts like a negative pressure that counteracts gravity. It does not interact electromagnetically in any way understood nowadays, and its nature is one of the most intriguing questions of

distinguir entre figs/sections etc....

modern physics. There are several models that intend to explain it, among them modifications to General Relativity. Numerous efforts and experiments are dedicated to studying and unraveling its nature, such as the Dark Energy Survey (DES; [3], [4]), the Dark Energy Spectroscopic Instrument (DESI; see 2 [5], [6]), the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST; [7]) the Euclid mission ([8]), and the upcoming Nancy Grace Roman Space Telescope ([9], [10]). These projects use a variety of observational techniques—including supernovae (see 1.5.6), baryon acoustic oscillations (see 1.5.5, 2.3.1), weak gravitational lensing (see 1.5.3, and galaxy clustering (see 1.5.7)—to probe the expansion history and structure of the Universe.

1.0.1 History of the Universe

Since the 1920s, when Edwin Hubble discovered that the Universe is expanding, several observations and experiments have been focused on the nature of this phenomenon. Together with Georges Lemaître, this led to the development of the Hubble-Lemaître equation, which describes this expansion mathematically: the velocity v at which a galaxy is moving away from us due to this effect is given by the distance to the galaxy multiplied by a constant (the Hubble constant), H_0, so that $v = H_0 d$. This equation has been crucial in understanding the expansion dynamics of the Universe over time starting with a hot, dense early Universe. The measurement of H_0 across different epochs has provided valuable insights into the various evolutionary stages of the Universe since the Universe was smaller, hotter, and more energetic. These stages are illustrated in Figure 1.2[11], which highlights key periods in the history of the Universe that are essential for reconstructing its life story.

aquí mayuscula/ abajo minuscula

- ⋆ **Inflation:** this initial period of the Universe happened when it was 10^{-34} sec to 10^{-32} sec old. During this epoch, it experienced nearly exponential growth, driven by the potential energy of a scalar field known as the inflaton. After this expansion, the inflaton field decayed, transferring its energy into particles and radiation in a process known as reheating. During this transition, quantum fluctuations in the inflaton field were stretched to cosmological scales and became the seeds of primordial density perturbations in the relativistic plasma, which later grew into the large-scale structure we observe today.

- ⋆ **Recombination:** the primordial plasma, composed mainly of protons and electrons coupled with photons through Thomson scattering, served as the medium in which temperature fluctuations propagated as sound waves producing density fluctuations. As the Universe expanded, its temperature dropped until photons decoupled from protons and electrons, allowing the formation of hydrogen atoms. This caused the patterns made by the density fluctuations to *freeze*, and leave an imprint in this first emitted light, which is the Cosmic Microwave Background Radiation (CMB), as well as in baryonic matter began to clump due to gravitational attraction, falling into dark matter gravity potential wells, eventually evolving into structure and forming the Baryon Acoustic Oscillations (BAO, explained in more detail in 2.3.1). These oscillations are observable in the distribution pattern of galaxies within the Large Scale Structure (LSS, see 1.5.2) of the Universe.

- ⋆ **Dark Ages and reionization:** after recombination, there is a period, called *Dark Ages* in which the Universe is not observable in any wavelength, except for some emissions in the 21cm line. During this period, matter collapses gravitationally into overdense regions, gradually giving rise to the first structures. These regions lead to the creation of dark matter halos, the formation of gas clouds (also collapsing gravitationally), and the emergence of the first stars. The radiation from these stars and early Active Galactic Nuclei (AGN) ionized the Intergalactic Medium (IGM). The phase of the Universe when the neutral hydrogen was ionized by the first generation of stars is called *reionization*.

⋆ **Accelerated Expansion of the Universe:** after several gigayears (Gyr), the universe's expansion began to accelerate, presumably due to the negative pressure attributed to the dominance of dark energy (or the cosmological constant Λ) over radiation and matter, counteracting gravitational attraction.

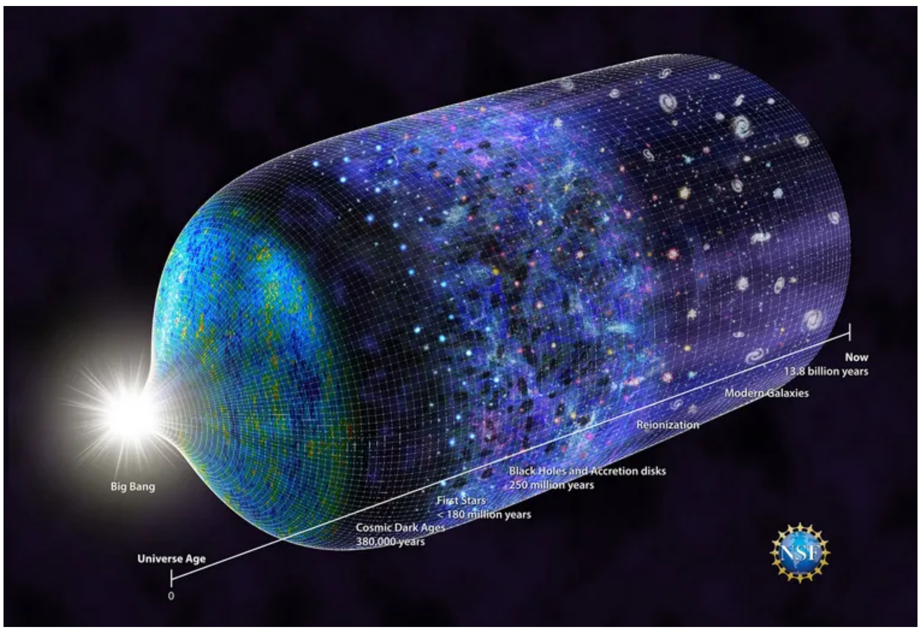


Figure 1.2: Timeline of the history of the Universe. Of particular interest, one can identify the stages of inflation (right after the Big Bang and before the emission of the CMB), recombination (associated with the CMB map), dark ages, reionization, and accelerated expansion of the Universe (labeled implicitly under the term "modern galaxies"). Credit: N.R.Fuller, National Science Foundation. [12]

i.e. Fig? section?

1.1 Cosmological Model

The observations explained in 1.5 support the ΛCDM cosmological model, which is the standard model in cosmology.

It is based on General Relativity and the cosmological principle, that states that the distribution of matter and radiation is homogeneous and isotropic at large scales, and establishes a universe described by the Friedmann-Lemâitre-Robertson-Walker (FLRW) metric:

\sin

$$ds^2 = c^2dt^2 - a(t)^2\left(\frac{dr^2}{1-kr^2} + r^2d\theta^2 + r^2sin^2\theta d\phi\right). \tag{1.1}$$

The first term of this expression represents the time dimension, while the next three correspond to the spatial components. Here, r, θ, ϕ are the spherical coordinates, and $\kappa = -1, 0, 1$ denotes the curvature of the Universe ($\kappa = -1$ indicates an open Universe, $\kappa = 0$ a flat Universe, and $\kappa = 1$ a closed Universe). The term $a(t)$, known as the *scale factor*, describes the "size" of the Universe relative to the present time, where $a = a_0 = 1$.

The starting point for deriving the Friedmann equation is *Einstein's field equations*:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}, \tag{1.2}$$

where $G_{\mu\nu}$ is the Einstein tensor, Λ is the cosmological constant, $g_{\mu\nu}$ is the metric tensor, and $T_{\mu\nu}$ is the energy-momentum tensor. Assuming a perfect fluid form for $T_{\mu\nu}$ and applying the FLRW metric, one solves the $(0,0)$ component (the time-time component) to obtain the *Friedmann equation*:

y Λ?

$$H^2 = \left(\frac{\dot a}{a}\right)^2 = \frac{8\pi G\rho}{3} - \frac{\kappa c^2}{a^2}, \tag{1.3}$$

with $H = \frac{\dot a}{a}$ representing the Hubble parameter, which determines the expansion rate of the universe, G as the newtonian gravitational constant, and ρ is the total energy density of the Universe. In the rest of this work, we adopt natural units where $c = 1$.

In addition to the Friedmann equation, the Einstein field equations applied to the FLRW metric also yield, from the $i = j$ components, the *continuity equation*:

segura? revisar!

$$\dot\rho = -3H(\rho + p), \tag{1.4}$$

where p represents the pressure of each fluid with energy density ρ, related by $p = w\rho$ ($w = 0$ for matter, $w = 1/3$ for radiation). Using equations 1.3 and 1.4 we obtain ~~a third one~~: siempre ?

$$\dot H = -4\pi G(\rho + p) + \frac{\kappa}{a^2}. \tag{1.5}$$

y Λ?

The ΛCDM model describes a universe composed of regular matter, as described by the Standard Model of Elementary Particles, alongside an unidentified form of non-relativistic matter known as Cold Dark Matter (CDM), which interacts solely through gravity. Additionally, the model accounts for accelerated expansion, attributed to a cosmological constant Λ, characterized by an equation of state given by $w = p/\rho = -1$, where p denotes pressure and ρ represents density. So, considering the equation:

$$\rho = \rho_i\left(\frac{a}{a_i}\right)^{-3(w+1)}, \tag{1.6}$$

si ?

valid for any fluid that follows the continuity equation, $\rho_\Lambda = \Lambda$, it follows that $p_\Lambda = -\rho_\Lambda = \Lambda$.

We define the *critical density* ρ_{crit} as the average matter density needed for the Universe to have a flat geometry ($k = 0$), assuming the cosmological principle. Following 1.3, the critical density is defined as:

no es la densidad crítica,
revisar

$$\rho_{crit} = \frac{3H_0^2}{8\pi G}, \tag{1.7}$$

with H_0 the Hubble constant, and G the newtonian gravitational constant.

The comparison between the different components of the Universe and the critical density is given by:

$$\Omega_i = \rho_i/\rho_{crit}, \tag{1.8}$$

where i refers to ordinary matter, dark matter, dark energy, etc. $\Omega_{total} = \sum_i \Omega_i = 1$ for a flat universe, $\Omega_{total} < 1$ for an open universe, and $\Omega_{total} > 1$ for a closed one.

Observations from the CMB (see 1.4.2 and 1.5) indicate that the Universe has a geometry very close to flat, and that its components are approximately in the proportions shown in figure 1.3.

aquí min.

- ⋆ $\Omega_b \approx 0.05$: around 5% is composed of ordinary matter[1]
- ⋆ $\Omega_m \approx 0.25$: 25% dark matter, a type of matter that interacts only gravitationally with ordinary matter.
- ⋆ $\Omega_\Lambda \approx 0.7$: the remaining 70% is Λ, dark energy, responsible for the expansion of the universe at an accelerated rate.

Depending on the observations used to obtain these results, the values can vary slightly.

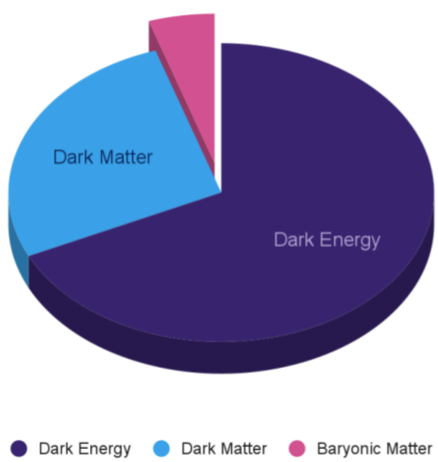


Figure 1.3: Pie chart illustrating the components of the universe and their approximate proportions according to the ΛCDM model. Around 5% is composed of ordinary matter, 25% dark matter, and the remaining 70% is dark energy. Depending on the probes used to obtain these results, the values can vary slightly. [13]

Aside from the ΛCDM model, there is a wide variety of alternative models that try to explain all the cosmological observations consistently. Among those, the $w_0 w_a$CDM model extends this

[1] Also called baryonic matter, and refers to the matter that we know.

framework by allowing a time-varying equation of state for dark energy, given by $w(t) = w_0 + w_a(t/1+t)$, cite - see Sébastien's corrections offering a more refined understanding of its role in cosmic expansion. Nonetheless, ΛCDM has been the most popular and, as said before, it is the standard model in Cosmology. It is the foundational framework used by the DESI survey. All the methodology and results of this research work assume this model.

1.2 Cosmological Redshift

In the 1920s Edwin Hubble observed that nearly all galaxies appear to move away from us, and that their recession velocities increase proportionally to their distance from us. This discovery leads to the conclusion of the expansion of the Universe. The recession velocity can be measured by means of the shift of the light towards the red end of the spectrum. This is called cosmological redshift. Objects that are farther away will have a larger recession velocity and hence a larger redshift.

Due to the change in energy levels of electrons, photons are emitted by sources of radiation (or absorbed by the medium between the source and the observer) by different elements of the periodic table at specific and unique wavelengths. These emissions and absorptions have an energy given by the equation $E = h\nu = h/\lambda$, where h is the Planck constant, ν is the frequency of the emission or absorption, and λ is the corresponding wavelength, with $c = 1$.

As the universe expands, the wavelength of these photons that propagate freely increases, as all physical distances expand. This causes the wavelength of the photons to shift to the red end of the electromagnetic spectrum. This implies that this red-shift of a photon is due to the fact that the Universe was smaller when the photon was emitted. The wavelength of a photon is inversely proportional to the scale factor. Thus, the wavelength at time t_0, denoted by λ_0, will differ from that at time t_1, denoted by λ_1, according to $\frac{\lambda_1}{\lambda_0} = \frac{a(t_1)}{a(t_0)}$.

In observational cosmology, when the source of the emission or absorption moves relative to the observer, the wavelength appears shifted: towards the blue end of the electromagnetic spectrum (blueshifted) if the source is approaching, or toward the red end (redshifted) if the source is receding. Observations of galaxies show that their spectra are redshifted, which is evidence of the expansion of the Universe.

To define the redshift z, we consider a light beam emitted by a source at time t_i from position r_i. This beam is detected by the observer at position $r = r_0 = 0$ at time t_0. According to General Relativity, $ds^2 = 0$ by definition, so from the FLRW metric:

$$dt = \frac{a(t)dr}{\sqrt{1-\kappa r^2}}, \tag{1.9}$$

and

$$\int_{t_i}^{t_0} \frac{1}{a(t)}\, dt = \int_{r_i}^{r_0} \frac{1}{\sqrt{1-\kappa r^2}}\, dr. \tag{1.10}$$

Let us also suppose that the wave begins to be emitted at time t_i (and will be observed at t_0), and at time $t_i + \delta t_i$, a different "part" of the same wave is emitted (which will be observed at $t_0 + \delta t_0$). Since r_0 and r_i are comoving, from (1.10):

$$\int_{t_i+\delta t_i}^{t_0+\delta t_0} \frac{1}{a(t)}\, dt = \int_{r_i}^{r_0} \frac{1}{\sqrt{1-\kappa r^2}}\, dx, \tag{1.11}$$

hence

$$\int_{t_i}^{t_0} \frac{1}{a(t)}\, dt = \int_{t_i+\delta t_i}^{t_0+\delta t_0} \frac{1}{a(t)}\, dt. \tag{1.12}$$

If we sum $\int_{t_0}^{t_i+\delta t_i} \frac{1}{a(t)}\, dt$ to both sides, we obtain:

$$\int_{t_i}^{t_i+\delta t_i} \frac{1}{a(t)}\, dt = \int_{t_0}^{t_0+\delta t_0} \frac{1}{a(t)}\, dt. \tag{1.13}$$

We note that δt can be as small as needed, then the change in $a(t)$ is negligible in the $t, t+\delta t$ interval, so it is considered constant. Hence, following the discussion in 1.40:

$$\frac{\delta t_0}{\delta t_i} = \frac{a(t_0)}{a(t_i)} = \frac{\lambda_0}{\lambda_i}. \tag{1.14}$$

From this expression ~~we~~ see that if $a(t_0) > a(t_i)$, then $\lambda_0 > \lambda_i$. This means that the observed wavelength is longer then it was when emitted (i.e. redshifted). This shift is a consequence of the expansion of the Universe, and objects that present this behavior are tracers of this expansion. This *cosmological redshift* (or simply *redshift*) z is defined as:

$$\frac{a(t_0)}{a(t_i)} = 1 + z = \frac{\lambda_0}{\lambda_i} \tag{1.15}$$

$$z = \frac{\lambda_0}{\lambda_1} - 1 = \frac{\lambda_0 - \lambda_1}{\lambda_1} \tag{1.16}$$

Due to the expansion of the Universe, objects that are farther away have higher receding velocities and, therefore, greater redshifts. Since light from high-redshift galaxies was emitted when the Universe was younger, observing these galaxies allows us to study the evolution of galaxies over cosmic time. Statistically, high-redshift galaxies are the progenitors of present-day galaxies, and any changes in galaxy number density or intrinsic properties with redshift provide a direct window into the formation and evolution of the galaxy population. With modern large telescopes, we can observe galaxies at $z > 10$, enabling us to explore the galaxy population from a time when the Universe was only about 10% of its current age.

1.3 Distances

This section is based on material extracted from [14].

The FLRW metric (Equation 1.1) provides a framework for calculating cosmological distances. Following the discussions in [15] and [16], $d\Omega = r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2$ defines the angular line element on a unit sphere. When scaled by r^2, it represents angular separations in spherical coordinates. Setting $d\Omega = 0$ corresponds to aligning the coordinate axes in a convenient way, restricting the analysis to purely radial motion, simplifying cosmological distance calculations. We then define the following distances:

Comoving Distance

The *comoving distance* Δr quantifies the separation between two points r_1 and r_2 in the Universe, factoring out its expansion. It remains constant with time for objects moving with the Hubble flow (see 1.5.4).

For a photon traveling radially (i.e., $d\theta = d\phi = 0$) along a null geodesic ($ds^2 = 0$) in the FLRW metric:

regreso?

$$ds^2 = -c^2 dt^2 + a(t)^2 \frac{dr^2}{1 - \kappa r^2} = 0, \tag{1.17}$$

we obtain the relation:

$$\frac{dr}{\sqrt{1 - \kappa r^2}} = \frac{c\, dt}{a(t)}. \tag{1.18}$$

Integrating both sides gives the comoving distance Δr between times t_1 and t_2:

$$\Delta r = \int_{t_1}^{t_2} \frac{c\, dt}{a(t)}. \tag{1.19}$$

In natural units where $c = 1$, this simplifies to:

ya se habia dicho

$$\Delta r = \int_{t_1}^{t_2} \frac{dt}{a(t)}. \tag{1.20}$$

Alternatively, one can express the comoving distance as a spatial integral over the radial coordinate:

$$\Delta r = \int_{r_1}^{r_2} \frac{1}{\sqrt{1 - \kappa r^2}}\, dr, \tag{1.21}$$

which yields:

$$\Delta r = \begin{cases} r_2 - r_1 & \text{if} \quad \kappa = 0, \\ \arcsin(r_2) - \arcsin(r_1) & \text{if} \quad \kappa = +1, \\ \text{arcsinh}(r_2) - \text{arcsinh}(r_1) & \text{if} \quad \kappa = -1. \end{cases} \tag{1.22}$$

This distance represents the coordinate separation between two points, unaffected by cosmic expansion, and serves as a foundational concept in defining other cosmological distances.

Physical Distance

The *physical distance* is the real distance between two points with comoving coordinates r_1 and r_2, and does change with the expansion of the Universe. It is given by:

$$D = \int_{r_1}^{r_2} dl, \tag{1.23}$$

with $dl = \frac{a(t)}{\sqrt{1-\kappa r^2}}\, dr$ the spatial part of 1.1, hence:

$$D(t) = a(t) \int_{r_1}^{r_2} \frac{1}{\sqrt{1-\kappa r^2}}\, dr \tag{1.24}$$

$$D(t) = a(t)\Delta r. \tag{1.25}$$

For two given times t_a and t_b, the physical distances are:

$$D(t_a) = a(t_a) \int_{r_1}^{r_2} \frac{1}{\sqrt{1-\kappa r^2}}\, dr, \tag{1.26}$$

$$D(t_b) = a(t_b) \int_{r_1}^{r_2} \frac{1}{\sqrt{1-\kappa r^2}}\, dr. \tag{1.27}$$

Using these two expressions, and considering that comoving distance does not change with the expansion of the Universe, the physical distance at time t_a, $D(t_a)$, can be determined if the scale factor a is known at both t_a and t_b, along with the physical distance at t_b, $D(t_b)$:

$$D(t_a) = \frac{a(t_a)}{a(t_b)} D(t_b). \tag{1.28}$$

Proper Distance

The *proper distance*, which refers to the distance a light beam travels, is given by:

$$d_p(t) = c \int_{t_i}^{t} dt = c(t - t_i). \tag{1.29}$$

For a light beam, $ds^2 = 0$ (as light travels along null geodesics). If $d\phi = 0$ and $d\theta = 0$, we have:

$$dt = \frac{a(t)dr}{\sqrt{1-\kappa r^2}}, \tag{1.30}$$

so:

$$d_p(t) = a(t) \int_{r_1}^{r_2} \frac{1}{\sqrt{1-\kappa r^2}}\, dr = a(t) \int_{t_i}^{t} \frac{1}{a(t')}\, dt'. \tag{1.31}$$

If we set $t_i = 0$ as the initial time and t_0 as the current time, the proper distance becomes the *distance to the horizon*, which defines the region of the universe that is causally connected to an observer:

$$d_H(t_0) = a(t_0) \int_{0}^{t_0} \frac{1}{a(t)}\, dt. \tag{1.32}$$

During different epochs of the Universe, various components such as radiation, matter, and dark energy coexist, each influencing its expansion. The *distance to the horizon* depends on the relative contributions of these components. Using the Friedmann equation:

$$H^2 = \frac{8\pi G\rho}{3} - \frac{\kappa}{a^2}, \tag{1.33}$$

where the total density ρ is the sum of the densities of radiation, matter, and dark energy:

$$\rho = \rho_r + \rho_m + \rho_\Lambda = \rho_{r_0}\left(\frac{a}{a_0}\right)^{-4} + \rho_{m_0}\left(\frac{a}{a_0}\right)^{-3} + \rho_{\Lambda_0}. \tag{1.34}$$

Multiplying and dividing this expression by the critical density today $\rho_{c_0} = \frac{3H_0^2}{8\pi G}$, and defining $\Omega_i = \frac{\rho_i}{\rho_c}$, we have:

$$\rho = \rho_{c_0}\left(\Omega_{r_0}\left(\frac{a}{a_0}\right)^{-4} + \Omega_{m_0}\left(\frac{a}{a_0}\right)^{-3} + \Omega_{\Lambda_0}\right), \tag{1.35}$$

so the Hubble parameter can be expressed as:

$$H = H_0\sqrt{\Omega_{r_0}\left(\frac{a}{a_0}\right)^{-4} + \Omega_{m_0}\left(\frac{a}{a_0}\right)^{-3} + \Omega_{\Lambda_0} - \left(\frac{a}{a_0}\right)^{-2}(1-\Omega_{T_0})}. \tag{1.36}$$

If the scale factor today is $a_0 = 1$ and $\Omega_{\kappa_0} = 1 - \Omega_{T_0}$, we can write the equation for the horizon distance as:

$$\int \frac{dt}{a(t)} = \frac{1}{H_0}\int \frac{da}{a^2}\frac{1}{\sqrt{\Omega_{r_0}a^4 + \Omega_{m_0}a + \Omega_{\Lambda_0} + \Omega_{\kappa_0}a^2}}. \tag{1.37}$$

Luminosity Distance

Certain objects in the Universe, such as variable stars and Type Ia supernovae (SNIa), have well-understood intrinsic luminosity L that can be determined from their observed flux F. By establishing a relationship between the flux and the luminosity of these objects, we can estimate their distance from us. This concept is referred to as the *luminosity distance* d_L. The relationship between F and L is given by:

$$F = \frac{L}{4\pi d_L^2}, \tag{1.38}$$

where L is:

$$L = \frac{E}{\delta t}. \tag{1.39}$$

In this expression, E represents the energy of the photons involved, and δt denotes a specific time interval. The observed flux changes due to the expansion of the Universe. Thus, for light:

$$\frac{\delta t_0}{\delta t_1} = \frac{\lambda_0}{\lambda_i} = \frac{a_0}{a(t_1)} = \frac{E_1}{E_0}, \tag{1.40}$$

where the quantities with the subscript 0 are measured at the present time, while those with the subscript 1 correspond to the time of emission.

If we observe a luminosity on Earth at the present time t_0, it would be given by:

$$L_0 = \frac{E_0}{\delta t_0}. \tag{1.41}$$

From (1.40),

$$E_1 = \frac{a_0}{a(t_1)} E_0, \tag{1.42}$$

and

$$\delta t_1 = \frac{a(t_1)}{a_0} \delta t_0, \tag{1.43}$$

then

$$L_0 = \frac{E_0}{\delta t_0} = \left(\frac{a(t_1)}{a_o}\right)^2 \frac{E_1}{\delta t_1} = \left(\frac{a(t_1)}{a_o}\right)^2 L_1. \tag{1.44}$$

As mentioned earlier, the physical distance is related to the comoving distance through the scale factor. For a given time t_1 and the present time t_0, this relationship can be expressed as:

$$d(t_1) = a(t_1)\Delta r, \text{ and } d_0 = a_0 \Delta r, \tag{1.45}$$

so for t_1:

$$d(t_1) = \frac{a(t_1)}{a_0} d(t_0) = a(t_1)\Delta r. \tag{1.46}$$

The flux observed at the present time is:

$$F_0 = \frac{L_0}{4\pi d_o^2} = \left(\frac{a(t_1)}{a_o}\right)^2 \frac{L_1}{4\pi a_0^2 (\Delta r)^2} = \frac{L_1}{4\pi a_0^2 (\Delta r)^2 (1+z)^2}, \tag{1.47}$$

and the luminosity distance for an expanding universe is defined as:

$$d_L = a_0 \Delta r (1+z). \tag{1.48}$$

From the Friedmann equation:

$$H^2 = H_0^2 \left(\frac{8\pi G\rho}{3H_0^2} - \frac{\kappa}{a^2 H_0^2}\right) = H_0^2 \left(\sum_i \Omega_{0_i} \left(\frac{a}{a_0}\right)^{-3(1+w)} - \frac{\kappa}{a^2 H_0}\right), \tag{1.49}$$

where the sum is over all the components of the Universe (baryonic matter, dark matter, radiation, and dark energy). Substituting $\kappa/a^2 H_0$ in terms of Ω_T:

$$H^2 = H_0^2 \left(\sum_i \Omega_{0_i} \left(\frac{a}{a_0}\right)^{-3(1+w)} + (1+\Omega_T) \left(\frac{a}{a_0}\right)^{-2}\right) \tag{1.50}$$

For light, $\Delta r = \int_{t_1}^{t_0} \frac{dt}{a(t)}$. With this, the comoving distance turns into:

$$\Delta r = \int \frac{da}{a^2 H_0 \left(\sum_i \Omega_{0_i} \left(\frac{a}{a_0}\right)^{-3(1+w)} + (1+\Omega_T)\left(\frac{a}{a_0}\right)^{-2}\right)^{(1/2)}} \tag{1.51}$$

and the luminosity distance into:

$$d_L = \frac{a_0(1+z)}{a_0} \int_{a_0}^{a} \frac{da}{a^2 \left(\sum_i \Omega_{0_i} \left(\frac{a}{a_0}\right)^{-3(1+w)} + (1+\Omega_T)\left(\frac{a}{a_0}\right)^{-2}\right)^{(1/2)}}. \tag{1.52}$$

Angular Distance

Lets consider an object of size D, at a distance d, that subtends an angular size θ. The angular-diameter size d_A is given by:

$$\theta = \frac{D}{d_A}, \text{ so that } d_A = \frac{D}{\theta} \tag{1.53}$$

To express d_A in terms of the FLRW metric, D represents the proper distance between two light signals emitted from two points with the same radial coordinate r_e at a time t_e, reaching the observer at t_0. D will then be the integral of the spatial part of the metric (1.1):

$$D = a_e r_e \int d\theta = \frac{a_0 r_e}{1+z}\theta \tag{1.54}$$

with $a_0 = a(t_0)$ and $a_e = a(t_e)$. Substituting this result in 1.53,

$$d_A = \frac{a_0 r_e}{1+z} = a_e r_e \tag{1.55}$$

In a space with no expansion, $d_A = d_L = d$, but in the cosmological context these values differ.

1.4 Statistical Tools and Cosmological Parameters

There is a set of fundamental quantities that describe the key properties and dynamics of the Universe at large scales. These are known as *cosmological parameters*. They characterize key aspects, like the expansion rate of the Universe, its composition (contents of matter, energy, dark matter, etc.), geometry, and the evolution of cosmic structures. They set the theoretical basis to interpret observations such as the cosmic microwave background radiation (CMB, see 1.0.1), galaxy distribution (see 1.5.2), and gravitational lensing (see 1.5.3), among others. Measuring these parameters with high precision allows us to understand the current state of the universe, as well as to reconstruct its history evolution and possible fate. This makes them an essential link between fundamental theories, cosmological models, and cosmological and astronomical observations. The following subsections will introduce the measured cosmological parameters. Before doing so, a couple of statistical tools used to determine them will be defined: the power spectrum and the correlation function.

1.4.1 Power Spectrum $P(k)$ and Correlation Functions (ξ)

As explained further in 1.5.2, at large scales, matter in the Universe is distributed in different structures, such as clusters of galaxies, filaments, voids, etc., forming the sponge-like-structure called the *cosmic web*. These structures do not have a random distribution, but follow patterns that can be measured with statistical tools that help us quantify how galaxies and gas group together, and how this changes in different evolutionary stages of the Universe.

Among these statistical tools we find the *2-point correlation function* ξ, which measures the excess probability to find galaxy pairs separated by a given distance r, with respect to a random distribution, and the *power spectrum* $P(k)$, the Fourier transform of ξ, which measures the amplitude of density fluctuations as function of wavenumber k.

Correlation Functions $\xi(r)$

As previously mentioned, the correlation function $\xi(r)$ quantifies how much more likely it is to find galaxy pairs separated by a distance r compared to a purely random distribution of galaxies. This quantifies galaxy clustering and provides a direct estimation of their distribution in real space. There are several estimators of the correlation function [17]. The most commonly used expression for the correlation function is the Landy & Szalay estimator.

$$\xi(r) = \frac{DD(r) - 2DR(r) + RR(r)}{RR(r)}, \tag{1.56}$$

where D means *data*, and R means *random*, DD is the number of galaxies separated by a distance r, and DR the number pairs of real galaxies separated at a distance r from one in a random catalog distribution.

Power Spectrum $P(k)$

As introduced before, the power spectrum $P(k)$ measures the amplitude of density fluctuations as a function of the wavenumber k, representing the distribution of matter in the Universe on different scales. This is expressed by the relation $P(k) = \langle|\delta_k|^2\rangle$, where δ_k represents the Fourier modes of the density fluctuation field. This specific form describes the matter power spectrum, which quantifies how matter clusters in Fourier space. In general, however, power spectra can be defined for different cosmological fields, such as temperature anisotropies in the cosmic microwave background or galaxy number counts.

Interpretation of $P(k)$

The wavenumber k is associated with spatial scales, where smaller values of k correspond to larger scales, and larger values of k correspond to smaller scales. The shape and amplitude of $P(k)$ depend on cosmological parameters defined previously, such as Ω_m (matter density parameter), Ω_b (baryon density parameter), H_0, and the equation of state of dark energy. The shape of the power spectrum features a turnover at small k (large scales), marking the transition from the linear to the nonlinear regime of structure formation. The slope at intermediate scales gives information about the matter content of the Universe, and the shape of $P(k)$ at large k (small scales) is affected by non-linear gravitational evolution and galaxy formation. The amplitude of $P(k)$ depends on Ω_m and the amplitude of primordial density fluctuations, and is affected by how galaxies trace (dark) matter distribution, given by the galaxy bias b.

Since $P(k)$ is the Fourier transform of $\xi(r)$, both contain the same fundamental information, but each function represents it in a manner tailored to different types of analysis.
To transform from real to Fourier space (from $\xi(r)$ to $P(k)$) we integrate the power spectrum over all wavenumbers:

$$\xi(\vec{r}) = \frac{1}{(3\pi)^3} \int P(k) e^{i\vec{k}\cdot\vec{r}} d^3\vec{k}. \tag{1.57}$$

Conversely, from Fourier to real space (from $P(k)$ to $\xi(r)$) we integrate the correlation function over all space:

$$P(\vec{k}) = \int \xi(r) e^{-i\vec{k}\cdot\vec{r}} d^3\vec{r}. \quad (1.58)$$

The measurement of $P(k)$ and $\xi(r)$ at different redshifts provides valuable information about the expansion of the Universe and the growth of structure.

1.4.2 Measured Cosmological parameters

There are several fundamental and derived quantities that describe different aspects of the Universe, such as its geometry, composition, and evolution. These encompass information about the energy density of matter, radiation, dark matter, dark energy, etc, primordial perturbations, the rate of expansion of the Universe, among others, and they define the basis of the cosmological models. This information is measured and constrained by different observations (explained in the next section, 1.5 and, as described before, some are embedded in $P(k)$ and $\xi(r)$). These measurements test theoretical models and predictions. The most important ones, classified by their interpretation are:

- ★ Density parameters (see 1.8 for definition of Ω_i):
  - ⋆ Ω_b: baryonic matter density parameter.
  - ⋆ Ω_m: total (baryonic and dark) matter density parameter.
  - ⋆ Ω_r: radiation density parameter, which includes Ω_γ - photons, and Ω_ν - relativistic neutrinos.
  - ⋆ Ω_Λ: dark energy density parameter.
  - ⋆ Ω_k: curvature parameter. Represents the spatial curvature of the Universe. $\Omega_k = 0$ corresponds to a flat Universe, $\Omega_k > 0$ implies a negative curvature, a hyperbolic geometry, and $\Omega_k < 0$ represents a positive curvature, which is a spherical geometry. In the ΛCDM model, $\Omega_k \equiv 0$. As a result, the total energy density satisfies $\Omega_T = 1$. This means that only the density parameters Ω_b, Ω_m, Ω_Λ, and Ω_r are needed to describe the energy content of the Universe. However, since $\Omega_T = 1$, only three of them are independent. For example, Ω_r can be expressed as:

$$\Omega_r = \Omega_T - (\Omega_b + \Omega_m + \Omega_\Lambda) = 1 - (\Omega_b + \Omega_m + \Omega_\Lambda). \quad (1.59)$$

- ★ Expansion rate parameters:
  - ⋆ H_0: the Hubble constant measured at the current stage ($z_0 = 0$). According to [18], $H_0 \approx 67.9 \frac{km}{sMpc}$
- ★ Primordial perturbations:
  - ⋆ A_s[2] is the amplitude of primordial scalar perturbations at a given scale k_*. It describes the overall strength of density fluctuations in the early universe.
  - ⋆ n_s is the spectral index. While A_s denotes the amplitude of primordial scalar perturbations, n_s describes how the amplitude changes with scale of the perturbations. $n_s \approx 1$ correspond to nearly scale-invariant fluctuations.

[2] Also $\Delta^2_{\mathcal{R}}(k)$, dimensionless power spectrum of primordial curvature perturbations, or related to $P_{\mathcal{R}}(k)$, primordial power spectrum, to which As is a normalization, as $P_{\mathcal{R}}(k) = A_s(\frac{k}{k_s})^{n_s - 1}$

  - ⋆ σ_8, a derived parameter, quantifies the the amplitude of matter fluctuations at scales of $8Mpc/h$ today. It is a normalization of the matter power spectrum representing the root mean square of such scales. They are derived from A_s after considering effects of structure formation and cosmic evolution.

- ★ Dark energy properties:
  - ⋆ w_0 is the parameter for dark energy in the equation of state. In the equation $w = p/\rho$, it represents the cosmological constant when $w = w_0 = -1$.
  - ⋆ w_a is a parameter that represents the time dependence of the dark energy equation of state.

- ★ Reionization:
  - ⋆ τ is the optical depth to reionization, the quantification of the cumulative effect due to scattering of the CMB (see 1.0.1) photons due to Thomson scattering along the line of sight (LOS). This measures the ionization of the neutral hydrogen in the intergalactic medium (see 3) after the first stars and galaxies. While τ is not a fundamental cosmological parameter, it is a critical nuisance parameter in CMB studies: its value affects the determination of A_s, and provides important information about the formation of the first stars and galaxies.

- ★ Early-Universe Derived Scales:
  - ⋆ r_d is the comoving sound horizon at the drag epoch, representing the maximum distance sound waves could travel in the photon–baryon plasma before baryons decoupled from radiation (see 1.5.5). It serves as a standard ruler in BAO measurements and depends on early-universe parameters such as Ω_b, Ω_m, and $H(z)$ before recombination. Though not a fundamental parameter, it is a key derived quantity in constraining cosmological models through LSS observations.

According to the results reported by Planck in 2018 ([18]), the values of these parameters are the ones shown in table 1.1.

1.5 Observational Probes

The widespread acceptance of the ΛCDM model is due to its ability to successfully explain numerous phenomena. While Λ explains the accelerated expansion of the Universe, CDM explains the observed kinematics in galaxies and clusters, which differ from those expected if only baryonic matter were considered.

As explained before, the Universe has been through a number of evolutionary stages. Each of them has been studied by diverse methods, and has provided probes that support the robustness of the ΛCDM model. Following the timeline of section 1.0.1, the most relevant probes are presented in the following subsections:

Parameter	**Planck (2018) cosmology** (TT, TE, EE + lowE + lensing)
$\Omega_m h^2$	0.14297
$+\Omega_c h^2$	0.12
$+\Omega_b h^2$	0.02237
$+\Omega_\nu h^2$	0.0006
h	0.6736
n_s	0.9649
$10^9 A_s$	2.100
Ω_m	0.31509
Ω_r	7.9638×10^{-5}
$\sigma_8(z=0)$	0.8119
r_d [Mpc]	147.09
r_d [h^{-1}Mpc]	99.08

Table 1.1: Cosmological parameters from Planck 2018. [18]

1.5.1 Cosmic Microwave Background

As explained in 1.0.1, the Cosmic Microwave Background Radiation (CMB) are the photons that decoupled from matter when e^- and existing nuclei combined to form the first neutral atoms. At that point, photons ceased to interact significantly with these particles and began to travel freely through space, permeating all the Universe. Nowadays, this first light is detected in the wavelength range of the microwaves, and provides very valuable information of the early Universe. It was predicted by George Gamow, Ralph Alpher, and Robert Herman in 1948, and first detected by Arno Penzias and Robert Wilson in 1964, which earned them the Nobel prize in Physics in 1978.

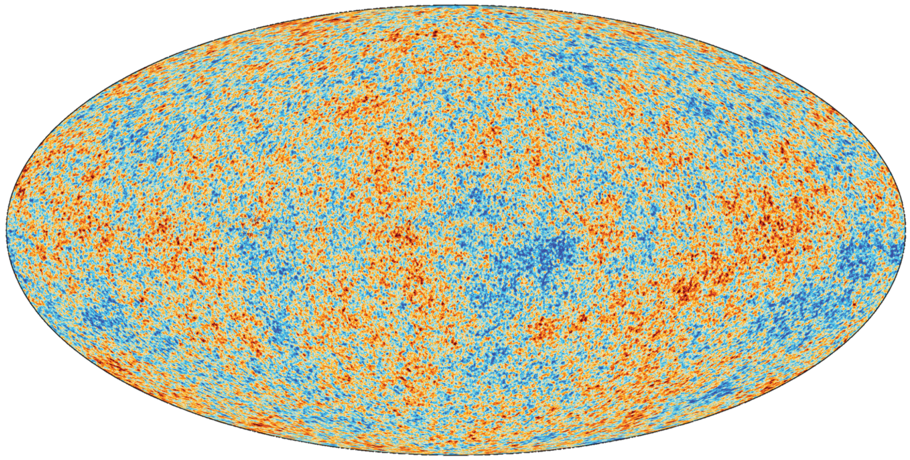
Derived from this event, a variety of experiments were built to analyze this radiation: ground-based, balloons, planes, and satellites. The most relevant are COBE (Cosmic Background Explorer[19]), which obtained the first all-sky map of the CMB, led by George Smoot and John Mather (who also got a Nobel prize for their results), BOOMERanG (Balloon Observations Of Millimetric Extra-galactic Radiation ANd Geophysics [20]), MAXIMA (Millimeter Anisotropy eXperiment Imaging Array [21]), ACBAR (Arcminute Cosmology Bolometer Array Receiver [22]), WMAP (Wilkinson Microwave Anisotropy Probe[23]), and *Planck* [24].

1.5.2 Large Scale Structure

The distribution of galaxies in the Universe shows a variety of structures, and contains information about the distribution of matter, the evolutionary history of the Universe and serves as powerful cosmological probe.

Following the Cosmological Principle, the Universe, at large scales, is considered to be homogeneous and isotropic. This is true at scales larger than 100Mpc/h. In smaller scales, we can find galaxies in different types of structures:

- ⋆ Groups: contain a few to tens of galaxies
- ⋆ Clusters: contain several hundreds to thousands of galaxies



no ref. en texto

Figure 1.4: Map of the Cosmic Microwave Background as seen by the Planck mission, as part of the results of Planck 2018[18]. Figure taken from [25]

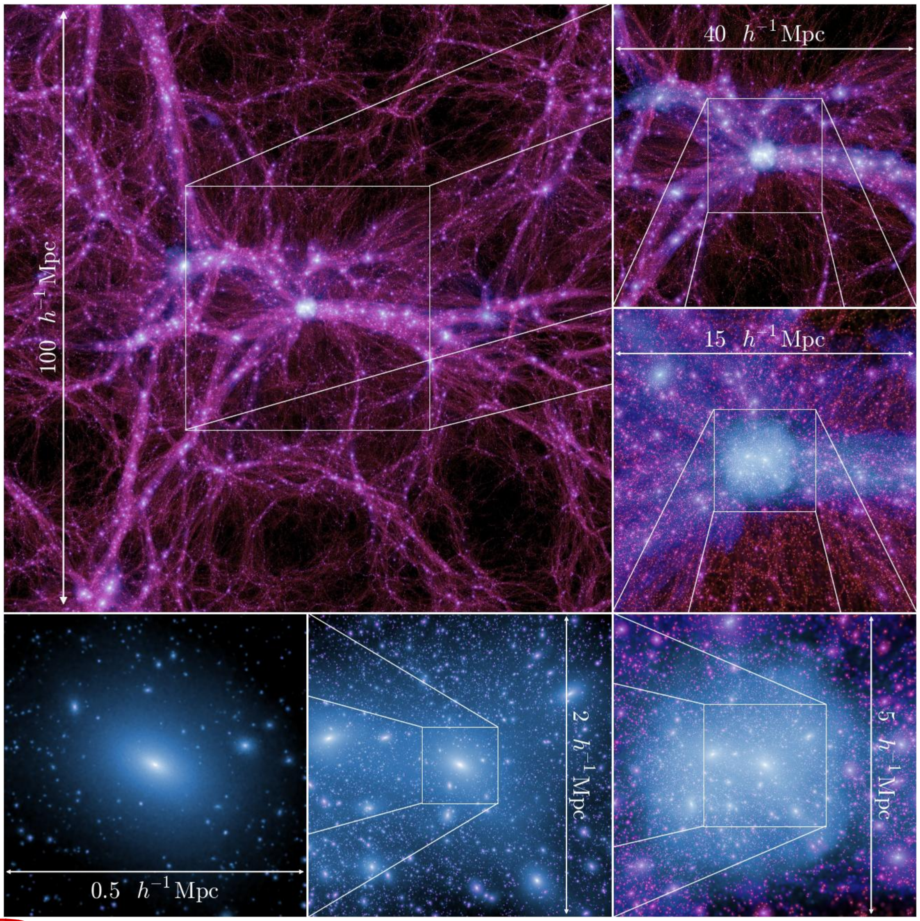
- ⋆ Filaments and walls/sheets: regions of clusters in thread-like structures that vary in thickness, with the thinner ones being filaments and the thicker ones being walls.
- ⋆ Voids: regions up to 100Mpc/h in diameter with very few to no galaxies surrounded by walls/sheets and filaments.

arriba section

At $z < 2$ a large percentage (considered to be $\approx 80 - 90\%$ [27]) of the baryonic matter is found as intergalactic gas. The reconstruction of the Cosmic Web requires mapping it, especially at redshifts above 2. The most abundant element conforming the IGM (see sec. 3) is HI, and it helps to make 3D maps of its distribution observing the absorption of light of background quasars. These are called tomographic studies, and there are numerous efforts dedicated to them. Among them, the COSMOS Lyman-Alpha Mapping And Tomography Observations (CLAMATO) survey [28], illustrated in figure 1.6 which aims to map the Cosmic Web at z=2.3, constructing three-dimensional maps of the IGM through Lyman-α forest tomography, extending beyond traditional H I Lyman-α forest analyses.

1.5.3 Gravitational Lensing

The light from distant sources is distorted by massive structures that lie in its path. These distortions behave analogously to those produced by optical lenses, and depending on the patterns made by such structures, their mass distribution can be reconstructed. This allows a better understanding of the nature of CDM. There are two types of gravitational lenses:
Strong lenses: as a result of the deflection of light by a massive foreground object, the shape of a background galaxy can appear significantly distorted. This often results in multiple images of the source, but strong lensing can also occur when only a single highly magnified and distorted image is observed — for example, an arc with no simple linear transformation from the original shape. An extreme case of this is an Einstein ring, in which the background source appears as *spread* in a circular or ring-like pattern due to perfect alignment between the source, lens, and observer, as in



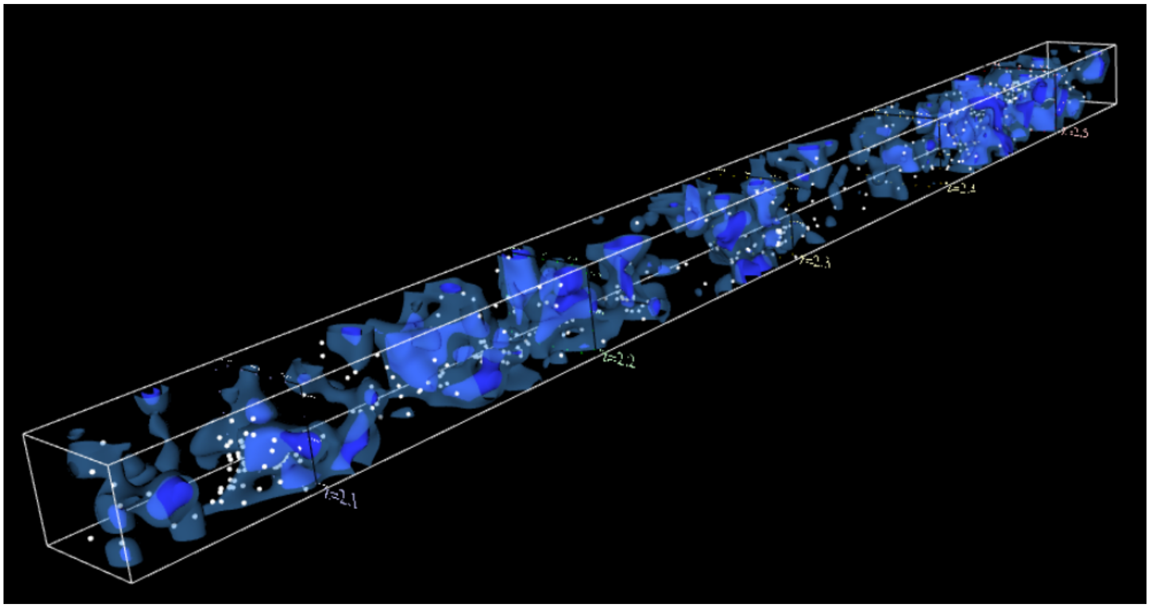
No descrita en texto

Figure 1.5: Zoom sequence from 100 Mpc/h to 5 Mpc_h of the Millenium-II Simulation, an N-body simulation of the dark matter in the Large Scale Structure, assuming a ΛCDM model. Figure taken from [26]

the illustration at the left in figure 1.7.
Weak lenses: the light of the source presents minor distortions, which are about 1%, as shown in the right illustration of figure 1.7. This can be measured statistically, examining correlations of shapes of galaxies measured in galaxy surveys. They provide a direct measurement of the distribution of mass. By means of different methods, the source can be reconstructed. [31]

1.5.4 Redshift Space Distorsions

Cosmic cartography is defined by three dimensions: right ascension (RA), declination (DEC, δ), and the third dimension is not a radial distance but given by redshift z. As seen in 1.2 these quantities are related by the Hubble expansion. The measurement of z has two components: cosmological redshift, caused by the expansion of the Universe, and an additional shift due to the peculiar



No descrita en texto

Figure 1.6: 3D visualization of the CLAMATO (COSMOS Lyman-Alpha Mapping And Tomography Observations)[29] tomographic map of the IGM COSMOS field, at $2.05 < z < 2.55$. The box is $20Mpc/h$ per side, the shades of blue correspond to densities of $\delta_F^{rec} = -0.08$ and $\delta_F^{rec} = -0.18$, and pale yellow dots are galaxy positions. This image is sourced from the interactive figure in [30]

velocity of the object - its motion relative to the Hubble flow[3]. These peculiar velocities arise from gravitational interactions with nearby structures and introduce Doppler shifts in the observed redshift. The peculiar velocity v is a tracer of the mass in the LSS [33]. Due to random motions on small scales, particles at the same distance have slightly different redshifts, which is observed as the elongation of structures along the line of sight (LOS), causing the so-called Fingers-of-God (FoG, see figure 1.8) effect, in which the structures seem to be pointing at the observer. Conversely, on large scales, objects fall towards overdense regions. This makes objects between the overdensity and the observer to look further away, while making objects on the other side of the overdensity look closer. This causes the overdensity to be magnified along the LOS. —a phenomenon known as the *Kaiser effect* or *Kaiser boost* [?]. These phenomena are collectively referred to as redshift-space distortions [34].

The recession velocity cz and distance d to a given galaxy are related by the Hubble law as:

$$cz \sim H_o d, \tag{1.60}$$

dos veces ya se definio como c=1

where c is the speed of light, and z is the redshift. We define the redshift distance s of a galaxy as $s = cz$, where z can be measured from its spectrum. The redshift distance s, and the true distance r differ, causing an apparent displacement along the line of sight in redshift space. As a consequence, the pattern in which galaxies cluster presents, in the redshift space, the so-called *redshift space distortions*, and this effect provides valuable information about large-scale motions of galaxies and structure growth.

[3] The smooth, large-scale recession of galaxies due to the expansion of the Universe, as described by Hubble law.

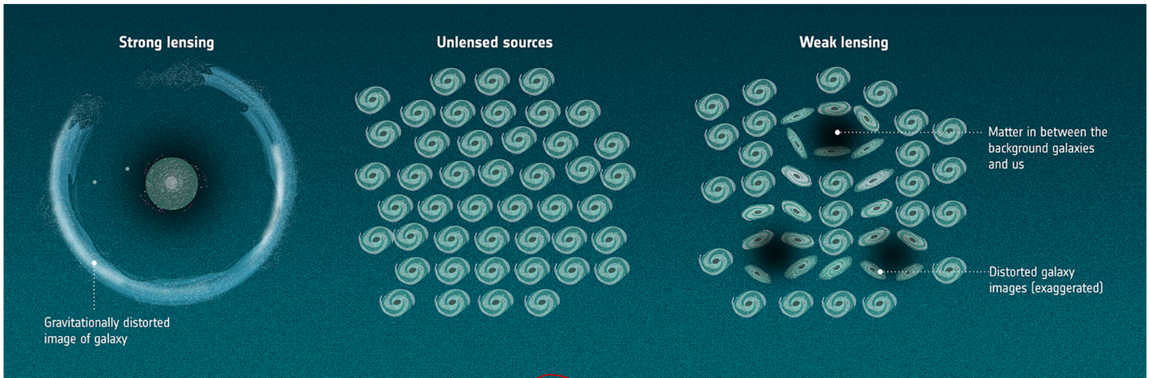


Figure 1.7: Representation of strong (left) and and weak (right) gravitational lensing. This figure illustrates the distortion of the shapes of the field galaxies shown in the center of the image by matter near the line of sight. Credits: ESA, [32].

We define the linear redshift distortion parameter β as the amplitude of coherent large-scale distortions. It is related to cosmology by:

$$\beta = \frac{f}{b}, \tag{1.61}$$

where b is the galaxy bias and f is the linear growth rate of structure. The latter is commonly approximated as $f = \Omega_m^\gamma$, with $\gamma \approx 0.55$ in General Relativity. Measuring β thus constrains both the matter density Ω_m and the bias b, and serves as a key probe of the growth of structure and gravitational dynamics [36][37].

1.5.5 Baryon Acoustic Oscillations

Before the recombination epoch, protons, electrons, and photons comprised a coupled fluid, forming a plasma where the radiation pressure of photons counteracted the attempt of gravity to collapse into dark matter potential wells. This interplay between radiation and matter caused oscillations in the plasma.

As the Universe expanded and the temperature decreased, photons ceased to interact with electrons. This led to the formation of neutral hydrogen, making the Universe transparent to photons. These photons freely traveled through the Universe, creating the Cosmic Microwave Background (CMB). In configuration space, a spherical sound wave propagated through the photon-baryon plasma in the early Universe. When this wave *froze* at the drag epoch, it left a characteristic scale —known as the sound horizon— imprinted in the spatial distribution of baryons. Gravitational interactions then drove the formation of structures, evolving into the Large Scale Structure (LSS, see 1.5.2) pattern. This frozen oscillation generated the Baryon Acoustic Oscillations (BAO). This frozen oscillation kept growing as a consequence of the expansion of the Universe only, and eventually generated patterns in the distribution of objects (galaxies, quasars, IGM clouds, etc.) in the Universe. This imprint is observed as a pronounced peak (the BAO peak) in the two-point correlation function (see 1.4.1) and the anisotropy spectrum of the CMB, as schematically shown in the left panel of Figure 1.9.[38]

a veces mayusla
otras min.

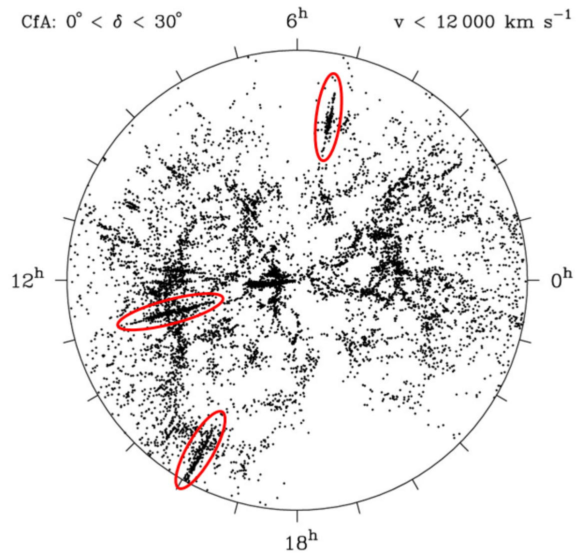


Figure 1.8: Due to random motions on small scales, particles at the same distance have slightly different redshifts, which is observed as the elongation of structures along the line of sight, causing the so-called Fingers-of-God (FoG) effect shown in red, in which the structures seem to be pointing at the observer. [35]

Figure 1.10 from [39] illustrates the stages of this process, as a function of redshift z, in which BAO were formed. Initially, electrons and protons were coupled in dark matter gravitational potential wells. The second panel depicts the coupled plasma being influenced by the radiation pressure of photons. When conditions allowed it, photons decoupled from the plasma, leaving an imprint of baryons and traveling freely through the Universe.

The position of the BAO peak, centered on a comoving distance r_d (sound horizon at the drag epoch, $r_d \approx 100h^{-1}$Mpc [42]), determines the ratio $D_M(z)/r_d$, where $D_M(z) = (1+z)D_A(z)$ is the comoving angular distance, and $D_H(z)/r_d$ (where the Hubble distance is $D_H(z) = c/H(z)$ at a given observed redshift). These ratios depend on cosmological parameters, constraining them through BAO peak observations. This frozen scale serves as a *standard ruler*, first measured by Eisenstein et al. [43] and Cole et al. [44], allowing distance measurements in the LSS clustering at different evolutionary stages of the Universe as a function of z, as shown in the right panel of Figure 1.9 [45] [46].

1.5.6 Type Ia Supernovae

Numerous efforts have been dedicated to understand the nature of the expanding Universe. To achieve this, precise methods for measuring distances are essential. Among these methods, one of the most important is the use of standard candles. The first objects used for this purpose were the type Ia supernovae (SN Ia), which consist of stellar explosions that follow the same pattern of formation and death, and all have approximately the same brightness. The results obtained demonstrated that the Universe expands at an accelerated rate. This earned a Nobel prize to Saul

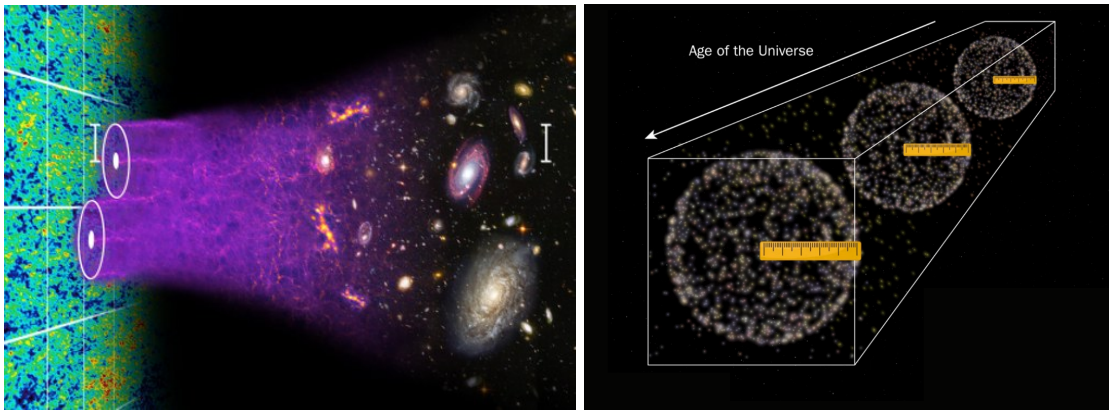


Figure 1.9: Left panel: Schematical evolution of CMB anisotropies into the distribution of matter imprinted in the LSS as BAO. Figure taken from [40] . Right panel: Representation of the Baryon Acoustic Oscillations as tracers of the expansion of the Universe. As the Universe ages, it expands, and so does the BAO scale. This makes it a remarkably useful tracer of the Dark Energy. Sourced from [41]. Credit: Claire Lamman/DESI collaboration.

Perlmutter, Brian Schmidt and Adam Reiss.

SNe Ia arise from binary star systems involving a white dwarf and another accompanying star. When the companion star is attracted to the white dwarf, it begins to accumulate mass until it exceeds a limit (the Chandrasekhar limit), triggering nuclear processes that enable the burning of heavy elements. This causes the white dwarf to explode, shining as brightly as the center of an average spiral galaxy. Since all SNe Ia are generated in a very similar process, they all have a similar brightness. Their light curves exhibit similar shapes, and by applying an empirical relation between their color and the stretch of their light curve, type SNe Ia can be standardized and used as precise distance indicators. When we compute their distance modulus using their apparent and absolute magnitudes, we can estimate their luminosity distance. They provide information and constraints of cosmological parameters at $0.001 < z < 1.5$. The most recent SN Ia experiments are the SCP (Supernova Cosmology Project [47]), and Pantheon sample [48].

1.5.7 Galaxy and Galaxy Cluster Kinematics

Rotation Curves of Spiral Galaxies:

The stars and cold gas in spiral galaxy disks move in circular orbits in the disk plane. The disk kinematics is described by their rotation velocity as a function of the galactocentric distance R rotation curves, $V_{rot}(R)$. These rotation curves present a flat behavior in the outer regions of the galaxies, meaning that their outer regions remain constant as R increases, suggesting the presence of a dark matter halo extending beyond the baryonic contents of the galaxy.

Galaxy Clusters:

Historically, galaxy clustering has been used as a cosmological probe, providing the first evidence of the existence of dark matter. Today they represent several cosmological tests. Among these, the measurement of BAO (2.3.1) and RSD (1.5.4), $P(k)$ across a range of scales, allowing the estimation of the deviations of a primordial power spectrum. Aside from the aforementioned, when comparing with predictions from ΛCDM simulations, galaxy clusters provide insights into the effect of dark matter, dark energy, and their connection[4] [16], [49]. Furthermore, gravitational lensing caused by galaxies and galaxy clusters directly measures the mass distribution of both: baryonic and dark matter. Observations of this effect test the ΛCDM model predicted relationships between the kinematics of galaxies and the mass distribution observed through gravitational lensing. Deviations from the observations reported by SNIa and BAO are also direct evidence of modified gravity as a cause of the accelerated expansion of the universe, hence galaxy clustering also tests constraints to the cosmological model.

Even though all the probes provide valuable information and complement each other, it is important to highlight that the most important observation for the development of this research work are the baryon acoustic oscillations.

[4]the connection between galaxies and halos in order to do cosmology with galaxies.

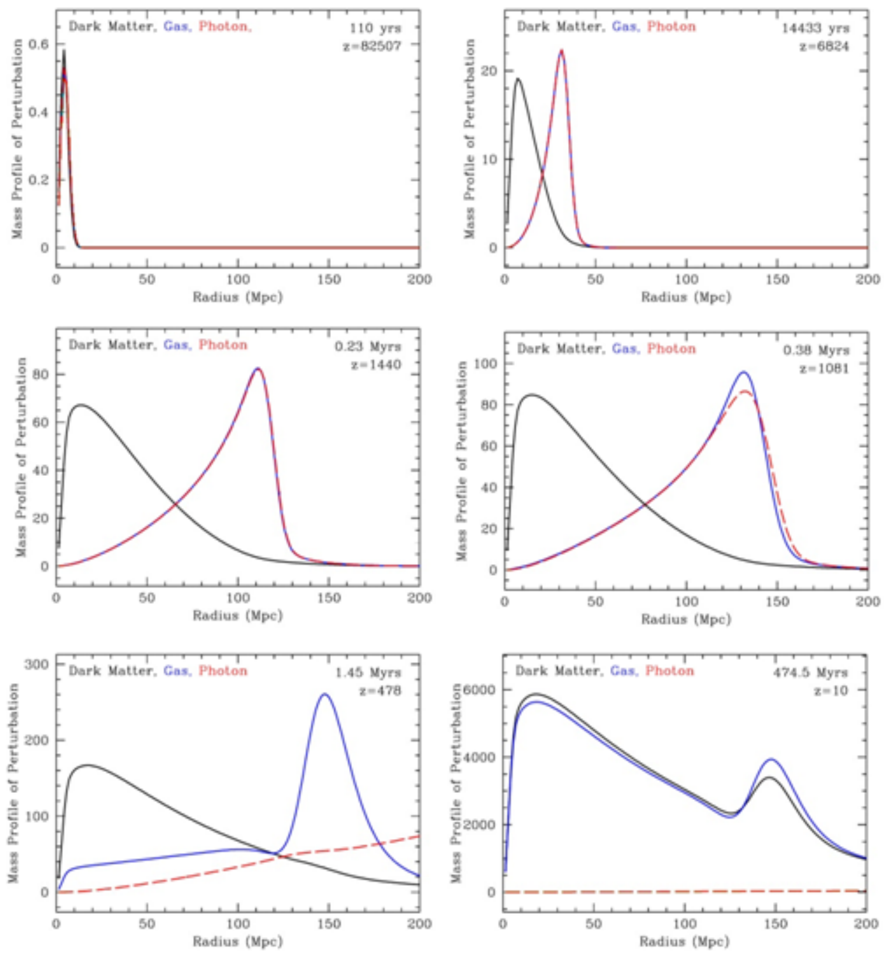


Figure 1.10: Snapshots of the process in which the BAO were formed in different epochs (redshifts z). The black lines represent dark matter, the blue ones represent baryons, and the red represent photons. The top left panel shows coupled photons and baryonic matter in the dark matter potential well, the top right panel shows how in the last oscillation of the coupled plasma photons started dragging the baryonic matter through the middle left panel. The middle right panel shows photons decoupling from this plasma as the CMB leaving a *fixed* BAO scale. The bottom left panel shows fully decoupled photons, and baryons in the fixed BAO scale. The right bottom panel shows how the interaction of dark and baryonic matter caused their distribution, observed by tracers (such as galaxies and IGM) nowadays. Figure taken from [39]

Chapter 2

The Dark Energy Spectroscopic Instrument (DESI)

A *redshift survey* is a census of a section of the celestial vault focused on measuring the redshift of astronomical objects, such as galaxies, galaxy clusters, AGN, quasars, and others. Using a cosmological model, the redshifts are used to estimate a the distance to such objects. Combining this information with their angular position, the surveys map the distribution of matter at different epochs of the Universe, providing valuable information of its evolution and thereof validating or falsifying our modeling and gravity theories.

The development of this thesis project has temporarily coincided with the transition of two such surveys: in one hand, it has been part of the analysis of the Final Data Release of the Extended Baryon Oscillation Spectroscopic Survey (eBOSS), and in the other hand, the preparation for the Dark Energy Spectroscopic Instrument (DESI), the start of the survey, its first data and public data releases. This chapter briefly describes the evolution and timeline of galaxy surveys, especially the aforementioned, and focuses on the DESI survey.

Timeline of Redshift Surveys

The LSS of the Universe arises from gravitational interactions and reflects the distribution of both baryonic and dark matter. Observational studies of the LSS provide key insights into cosmic evolution and the accelerated expansion of the Universe. Two fundamental probes derived from galaxy redshift surveys are Baryon Acoustic Oscillations (BAO), used as a standard ruler to measure the expansion history, and Redshift-Space Distortions (RSD), which probe the growth of structure and test gravity on large scales. Resolving these features requires high-precision spectroscopic redshift measurements, prompting the development of numerous redshift surveys with varied strategies and wavelength coverage tailored to specific scientific goals.

The first large spectroscopic galaxy sample, the Harvard-Smithsonian Center for Astrophysics (CfA) [50] in the 1970s, consisted of several thousand galaxies for clustering analysis. Since then, and with the development of instruments such as spectrographs and CCDs, several redshift surveys have been designed and conducted, aiming to map a larger amount of objects in a wider and deeper sample of the Universe. Among these:

- ⋆ CfA2 Redshift Survey [50] (1999): Completed 18,000 redshifts over 17,000 deg^2
- ⋆ 2dF Galaxy Redshift Survey (2dFGRS) [51] (2001, 2003 for final release): Completed 220,000 redshifts over 1500 deg^2
- ⋆ 6dF Galaxy Survey (6dFGS) [52] (2009): Completed 125,000 redshifts over 17,000 deg^2
- ⋆ Baryon Oscillation Spectroscopic Survey (BOSS) [53](2002-2014): Completed 1,660,000 redshifts over 10000 deg^2
- ⋆ Extended BOSS (eBOSS) [54] (2014 – 2019): Completed 975,000 redshifts over 6000? deg^2
- ⋆ Dark Energy Spectroscopic Instrument (DESI)[5][6]: (2019 - ongoing): Aims to measure 40,000,000 spectra over 14,000 deg^2.[55][56]

Figure 2.1, shows the area (in deg^2) covered and the number of spectra per deg^2 observed by the main spectroscopic surveys. While this plot does not explicitly show a timeline, it reflects how spectroscopic galaxy surveys have progressively pushed the boundaries of both area and spectral density, culminating in projects like DESI, currently the largest ground-based spectroscopic survey in terms of both sky coverage and redshift depth. This progression is the result of steady technical and scientific improvements over time.

2.1 The DESI Survey

Since May 2021, the Dark Energy Spectroscopic Instrument (DESI) is conducting a five-year spectroscopic survey aimed to be "the most precise measurement of the expansion of the Universe ever obtained" [[58], [59]]. DESI is a spectrograph that operates robotically a carefully choreographed arrangement of 5020 optical fibers in the focal plane of the 4m Mayall telescope, capturing up to 5000 spectra simultaneously across a wavelength range extending from 360 nm to 980 nm. The fibers feed into three-arm spectrographs with a resolution $R = \lambda/\Delta\lambda$ ranging from 2000 to 5500. It is designed to be a Stage IV [1] dark energy survey during its five years of operations, obtaining spectra of over 40 million galaxies and quasars, mapping over 14,000 deg^2 from nearby stellar objects to redshifts of $z > 3.5$. The expansion history or distance-redshift relationship of the Universe at this redshift range will be measured using the standard ruler technique with BAO, and several surveys with successively increasing sample sizes have followed their lead measuring it at different redshifts. DESI will also measure RSD which will give valuable information about the growth of structure from peculiar velocities of galaxies.

To achieve these goals, it observes seven different classes of targets, selected based on the Legacy Surveys [61], [62]:

[1] Stage IV refers to the classification system proposed by the Dark Energy Task Force (DETF)[60], which ranks surveys according to their scale and precision. Stage IV surveys represent the most advanced generation, aiming to deliver an order-of-magnitude improvement in constraints on dark energy compared to previous efforts. They typically cover large areas of the sky, use multiple cosmological probes (such as BAO, RSD, and weak lensing), and are designed to test the expansion history and structure growth of the Universe with high precision. DESI meets these criteria by targeting an unprecedented volume of the Universe with spectroscopic redshift measurements.

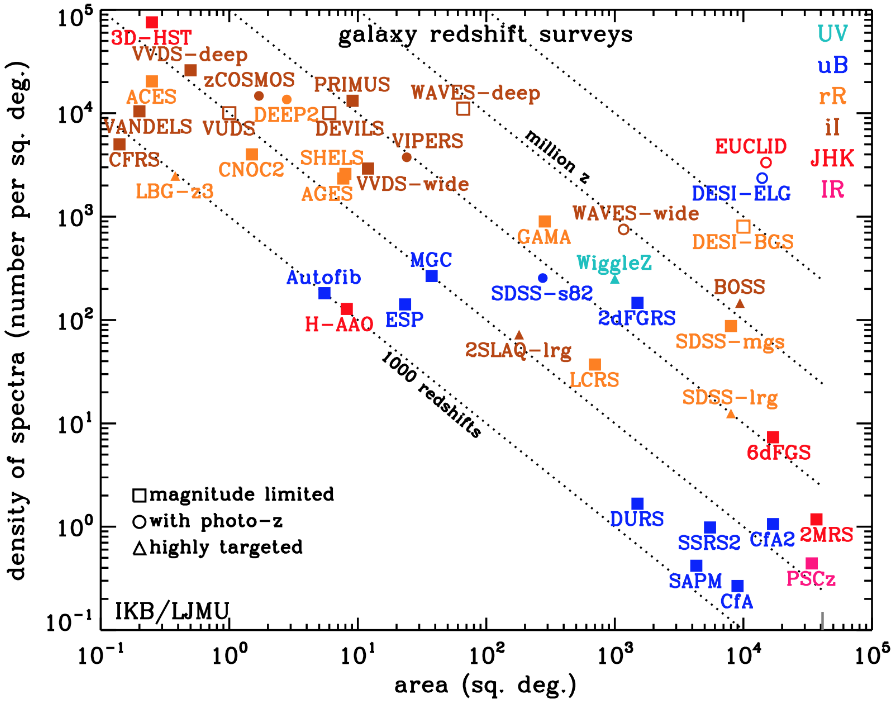


Figure 2.1: Area (in deg^2) covered and the number of spectra observed per deg^2 by the main spectroscopic surveys, comparable with DESI (ELG and BGS surveys within the DESI survey, explained in 2.1). Squares represent magnitude-limited surveys, circles are surveys involving color cuts for photometric redshift selection, triangles are highly targeted surveys. The filled symbols represent completed surveys. The color code refers to the wavelength selection of each survey. For more information see [57].

- ⋆ Stars: sample of more than ten million stars in galactic latitudes of $b \pm 20$ of up to a magnitude of 19 visible with exposures of 8-10 min with a $S/N = 25$ that will provide valuable spectra with information to probe the ΛCDM model at low scales, by the radial velocity, effective temperature, surface gravity, chemical abundance distribution, and approximate age of different stellar populations mapping the spacial distribution, kinematics, and chemical composition of the sample. These spectra will enrich the dataset obtained by other surveys like GAIA [63].

- ⋆ Bright sample of low redshift galaxies (BGS - Bright Galaxy Sample): magnitude limited sample of more than ten million galaxies in a redhsift range between $0 < z < 0.6$ (although the best BAO and RSD measurements will be in the range of $0 < z < 0.4$), observed during the time when the sky is too bright and hampers the observation of fainter objects. It will contain quiescent and star-forming galaxies [64].

- ⋆ Luminous red galaxies (LRG): sample of galaxies that present high luminosities and are red in the rest-frame optical wavelengths. They have high stellar masses and are not forming stars. They also present high clustering, which makes them a good BAO tracer. The sample is constituted by approximately 8 million galaxies in a redshift range of $0.4 < z < 1$ [65].

- ⋆ Emission Line Galaxies (ELG): sample of approximately 17 million luminous star-forming galaxies, typically late-type spiral and irregular (although ELG refers to galaxies actively forming stars at a sufficiently high rate) at redshifts $0.6 < z < 1.6$ (where the tightest constraints are expected to be in the redshift range of $1.1 < z < 1.6$) whose spectra present the [O II] emission doublet (rest-frame wavelengths of $3726\mathring{A}$ and $3729\mathring{A}$) with $S/N \approx 7$, and strong emission lines from ionized H II regions surrounding short-lived but luminous, massive stars. This survey is the largest sample of DESI targets. Due to the active star formation, they tend to be bluer than galaxies with older stellar populations such as LRGs [66].

- ⋆ Quasars (QSO) (also known as clustering quasars, or direct tracers of dark matter): sample of the most luminous extragalactic sources known, reaching 1000 times the brightness of a normal galaxy. Their luminosities are a consequence of the accretion of supermassive black holes in the center of galaxies, and hence associated with Active Galactic Nuclei (3.1.1). These objects are bright enough to outshine their host galaxy. Due to the size of their emitting regions they are observed as point sources in images. They are used as point tracers of matter clustering, and represent the targets at highest redshift observed by DESI, at $0.9 < z < 2.1$ [67] [68].

- ⋆ Lyman-alpha quasars (or absorption quasars): 8 million quasars at redshifts above 2.1 used to probe the intergalactic medium using the neutral hydrogen absorptions at $\lambda = 1216\mathring{A}$ that builds the Lyman-α forest (described in more detail in 3.2.1). These are observed multiple times to improve the measurements of the Lyman-α forest with the required S/N.

These objects will be observed in the footprint shown in figure 2.2.

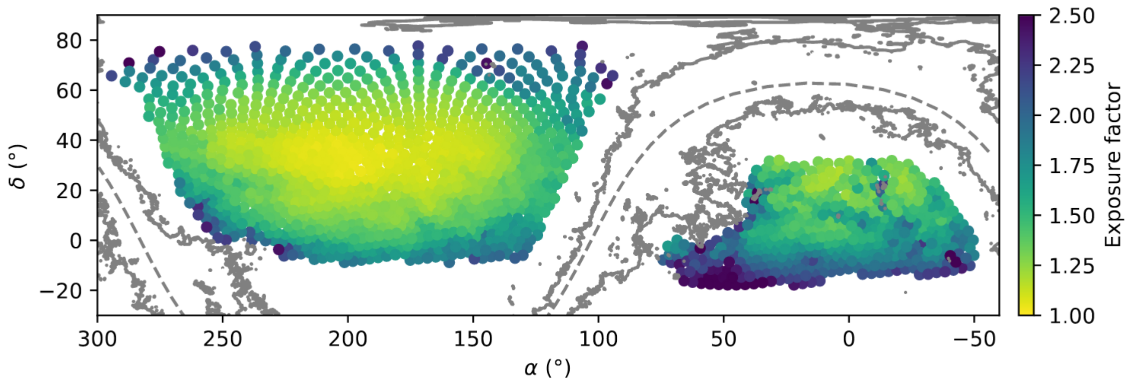
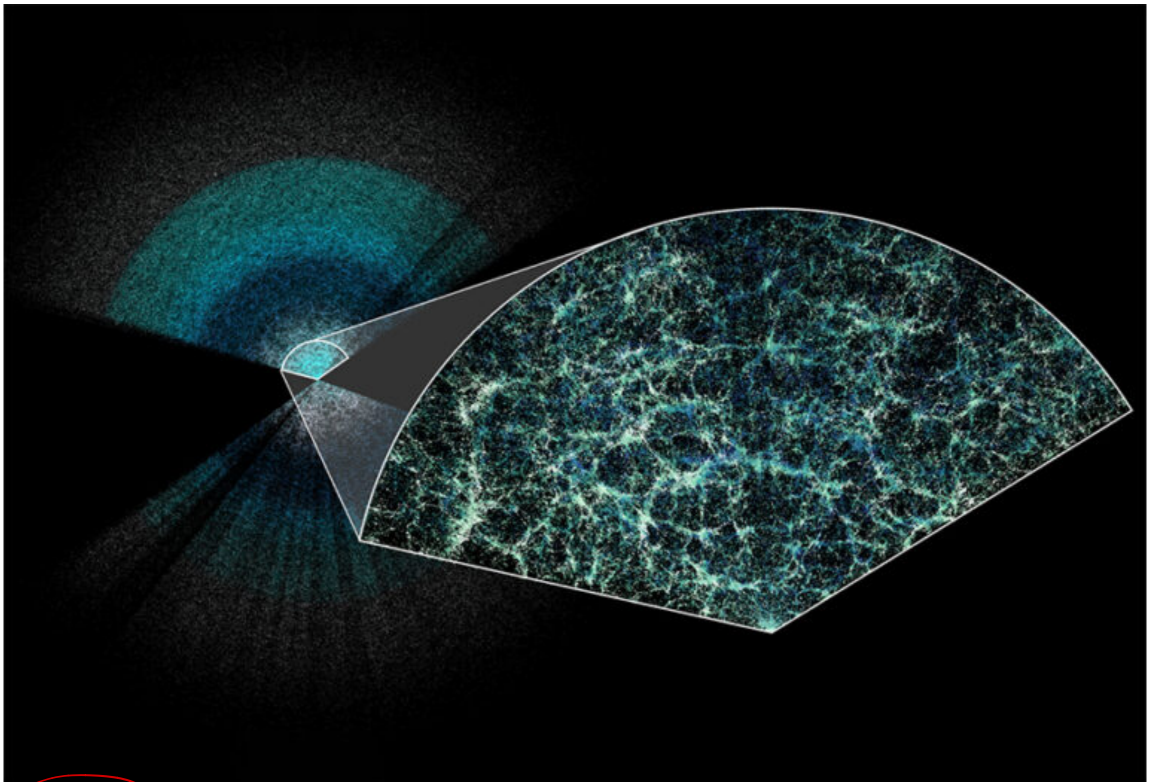


Figure 2.2: Forecasted footprint of the Dark Energy Spectroscopic Instrument (DESI) survey. The color map represents the time required to reach a specific galaxy depth, relative to the observation at zenith without galactic extinction. The gray dashed line indicates the Galactic plane, while the gray contour is the $E(B-V) = 0.3mag$. [69]

According to [71], $\approx$ 70% of the main DESI survey targets are identified as quasars, $\approx$ 16% as galaxies, $\approx$ 8% stars, and $\approx$ 8% are inconclusive spectra due to low quality and the lack of identifiable reliable features. Achieving the goal of observing this extensive sample of objects



mencionar texto

Figure 2.3: Tracers observed by DESI and building the cosmic web, closed-up in a circular wedge. Distribution of tracers, from the core (observer) to the outer edges: Bright Galaxy Sample (BGS), Luminous Red Galaxies (LRGs), Emission Line Galaxies (ELGs), Quasars, and absorber quasars (Lya forest quasars). Credit: Claire Lamman, DESI collaboration [70].

within a five-year time-frame requires a significantly high survey speed (number of measurements per time unit). The parameters involved in the realization of this objective are the exposure time of each observation, time between consecutive exposures and number of objects observed in each exposure. The design and construction of DESI allows the instrument to [58]:

1. Take science exposures of 1000 seconds in nominal conditions (although the longest exposures expected are of 1200 seconds).

2. Obtain 5000 spectra during each exposure. This is made possible by the 5020 robotic positioners of the focal plane, each equipped with an optical fiber, arranged in ten petals, with 500 positioners allocated to each petal.

3. Change between fields in ≈ 2 min, thanks to the stability of the system and the robotic positioner system that controls the optical fibers.

This instrument is installed in the Nicholas U. Mayall 4-meter Telescope, at I'olkam Du'ag (Manzanita Bush Mountain) in the Tohono O'odham Nation, also known as Kitt Peak, Arizona.

It is operated by the National Optical-Infrared Astronomy Research Laboratory (NOIRLab) of the National Science Foundation (NSF) of the USA.

2.2 Instrument facilities

The Mayall, the largest of the 22 telescopes located at the Kitt Peak National Observatory (KPNO) and commissioned fifty years ago, is a reflector telescope with a 4m diameter primary mirror, positioned on an equatorial mount. It was reconditioned for the installation of the DESI instrument (February 2018) and saw its new first light in October 28, 2019. Figure 2.4 shows a scheme of the facilities of the 4 m Mayall telescope and the instrumental arrangement built for the DESI instrument. We can identify six main structures that conform it: corrector, focal plane, fibers system, spectrographs, instrument control system, and data systems.

* Corrector: its function is to provide a larger field of view to the focal plane, allowing an array of 5000 optical fibers in 8 deg^2. It consists of 4 lenses made of fused silica. The two smaller ones are single-surface spherical, while the others are spherical. They were designed to meet the needs of the DESI lenses. Additionally, they have an atmospheric dispersion compensator to calibrate blurry areas away from the zenith.

* Focal plane: It is located at the top of the telescope and has 5000 robotic positioners, each equipped with an optical fiber. These positioners adjust to direct towards individual sources, ensuring that the fibers collect their respective light, with the additional challenge of avoiding collisions between them. With an approximate diameter of 1 meter, the focal plane is composed of ten petals, each with 500 positioners. Additionally, the focal plane includes sensors and lights to ensure the proper alignment of the positioners.

* Fiber system: The fibers, each 47.5 meters long, carry the light from the 5000 collected targets at the focus to the spectrographs. Each fiber measures $107\mu m$ in diameter (roughly the thickness of a human hair) and each is mounted on one of the positioners. Afterwards, they are grouped into sets of 500.

* Spectrographs: each of the ten spectrographs is fed by a cable of 500 fibers. The light from the fibers is separated into three cameras that are each sensitive to different wavelength ranges and have different resolutions ($R = \lambda/\Delta\lambda$):

    · $360 < \lambda \leq 555\mathrm{nm}; \mathrm{R} = 2000 - 3200$

    · $555 < \lambda \leq 656\mathrm{nm}; \mathrm{R} = 3200 - 4100$

    · $656 < \lambda \leq 980\mathrm{nm}, \mathrm{R} = 4100 - 5000$

  The detectors are arrays of 4096 x 4096 pixels of 15μm, each in its own cryostat which is at very low and stable temperatures.

* Instrument control system: controls and monitors all functions required for the operation of DESI. From the data acquisition, connection to the Mayall Telescope, monitoring infrastructure, maintenance of a database of operations, guiding, coordinate transformations, subsystem coordination and user interfaces. It also has the DESI Online System, that contains the user interfaces for exposure sequences, real-time telemetry, and image previews.

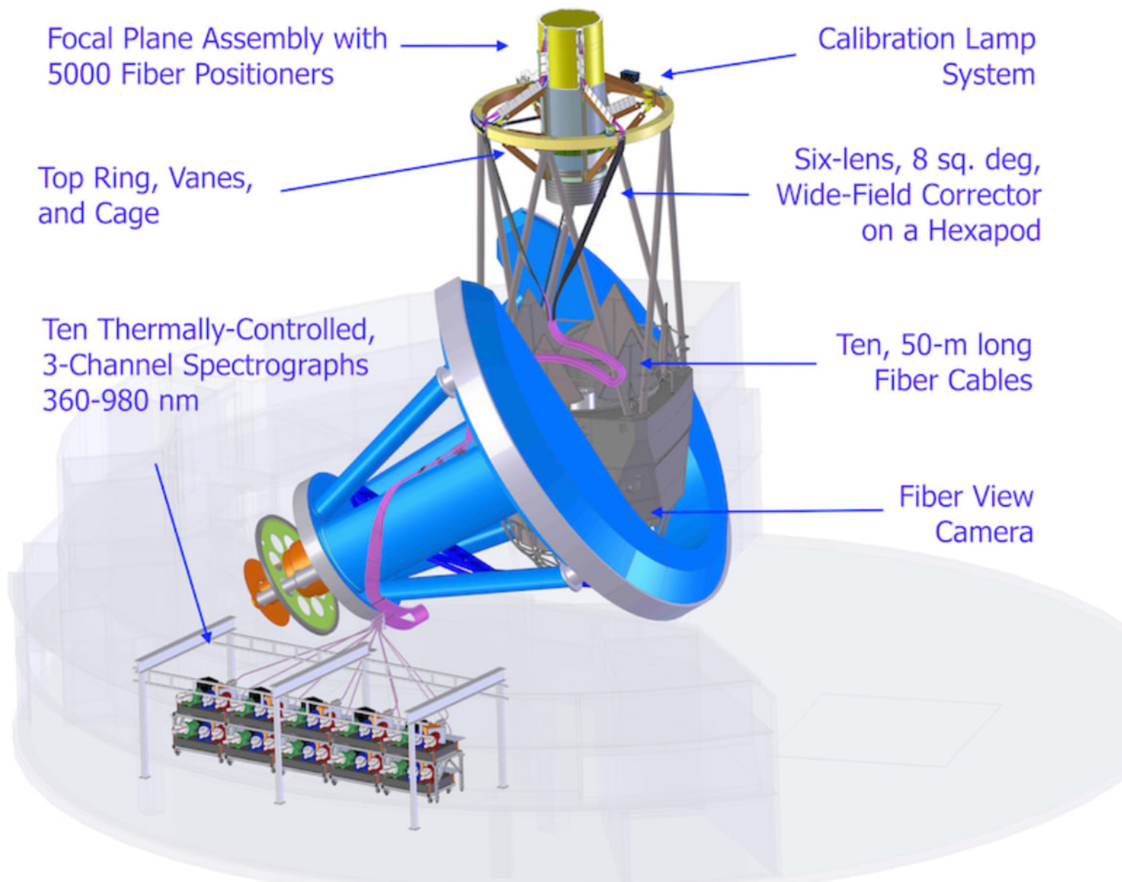


Figure 2.4: Scheme of the Dark Energy Spectroscopic Instrument in the Nicholas U. Mayall 4-meter Telescope, at I'olkam Du'ag (Manzanita Bush Mountain) in the Tohono O'odham Nation, also known as Kitt Peak, Arizona, with its main components (described in this chapter: corrector, the structure that supports it, focal plane with positioners, fiber system and cables, and spectrographs) [72].

∗ Data system: the data are monitored by the observers within the first three minutes after being obtained. Then the raw data are transferred to the National Energy Research Scientific Computing Center (NERSC) at Berkeley Lab for processing. Then the processed spectra are extracted, calibrated, classified, and their redshift measured. 10 TB are expected to be transferred from KPNO to NERSC per year, as well as stored at NOIRLab in Tucson, Arizona. After the data have been processed at NERSC investing millions of CPU hours, $\approx$ 100 TB will be available per year, containing the contents of the data releases, expected to be within a yearly cadence throughout the five years of DESI operations.

2.3 Scientific Goals of the DESI Survey

The scientific goals of DESI are focused on advancing our understanding of dark energy, galaxy formation, and the LSS. DESI aims to map the distribution of millions of galaxies and quasars across a wide redshift range, enabling precise measurements of the expansion history of the Universe. One of its key goals is to constrain the properties of dark energy by studying the expansion rate at different cosmic epochs and the growth of cosmic structures. An example of this is shown in figure 2.5, which shows the comparison of $H(z)/(1+z)$ from BOSS/eBOSS and DESI [5]. Section **??** mentions that dark energy, represented by Λ, is described by three possible scenarios: a cosmological constant, equivalent to static dark energy with $w = -1$; a dynamical dark energy with $w(a) \neq -1$; or a failure of General Relativity. DESI is designed to address this fundamental question about the nature of the Universe.

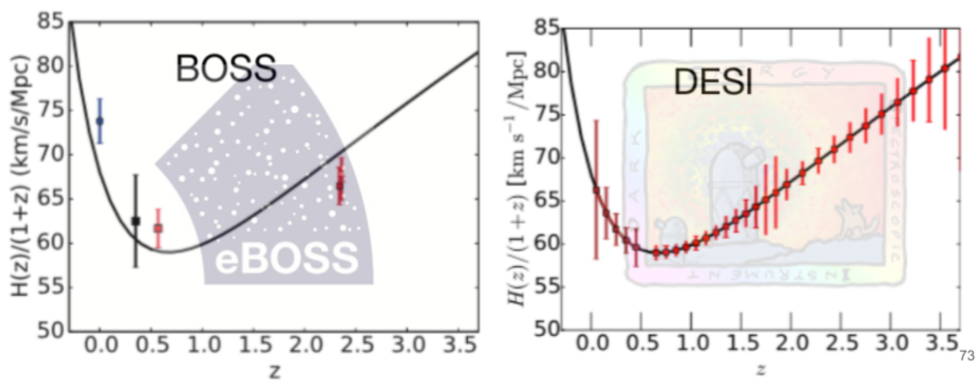


Figure 2.5: Comparison of $H(z)/(1+z)$ measurements from BOSS/eBOSS (left) and DESI (right). The figure shows the Hubble parameter evolution as a function of redshift, with data points representing BAO measurements from different surveys. The DESI results (right) are expected to provide significantly improved precision, particularly at higher redshifts, highlighting its capability to map the expansion history of the Universe with unprecedented accuracy.[5]

By probing the Universe with unprecedented precision, DESI will test cosmological models that seek to explain the nature of dark energy. A primary approach involves measuring BAO (see 2.3.1) and RSD 1.5.4. BAO provides a "standard ruler" for gauging cosmic distances, while RSD provides insight into the growth of cosmic structures. This combined analysis allows DESI to assess the rate of structure formation, which is sensitive to both dark energy and the properties of dark matter.

In addition to its primary focus on dark energy, DESI will also contribute to our understanding of galaxy evolution, by providing a vast catalog of galaxy redshifts that will allow researchers to study how galaxies have evolved in various environments over cosmic time. Furthermore, it will offer new insights into the distribution of matter in the Universe, including constraints on neutrino masses and tests of modifications to general relativity. By linking theory with observations, the rich dataset obtained by DESI will serve as a valuable resource for cosmological studies, enhancing our understanding of fundamental physics.

2.3.1 Baryon Acoustic Oscillations

In the context of DESI, BAO play a central role in addressing key scientific questions about the expansion of the Universe and the nature of dark energy. DESI aims to measure the BAO signal with unprecedented precision, across a wide redshift range, providing a precise standard ruler for measuring cosmic distances and improving constraints on the expansion history of the Universe.

One of the primary goals is to refine our understanding of the accelerated expansion of the Universe due to dark energy. By using the BAO signal at different redshifts, DESI can measure the expansion rate at various epochs, providing essential data to model the evolution of dark energy. BAO also allows DESI to independently test the ΛCDM model and investigating new physics, offering a precise way to measure the Hubble constant and compare it with results from other cosmological probes, such as the CMB.

Furthermore, these measurements can be used to constrain neutrino masses and the potential effects of modified gravity. Neutrinos, with their unique properties, leave subtle imprints on the LSS of the Universe. By comparing BAO signals across time, DESI can provide new insights into the role neutrinos play in the cosmos.

2.3.2 Redshift Space Distorsions

Redshift Space Distortions (RSD) are crucial to understand how galaxies move under the influence of gravity. DESI measures peculiar velocities caused by gravitational attraction, offering direct insights into the growth of cosmic structures. One of the main scientific goals of studying RSD with DESI is to probe gravity on cosmological scales, testing predictions from General Relativity and exploring potential modifications.

RSD measurements also help distinguish between different models of dark energy by examining how the large-scale structure of the universe grows over time. Dark energy influences this growth by opposing gravitational collapse, so tracking the rate of structure formation provides an independent test of dark energy models.

In addition to probing gravity and dark energy, the RSD analysis of DESI will improve the understanding of galaxy formation and evolution. By disentangling the effects of large-scale clustering from those of small-scale, nonlinear motions, RSD measurements help reveal how galaxies move and form in different cosmic environments.

Finally, RSD data will complement other cosmological probes, like Baryon Acoustic Oscillations (BAO), offering a cross-check on the expansion history and improving precision in measuring the universe's fundamental parameters. Together, RSD and BAO measurements will provide a powerful tool for understanding the interplay between gravity, dark matter, and dark energy across cosmic time.

2.4 DESI Y1

2.4.1 Data

The data used for the cosmological analysis of this research is DESI DR1, the first data release of this survey. In June 2021, DESI officially started mapping the 14,000 deg^2 of the celestial vault, in the footprint shown in 2.2, which is planned to be covered in a five year time lapse. The DR1 dataset is composed of other sub-datasets: commissioning and Survey Validation, which are all data taken

in what?

between December 14, 2020 - June 10, 2021, used for the validation of the survey design and strategy for the DESI Main Survey, that started in May 14, 2021. The first year of the five that conform the DESI operations concluded in June 13, 2022. The data were processed by the data reduction pipeline and the redshift classification algorithms in [73]. DR1 is composed of the following amount of spectra: 6,279,965 galaxies conforming the BGS, 3,926,272 ELG, 2,829,517 LRG, 1,340,073 QSO, and 3,641,276 stars of the MWS. They have been observed accross the footprint shown in figure 2.6, which shows the completeness of the DESI Main Survey by June 13, 2022 for the dark, bright, and backup programs [74]. The second and third data releases of DESI will correspond to the third and fifth years of operations respectively. The DESI collaboration is currently analyzing the data of the third year of operations and extracting cosmological information from them, while observers at the Mayall are already obtaining the dataset that will conform the third and final data release.

Quasar catalog

The different Lyman-α analyses require a refined sample of quasars with very precisely determined redshifts. For this, DESI follows a thorough procedure to identify quasar spectra and determine such feature with the requirements met. This procedure is thoroughly detailed in Section 2.1 of [75].
The quasar catalog was made through three classifiers: RedRock [76], a template fitting code that uses Principal Component Analysis using templates for galaxies, stars and quasars in a wide range of redshifts, QuasarNet [77], using a deep convolusional neural network that identifies potential quasar emissions and estimates the redshift, and finally a second iteration of RedRock for targets confirmed as quasars. This catalog is part of the first DESI Data Release (DR1) catalog, released on March 2025.

The left panel of figure 2.7 shows the spacial distribution comparison of the DESI DR1 (green dots) and the SDSS DR16 (red curve) [78] quasars. The right panel of figure 2.7 shows the redshift distribution of the DESI DR1 (orange) and the SDSS DR16 (green) quasars, together with the DESI Lyman-α pixels.

2.4.2 Cosmological Results

As mentioned previously, one of the primary goals of the DESI experiment is the precise measurement of cosmological parameters and test of the ΛCDM model. By measuring BAO and RSD imprinted in the LSS distribution of galaxies and the IGM, DESI provides valuable insights into the expansion history of the universe and the growth of structure. DESI observations of LSS tracers, including galaxies, quasars, and absorbers, provide constraints on fundamental quantities such as H_0, Ω_m, Ω_b, and w. Additionally, the Lyman-α forest, as well as other absorbers (which will be introduced and explained in the following chapters), enables constraints on $P(k)$ and complements galaxy clustering analyses performed at lower redshifts.

Alternative Models

As explained in **??**, the ΛCDM model has been remarkably successful at explaining numerous observational phenomena. Despite its success, it faces several challenges, such as the nature of dark energy and dark matter, and difficulties in explaining small-scale structures in galaxy formation. These open questions have led to the exploration of alternative cosmological models.

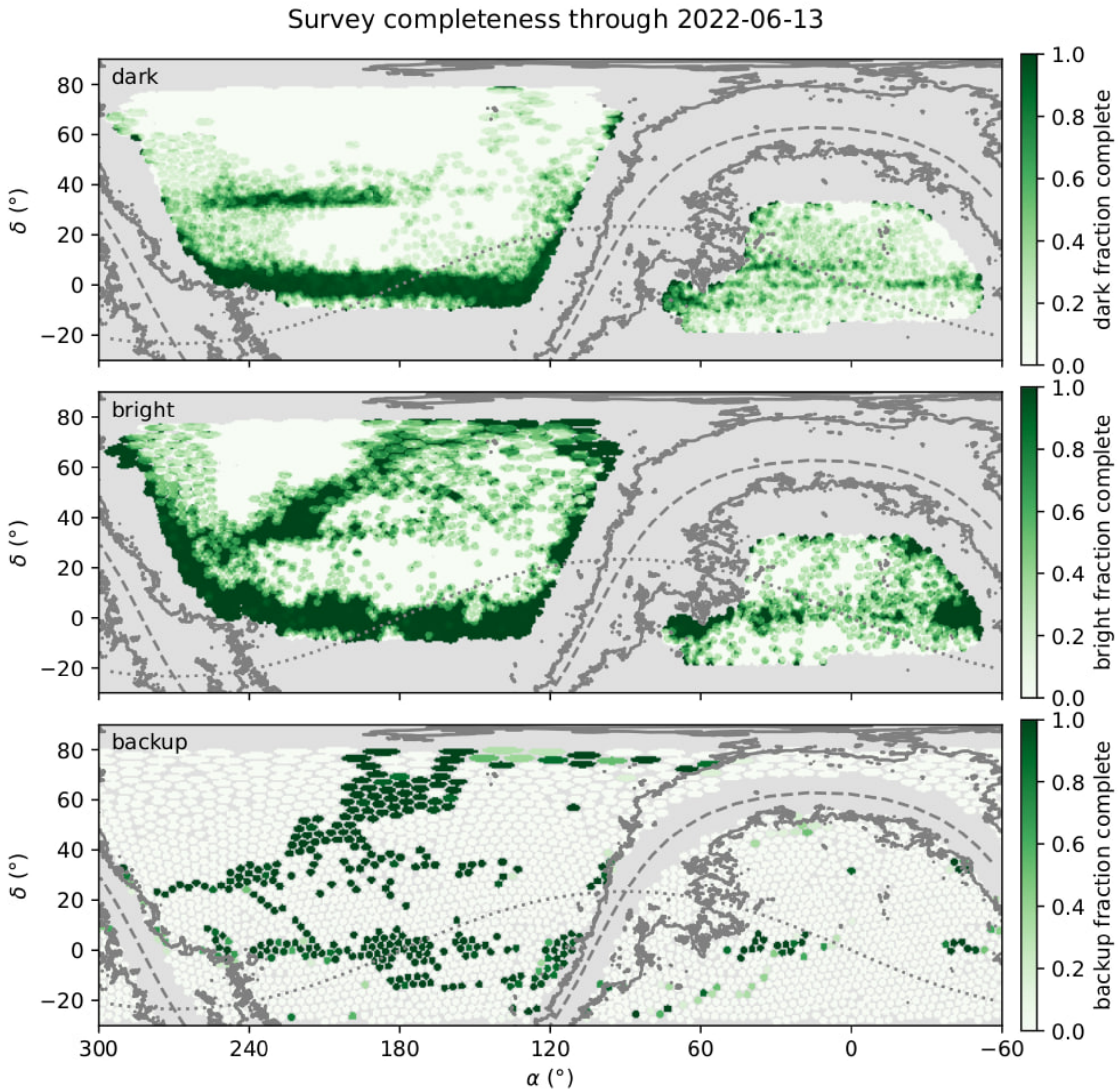


Figure 2.6: Footprint of the first year of data of the Main Survey obtained with DESI. This sample is composed of 6,279,965 BGS, 3,926,272 ELG, 2,829,517 LRG, 1,340,073 QSO, and 3,641,276 stars of the MWS observed by June 13, 2022. This plot shows the completeness of the dark, bright, and backup programs in shades of green, and the unobserved region in white. The gray dashed line illustrates the Galactic plane, while the gray contour is the $E(B-V) = 0.3 mag.$ [74]

Some alternative models focus on modifying the properties of dark energy. In the ΛCDM model, dark energy is described by a cosmological constant with a fixed equation of state $w = -1$. However, dynamical dark energy models, such as quintessence, allow the equation of state to evolve over time,

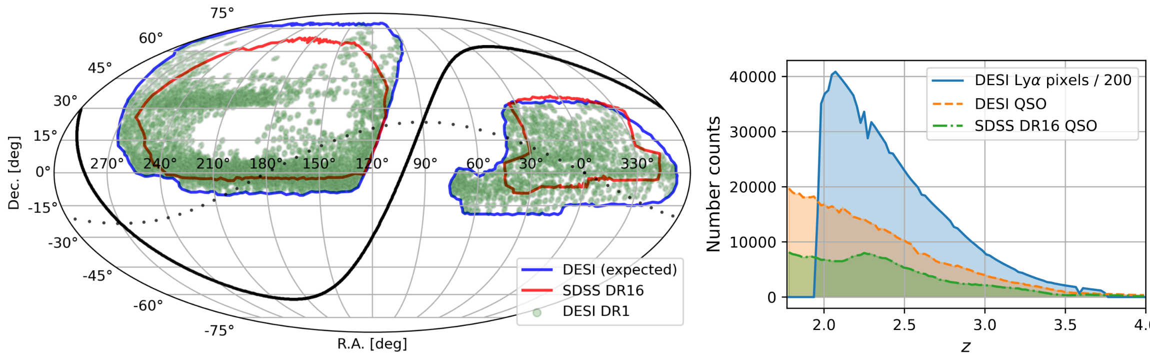


Figure 2.7: Left panel: footprint of the DESI DR1 quasars (green dots) and expected final footprint (surrounded by the blue curve), compared with the one reported by SDSS DR16 (surrounded by red curve). Right panel: redshift distribution of the DESI DR1 (orange) and the SDSS DR16 (green) quasars, together with the DESI Lyman-α pixels. [75]

offering a more flexible approach to fitting cosmological data. Another class of alternatives, known as modified gravity models, proposes that the accelerated expansion of the Universe may result from deviations from general relativity, rather than invoking dark energy. Examples of such models include $f(R)$ gravity. These alternative models aim to address the limitations of ΛCDM and provide new perspectives on the fundamental nature of the Universe.

Among the popular models, we find the $w_0 w_a$ model, a generalization of the ΛCDM model that allows a more flexible description of dark energy. As explained before, in ΛCDM, dark energy is characterized by a constant equation of state parameter $w = -1$, corresponding to a cosmological constant. The $w_0 w_a$ model introduces two parameters, w_0 and w_a, to describe the equation of state of dark energy, which can evolve over time. The equation of state in the $w_0 w_a$ model is given by:

$$w(z) = w_0 + w_a \frac{z}{1+z} \tag{2.1}$$

where z is the redshift, w_0 is the present value of the dark energy equation of state, and w_a characterizes its possible time evolution. The $w_0 w_a$ model allows for dynamical dark energy, meaning that w can evolve as the Universe expands. This flexibility enables the $w_0 w_a$ model to better fit a broader range of cosmological data, especially in cases where observations suggest deviations from the simple cosmological constant hypothesis, providing a more general framework for testing cosmological theories, helping to refine our understanding of the expansion history of the Universe, and the role and nature of dark energy.

The $w_0 w_a$ model appears to be slightly favored by recent DESI data analysis. The precise measurements of DESI of the LSS, including BAO and RSD, enable tighter constraints on the time evolution of dark energy. DESI data, with its extensive redshift coverage and high precision, indicate a slight preference for models where w_0 deviates slightly from -1 and w_a allows for evolution, suggesting that the assumption of a cosmological constant may not fully describe the nature of dark energy. This evolving dark energy scenario provides a better fit to the data across multiple redshift bins, particularly at higher redshifts, where the influence of dark energy may evolve more significantly than predicted by the ΛCDM model.

Nonetheless, the results derived from the Lyman-α forest do not seem to strongly favor either the w_0w_a or the ΛCDM model. Both models appear to fit the Lyman-α forest data similarly well, showing no significant preference for one over the other. This can be seen qualitatively in figure 2.8 (where the results from the Lyman-α forest analysis is shown in violet). This indicates that the Lyman-α forest data alone is insufficient to distinguish between these competing models[2], and the whole analysis followed by the DESI Lyman-α working group, and the development of the research presented in this thesis considers a ΛCDM model.

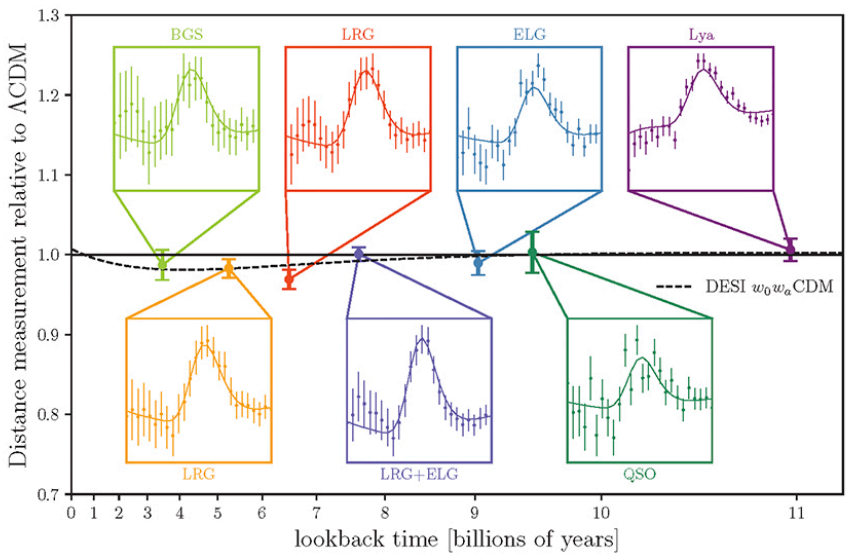


Figure 2.8: Hubble diagram incorporating BAO measurements from all tracers. The results show a slight preference for a w_0w_a model. Credit: Arnaud de Mattia/DESI collaboration.[79]

The measurement of Ω_m reported by DESI BAO, obtained with the first year of observations [**?**] is listed in table 2.9 together with the measurement combined with other external datasets and priors.

[2] There is no bayesian analysis performed with DESI Lyman-α data alone, since it is made together with other tracers.

model/dataset	Ω_m	H_0 [$\mathrm{km\,s^{-1}\,Mpc^{-1}}$]	$10^3\Omega_\mathrm{K}$	w or w_0	w_a
Flat ΛCDM					
DESI	0.295 ± 0.015	—	—	—	—
DESI+BBN	0.295 ± 0.015	68.53 ± 0.80	—	—	—
DESI+BBN+θ_*	0.2948 ± 0.0074	68.52 ± 0.62	—	—	—
DESI+CMB	0.3069 ± 0.0050	67.97 ± 0.38	—	—	–
ΛCDM+Ω_K					
DESI	0.284 ± 0.020	—	65^{+68}_{-78}	—	—
DESI+BBN+θ_*	0.296 ± 0.014	68.52 ± 0.69	$0.3^{+4.8}_{-5.4}$	—	—
DESI+CMB	0.3049 ± 0.0051	68.51 ± 0.52	2.4 ± 1.6	—	—
wCDM					
DESI	0.293 ± 0.015	—	—	$-0.99^{+0.15}_{-0.13}$	—
DESI+BBN+θ_*	0.295 ± 0.014	$68.6^{+1.8}_{-2.1}$	—	$-1.002^{+0.091}_{-0.080}$	—
DESI+CMB	0.281 ± 0.013	$71.3^{+1.5}_{-1.8}$	—	$-1.122^{+0.062}_{-0.054}$	—
DESI+CMB+Panth.	0.3095 ± 0.0069	67.74 ± 0.71	—	-0.997 ± 0.025	—
DESI+CMB+Union3	0.3095 ± 0.0083	67.76 ± 0.90	—	-0.997 ± 0.032	—
DESI+CMB+DESY5	0.3169 ± 0.0065	66.92 ± 0.64	—	-0.967 ± 0.024	—
w_0w_aCDM					
DESI	$0.344^{+0.047}_{-0.026}$	—	—	$-0.55^{+0.39}_{-0.21}$	< -1.32
DESI+BBN+θ_*	$0.338^{+0.039}_{-0.029}$	$65.0^{+2.3}_{-3.6}$	—	$-0.53^{+0.42}_{-0.22}$	< -1.08
DESI+CMB	$0.344^{+0.032}_{-0.027}$	$64.7^{+2.2}_{-3.3}$	—	$-0.45^{+0.34}_{-0.21}$	$-1.79^{+0.48}_{-1.0}$
DESI+CMB+Panth.	0.3085 ± 0.0068	68.03 ± 0.72	—	-0.827 ± 0.063	$-0.75^{+0.29}_{-0.25}$
DESI+CMB+Union3	0.3230 ± 0.0095	66.53 ± 0.94	—	-0.65 ± 0.10	$-1.27^{+0.40}_{-0.34}$
DESI+CMB+DESY5	0.3160 ± 0.0065	67.24 ± 0.66	—	-0.727 ± 0.067	$-1.05^{+0.31}_{-0.27}$
w_0w_aCDM+Ω_K					
DESI	0.313 ± 0.049	—	87^{+100}_{-85}	$-0.70^{+0.49}_{-0.25}$	< -1.21
DESI+BBN+θ_*	$0.346^{+0.042}_{-0.024}$	$65.8^{+2.6}_{-3.5}$	$5.9^{+9.1}_{-6.9}$	$-0.52^{+0.38}_{-0.19}$	< -1.44
DESI+CMB	$0.347^{+0.031}_{-0.025}$	$64.3^{+2.0}_{-3.2}$	-0.9 ± 2	$-0.41^{+0.33}_{-0.18}$	< -1.61
DESI+CMB+Panth.	0.3084 ± 0.0067	68.06 ± 0.74	0.3 ± 1.8	-0.831 ± 0.066	$-0.73^{+0.32}_{-0.28}$
DESI+CMB+Union3	$0.3233^{+0.0089}_{-0.010}$	66.45 ± 0.98	-0.4 ± 1.9	-0.64 ± 0.11	$-1.30^{+0.45}_{-0.39}$
DESI+CMB+DESY5	0.3163 ± 0.0065	67.19 ± 0.69	-0.2 ± 1.9	-0.725 ± 0.071	$-1.06^{+0.35}_{-0.31}$

Figure 2.9: Cosmological parameter constraints from DESI DR1 BAO measurements, combined with external datasets and priors, like SNIa experiments (Pantheon, Union3, and Dark Energy Survey Y5), BBN, and CMB, are presented for the baseline flat ΛCDM model as well as for model extensions that include spatial curvature and two parameterizations of the dark energy equation of state. The reported values correspond to marginalized means with 68% credible intervals, including upper limits where applicable. It is important to note that DESI data alone constrain $r_d h$, rather than H_0. Table obtained from [?]

Chapter 3

Quasars and IGM as Large Scale Structure Tracers.

This chapter describes the physics underlying the Lyman-α forest, explaining its role as an important tracer of the Intergalactic Medium (IGM), Dark Matter, and the Large Scale Structure 1.5.2 It culminates by exploring the importance of the Lyman-α forest as a tracer of BAO at redshifts above 2.1 as well as its relevance and current state in cosmology.

3.1 Quasars

3.1.1 Active Galactic Nuclei

According to [16], "normal" galaxies are those whose spectrum is dominated by the contribution of the black body radiation of all the individual stars conforming them. Their spectrum is typically in the wavelength range of $\sim 4000\mathring{A}$ and $\sim 20000\mathring{A}$, which can be extended if it has active star formation (in smaller wavelengths) and heated dust (by these mentioned young stars).

There exists a subset of galaxies characterized by a *spectral energy distribution* (SED) significantly broader than that of a "typical" galaxy, with wavelengths ranging from X-ray to radio, even when considering gas and stellar formation processes. These galaxies frequently present intense and broad emission lines, and are named *active galaxies.* The source of their activity originates from the central region known as the *Active Galactic Nucleus* (AGN). Despite occupying a very small spatial extent, the luminosity of the AGN can exceed that of their host galaxy by orders of magnitude.

To classify an object as an AGN, it should present at least one of the following properties:

- A compact nuclear region with a significantly greater luminosity compared to an area of equivalent size within a normal galaxy.
- Non-stellar (non-thermal) continuum emission
- Strong emission lines
- Variability in continuum emission and/or in emission lines over relatively brief periods.

The active galaxies are classified into different types according to their observational properties: Seyfert galaxies, radio galaxies, quasars, BL Lac (BL Lacertae, named after its prototype) objects, and OVVs (optically violently variables). Examples of different AGN spectra are shown in figure 3.1.

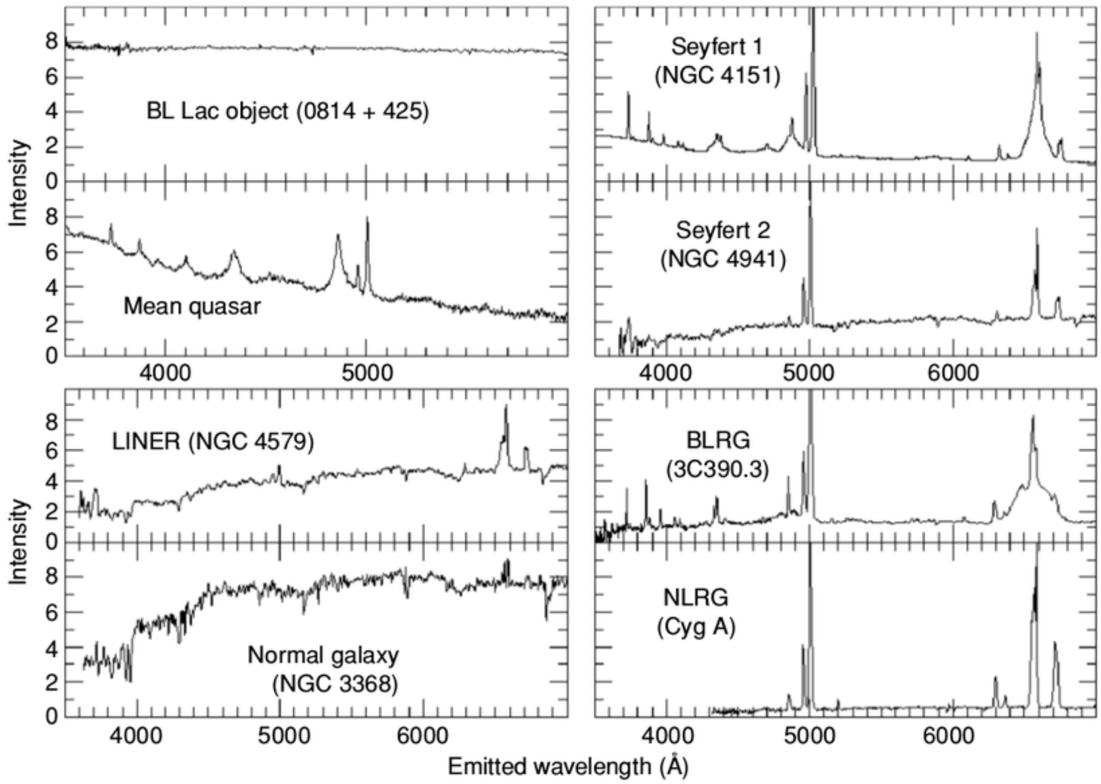


Figure 3.1: Examples of spectra of different types of Active Galactic Nuclei and a normal galaxy. Credit: Bill Keel. Figure taken from [80]

Por que tiene esta forma ?
Cita [ ]

3.1.2 AGN spectra

The AGN spectra are composed of two main ingredients: the continuum and the emission lines. The AGN spectra are broad, going from gamma rays up to radio. The overall SED can be described by a power law $F_\nu \propto \nu^{-\alpha}$, with $0 \leq \alpha \leq 1$. The SED also presents depressions and bumps, being the *blue bump* (at $\approx 10^{15} - 10^{16} Hz$) and the *broad bump* (at $\approx 10^{20} - 10^{21} Hz$) the most important/prominent ones. The wide energy range of the spectrum suggests there are many physical procesess behind their emission, most of them involving relativistic electrons.

The relativistic electrons are thought to be generated in the first order Fermi accelerations that occur in shocks from supersonic flows near the central supermassive black hole (SMBH) in the center of the AGN. The Fermi acceleration is the acceleration that charged particles face when they are reflected (passing, scattering and repeated bouncing) by an interstelar magnetic field, resulting in a power-law energy distribution for the accelerated particles. The motion of these relativistic electrons with a power-law energy distribution in a magnetic field produces synchrotron radiation (generated by relativistic electrons spiraling in a magnetic field) with a power-law spectrum that can be of many decades in frequency in the radio band. The X-ray band energies are explained for photons emitted in inverse Compton process (that occurs when a charged particle gives part of its energy to a photon).

Emission lines are produced by the transitions of excited atoms. These lines are used to infer the physical properties of the emiting gas (density, temperature, composition, etc.) and the sources. Lines are divided into permitted and forbidden/improbable (there are some semiforbidden lines with intermediate spontaneous transitions) depending on the rate of spontaneous lines between energy levels responsible for the emission.

Line ratios are used to distinguish an AGN from a star-forming galaxy, which has emissions due to HII regions generated by young massive stars. The temperature and level of ionization are expected to be higher in an AGN because they have more UV flux, and because the radiation field in an AGN has more high energy photons.

3.1.3 Quasars and QSO

The term *quasar* comes from the term "Quasi-Stellar Radio Source", and it refers to compact radio sources with a non thermal continuum and strong, broad emission lines, and luminous nuclei that dominate in the blue/UV section of the spectrum and are often variable and sometimes have jets. Their luminosities can reach 1000 times the one of a normal galaxy. They usually outshine their host galaxy and it is very hard to observe, nonetheless nearby quasars have been detected in both, elliptical or spiral galaxies, and in ~~disturbed~~ interacting systems.

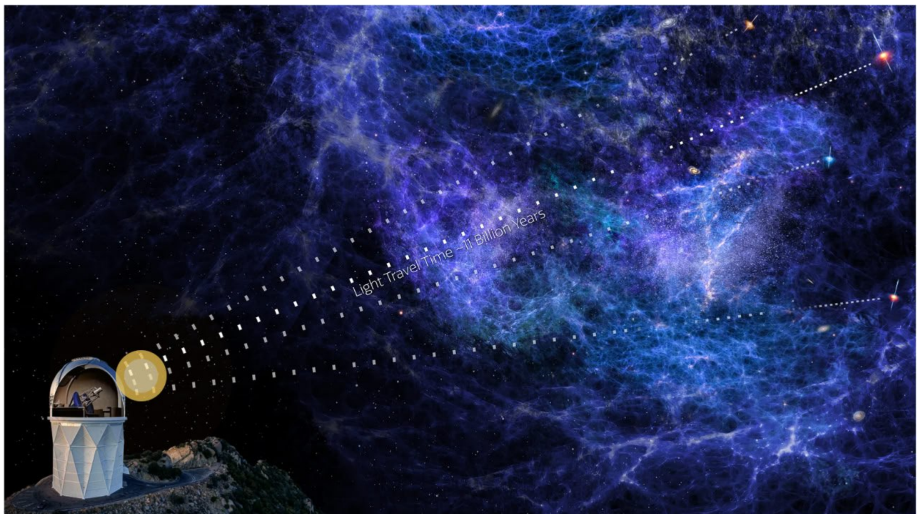
Usually the terms quasar ~~and~~ QSO are used interchangeably, although the term *quasar* refers to a radio-loud source and *QSO* ~~to a~~ radio-quiet one.

3.2 Intergalactic Medium

Galaxies are thought to consist of three primary constituents: dark matter, stars, and gas. The gas is broadly categorized into two components based on its relationship with the galaxy: the interstellar medium (ISM) and the intergalactic medium (IGM). The ISM represents gas directly associated with the galaxy, while the IGM comprises halo gas distributed outside the galaxy but within the host dark matter halo, along with gas unrelated to dark matter halos. Some authors also define the IGM as "anything outside the virial radius of galaxies and clusters (the medium between halos rather than the medium between galaxies)" ([81]). Interactions between the ISM and IGM play crucial roles in the formation and evolution of galaxies through various processes. Consequently, understanding the IGM becomes imperative in comprehending galaxy formation and evolution. Prior to the emergence of stars and galaxies, all baryons were part of the IGM, which later, upon virialization, got accreted by dark matter halos, eventually transforming into stars and the ISM. At present times, more than 50% ~~of the~~ baryonic matter is considered to be in the IGM, less than 10% is found in stars (according to BBN predictions and CMB observations), and the rest making the cold molecular and atomic gas ~~and~~ the hot gas in intracluster medium (ICM).[16]

The IGM is a very important tool for the study of the evolutionary stages of the Universe. It provides information of cosmological events when analyzed at different redshifts, and origins of structures that turn into galaxies. This allows the potential improvement of constraints of the cosmological initial conditions. All the different characteristics of the IGM, including its density, temperature, chemical composition,and spatial distribution, are studied through its interaction (absorption) with the light of background sources and the radiation field produced by it.

The IGM is a valuable tracer of the distribution of baryonic matter and dark matter. When light of a faraway quasar goes through the IGM, electrons changing their orbital levels leave their imprint in the spectrum of the quasar. The density fluctuations of IGM are observed as a continuous absorption field with respect to the unabsorbed emission of quasar spectra.



No se usa en el texto

Figure 3.2: Artistic interpretation of the detection of light from background quasars, travelling through and being absorbed by the IGM, until it is detected at the DESI instrument at the Mayall telescope. Credit: KPNO/NOIRLab/NSF/AURA/P. Marenfeld and DESI Collaboration [82]

3.2.1 Lyman-α and Lyman-β Forests

The Lyman series is a spectral sequence of transitions within the hydrogen atom, manifesting when an electron shifts its energy level from $n \geq 2$ to $n = 1$. This transition leads to the emission of photons within the ultraviolet segment of the electromagnetic spectrum. Each transition within this series is denoted by a Greek letter, corresponding to the difference between the initial and final energy levels. For instance, the transition from $n = 2$ to $n = 1$ is termed Lyman-α, from $n = 3$ to $n = 1$ is Lyman-β, from $n = 4$ to $n = 1$ is Lyman-γ, and so forth.

Among these transitions, Lyman-α is particularly notable. It occurs in hydrogen overdensities within the IGM, where atoms absorb photons emitted by a source, typically a QSO. These absorptions leave behind a distinctive print' in the form of absorption lines at a rest-frame wavelength of $\lambda_{Ly\alpha} = 1215.67\AA$ within the spectra corresponding to different wavelength depending on the redshift of the IGM overdensity. Figure 3.3 shows a quasar spectrum at a redshift of $z = 4.646$, where the Lyman-α emission line is shown in a blue dotted line, and the Lyman-α forest is highlighted in a light violet shade. There are other forests formed analogously, named after the transition in the Lyman series that originates them. Among these, the Lyman-β forest, with an absorption wavelength of $\lambda_{Ly\beta} = 1025.18\mathring{A}$, also shown in a blue dotted line in figure 3.3 (explained in more detail in section 6.1) is of peculiar interest as a potential LSS tracer.

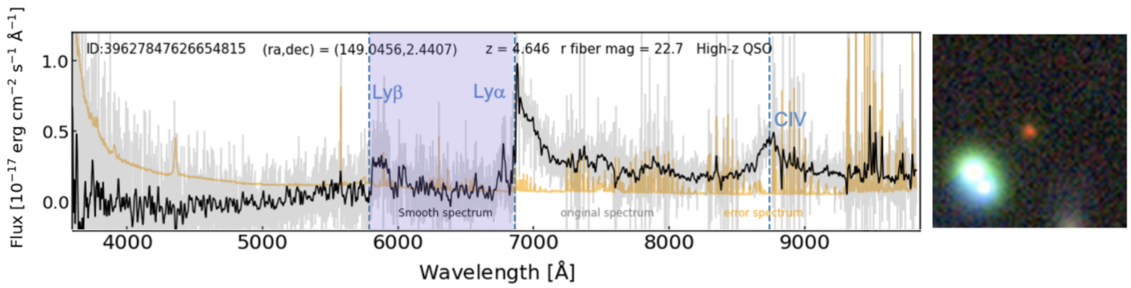


Figure 3.3: The left pannel is the spectrum of a quasar observed by DESI at a redshift of $z = 4.646$, where the Lyman-α and Lyman-β emission lines at rest-frame wavelengths of $\lambda_{Ly\alpha} = 1215.67\mathring{A}$ and $\lambda_{Ly\beta} = 1025.72\mathring{A}$ are shown in blue dotted lines, and the Lyman-α forest is highlighted in a light violet shade. The right pannel shows the 18"x18" thumbnail image of the quasar whose spectrum is that of the left panel, observed by the DESI Legacy Imaging Surveys ([71] and [83]).

3.2.2 Absorber Forests

The process responsible for generating the Lyman-α forest can also produce other forests with alternative absorption lines. These forests serve as valuable probes, offering complementary insights into various aspects of the history of the Universe at different epochs, including their role as LSS tracers and the evolution of cosmic structure. Of particular relevance to this thesis is the exploration of their potential as a source of information and constraints on cosmological parameters through the IGM. Several efforts have been made in this research line studying different regions of the Mg II forest [84] and proposing the C IV forest as BAO tracer ([85], followed by a handful of studies that aim to compute several quantities looking to obtain cosmological information from it, e.g. [42] and [86]). This research explores different angles of the use of Si IV, C IV, Mg II, and Ly-β forests as sources of cosmological information. This is explained in detail in 7.

3.3 Cosmology with the Lyman-α forest

One of the most surprising discoveries in the understanding of the Universe has been its accelerated expansion. The existence of new physics resulting in the accelerated expansion of the Universe at late times is a fact. This breakthrough has ignited dedicated efforts, using various techniques, to quantify the properties of the acceleration. The most popular way to address this quantity is the dark energy, a fundamental constituent of the universe of which we have almost no information except that it has negative pressure and is believed to drive its expansion. The measurements that probe the expansion history of the universe reveal changes in the energy density of its components as the universe expands, thereby allowing a deeper understanding of their nature and characteristics.

The main probes for this analysis are the type Ia supernovae (SN Ia), relying on the fact that they are standardizable candles, and BAO, using the principal of standard rulers. The most used tracer of BAO are galaxies, and they provide a robust probe at redshifts up to $z < 2$. However, there is a need to obtain cosmological information of the younger universe. This is a significant scientific incentive to measure the expansion history of the Universe at higher redshifts. This is provided by means of the Intergalactic Medium (IGM), using mainly the Lyman-α forest at redshifts of $z > 1.8$. This was first proposed by McDonald (2003) and Weinberg (2003) and first studied in detail by McDonald & Eisenstein (2007) [38]. [87], [88], [89], [90].

With the start of the millennium, the number of studies using the Lyman-α forest to trace the LSS have grown significantly. Particularly, it has been widely used in the development of cosmology with analyses such as tomographical studies of the IGM [e.g. 28, 91–97], measurements of the line-of-sight one-dimensional flux power spectrum (P_{1D}) [e.g. 98–107], constraints of neutrino masses [e.g. 108–110] and dark matter models [e.g. 110–115], and the measurement of the BAO scale with correlation functions [e.g. 78, 87, 116–123], technique in which this research is centered.

The use of the latter technique has evolved from early studies that used samples of quasars enough only to identify Lyman-α absorption and correlate them along individual lines of sight (LOS), to the probe of the full three-dimensional distribution of neutral hydrogen with a sample of 10,000 QSO of BOSS, and the evidence for the expected long-range correlations, including the RSD [124], until successfully achieving the first measurement of the BAO peak reported in [117] in 2013 with the ninth Data Release of the SDSS, composed of a sample of $\approx 60,000$ QSO at $z \sim 2.3$ ([125], [126]).

The DESI survey seeks to achieve exceptional precision in mapping BAO across a broad spectrum of redshifts during a five-year campaign. During the observations of the first year of operations, it has performed the BAO analysis at $z = 2.33$ computing auto-correlations of the Lyman-α forest dataset and its cross-correlation with quasar positions using a sample of over 450,000 Lyman-α forest spectra and more than 700,000 quasars over an area of 9500 deg^2 from its first data release (DR1). This is the largest dataset ever obtained for this type of analysis, duplicating the sample of the previous eBOSS DR16, which has been the state of the art since [75], [78].

Chapter 4

Correlation Function Measurements from Quasars and IGM Tracers

The IGM is a powerful probe of the LSS by tracing neutral hydrogen and other abundant elements in the Universe. The statistical properties of the LSS can be studied through the flux correlation function $\xi_F(r)$, which measures the spatial correlation between absorption features and provides information about the IGM clustering, and the flux power spectrum $P_F(k)$, which provides information about growth of structure and other features from which cosmological parameters and some properties of the IGM can be inferred. Throughout this chapter, these tools will be explained, as well as its use for this research work.

4.1 IGM Absorption Fluctuations

In the same way as the correlation function $\xi(r)$ is computed with positions of galaxies, flux correlation functions can be derived in a similar manner. However, instead of galaxy positions, these functions rely on the locations of IGM density fluctuations, traced by the flux fluctuations due to the optical depth of the absorption that generates the different forests.

Since neutral hydrogen is the most abundant element in the Universe, the best and most used tracer of IGM has been the Lyman-α forest. Nonetheless, since the objective of this thesis work is to explore other transitions as IGM tracers and sources of cosmological parameters, the statistical tools to study absorber forests has been generalized to analyze any transitions whose physics are consistent with the aforementioned. Following the notation used in [75], we define the following regions to develop the analysis of this research:

- ⋆ Region A: the restframe wavelength range of the QSO spectrum from $1040\mathring{A}$ to $1205\mathring{A}$
- ⋆ Region B: the restframe wavelength range of the QSO spectrum from $920\mathring{A}$ $1020\mathring{A}$

These regions are delimited by the corresponding transitions in restframe wavelength, and should be differentiated from the absorption produced by the Lyman-α and Lyman-β transitions, conforming the corresponding forests. As we define these regions, we can better understand the *contaminants*, such as metals in region A (misunderstood by the analysis pipelines as Lyman-α and biasing the shape of the correlations [**?**]), or Lyman-α lines in region B. This is illustrated in figure 4.1. It should be noted that when this work mentions the analysis of the Lyman-β forest, it refers to the Lyman-β forest in region B, and does not consider Lyman-α absorption contaminating it.

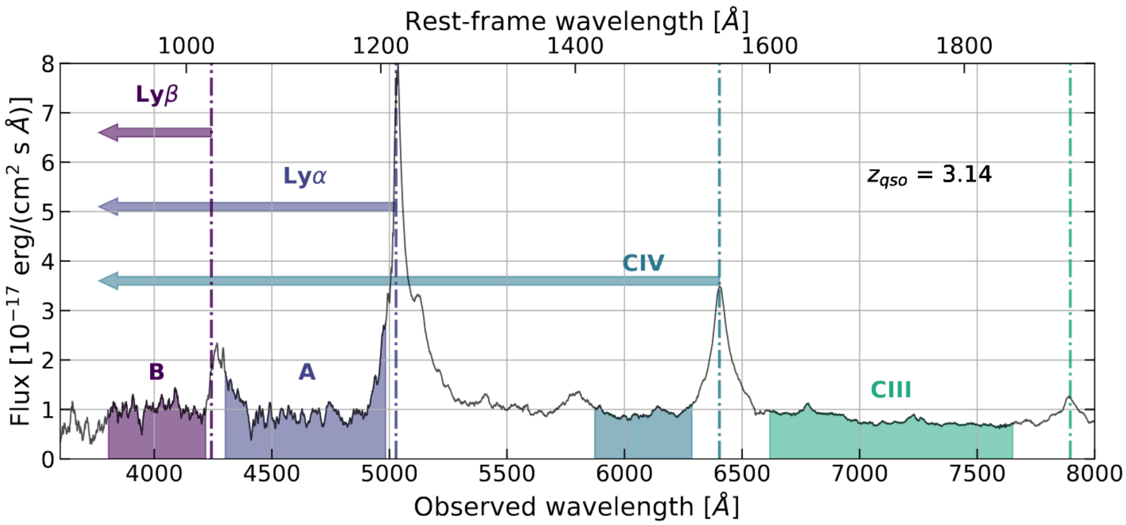


Figure 4.1: Quasar spectrum from the first year of the DESI survey, at a redshift of $z = 3.14$. The dashed lines indicate the positions of prominent emission lines. The colored regions on the blue side of each emission line correspond to the associated absorption forests, named after the respective ions: light green marks the C III forest, defined by the C III emission line; light blue corresponds to the C IV forest; indigo highlights the Lyman-α forest and Region A; and purple denotes the Lyman-β forest and Region B [75].

The following paragraphs provide a step-by-step explanation on how to compute the correlation functions taking as a starting point the measurement of the flux of the QSOs in the catalogs. For further reference, this process is thoroughly detailed in [[127], [75]][1][[120], [121], [123], [122], [78]][2].

The flux f of an object is the amount of radiation that we observe from it at different wavelengths λ. The observed flux of a quasar is denoted as $f_q(\lambda)$. The unabsorbed quasar continuum, denoted $C(\lambda)$ (or specifically $C_q(\lambda)$ for a quasar), corresponds to the intrinsic emission spectrum of the object, representing the photon flux emitted across all wavelengths in the absence of any intervening absorption. $\overline{F}(\lambda)$ is the mean transmission of IGM at a given wavelength λ.

The transmitted flux fraction F is given by the relation $F = e^{-\tau}$, with τ being the optical depth (the quantification of how much light from the source is absorbed by the IGM when passing through it). The optical depth τ is connected to local matter overdensities δ_m. To establish this connection, we rely on two key physical assumptions: first, under adiabatic expansion, the temperature T of

[1] DESI pipeline analysis.
[2] BOSS and eBOSS pipeline analysis.

the low-density IGM is tightly correlated with its density ρ, following the relation

$$\frac{d\ln T}{d\ln \rho} = \gamma - 1. \tag{4.1}$$

Second, assuming photoionization equilibrium, the number density of neutral hydrogen atoms scales with the gas density and temperature as $n_{\mathrm{HI}} \propto \rho^2 T^{-0.7}$. Since the optical depth is directly proportional to n_{HI}, these assumptions together allow us to express τ as a function of the local matter density. This leads to the formulation known as the aforementioned fluctuating Gunn–Peterson approximation (FGPA) which models the optical depth as

$$\tau(z, \mathbf{x}) = \tau_0(z)\left[1 + \delta(z, \mathbf{x})\right]^{\alpha(z)}, \tag{4.2}$$

where $\tau_0(z)$ is a redshift-dependent normalization factor encapsulating the photoionization rate and thermal state of the gas, and $\alpha(z) = 2 - 0.7(\gamma(z) - 1)$ reflects the slope of the temperature–density relation. Both $\tau_0(z)$ and $\alpha(z)$ are treated as free parameters in practice, and the method for choosing them is explained in [128].

The fluctuations around the mean transmitted flux fraction are called *transmitted flux field*, they are given as a function of wavelength $\delta_q(\lambda)$, and calculated as:

$$\delta_q(\lambda) = \frac{f_q(\lambda)}{C_q(\lambda)\overline{F}(\lambda)} - 1. \tag{4.3}$$

This is the relative ratio of the absorbed flux to the expected flux. The product $\overline{F}(\lambda)C_q(\lambda)$ is the mean expected flux for a quasar q at a wavelength λ. Following [75] and [127], this is estimated with the approximation:

$$\overline{F}(z)C_q(\lambda) = \overline{C}(\lambda_{RF})\left(a_q + b_q\frac{\Lambda - \Lambda_{min}}{\Lambda_{max} - \Lambda_{min}}\right), \text{with } \Lambda = \log\lambda. \tag{4.4}$$

In this expression, $\overline{C}(\lambda_{RF})$ is an estimate of the mean continuum, the rest-frame ($\lambda_R F$) spectra of all quasars. The pair of parameters (a_q, b_q) describe the diversity of quasar spectra and the redshift evolution of the mean transmission $\overline{F}(\lambda)$. The flux variance (total variance of the data, later used to calculate the pixel weights described in 4.2.1) σ_q^2 in the pixels is given by:

$$\sigma_q^2 = (\overline{F}(\lambda)C_q(\lambda))^2\sigma_{LSS}^2(\lambda) + \eta_{pip}(\lambda)\sigma_{pip,q}^2(\lambda). \tag{4.5}$$

In this expression, the first term represents the intrinsic variance of the Lyman-α forest. The second term represents the variance from the instrumental noise, and is conformed by the noise variance estimated by the pipeline $\sigma_{pip,q}^2$, normalized with the expected flux, and multiplied by a function that corrects possible mis-calibrations of the instrumental noise, $\eta_{pip}(\lambda)$. The functions $\eta(\lambda)$ and $\sigma_{LSS}^2(\lambda)$, the intrinsic variance of the Lyman-α forest, are iteratively fitted with the mean quasar continuum $\overline{C}(\lambda_{RF})$ and the parameters (a_q, b_q). These parameters are typically computed in around five convergent iterations[3] by the code PICCA (Package for IGM Cosmological-Correlations Analyses), a software package designed to process Lyman-α forest data for LSS analyses taking quasar spectra as input to produce 1D and 3D correlation functions (both auto- and cross-correlations) between quasars and absorption features [129].

[3]explained in more detail in [127]

4.2 The Flux Correlation Functions

In general, there are two types of datasets considered in the analysis of this research. One of them is conformed by the *fluctuations* $\delta(\lambda)$ of absorber forests, described in chapters 2 and 5. The fluctuations in question are mainly those produced by the Lyman-α, Lyman-β, C IV, Si IV, and Mg II transitions. The other dataset consists of quasars that make up the sample of *absorption quasars* in the DESI DR1 catalog, as reported by [75]. This gives a combination of 6 correlations to perform the analysis of these works:

- Auto-correlations: Lyman-α, Lyman-β, C IV, Si IV, Mg II correlating each with fluctuations of the same species.
- Cross-correlations: Lyman-α, Lyman-β, C IV, Si IV, Mg II fluctuations correlated with quasar positions

These are measured in bins of comoving separation along and across the line of sight ($r_{\parallel}$ and $r_{\perp}$ respectively), obtained from the angular and redshift separations using the standard cosmological model explained in Chapter 1 with cosmological parameters from table 1.1 of [18]. Their computation is described in the following subsections, following [127] and [75]:

4.2.1 Auto-correlation

In order to measure the auto-correlation function, redshift and angular separations ($\Delta z, \Delta\theta$ respectively) are transformed into longitudinal (along the LOS) and transverse comoving separations ($r_{\parallel}$ and $r_{\perp}$). For a pair of measurements (pixel-pixel or pixel-quasar) denoted by (i, j) at redshifts (z_i, z_j) and separated by an angle $\Delta\theta$ we calculate $r_{\parallel}$ and $r_{\perp}$ as:

$$r_{\parallel} = (D_C(z_i) - D_C(z_j)) \cos\left(\frac{\Delta\theta}{2}\right), \tag{4.6}$$

$$r_{\perp} = (D_M(z_i) + D_M(z_j)) \sin\left(\frac{\Delta\theta}{2}\right) \tag{4.7}$$

For a given pixel i, the redshift is calculated assuming the wavelength of a known absorption, denoted by m such that $z_i = \frac{\lambda_{obs}}{\lambda_m} - 1$. In equation 4.6, $D_C(z)$ is the comoving distance and $D_M(z)$ the angular (or transverse) comoving distance. Assuming a fiducial cosmology with $\Omega_k = 0$, $D_M = D_C$ [75], [127]. A fiducial cosmology refers to a set of cosmological parameters adopted as a reference model when analyzing data. These parameters are used to convert observables, such as RA, DEC,, and z, into physical quantities like distances. While the fiducial cosmology may not exactly match the true underlying cosmology, it provides a consistent framework for measurement. Deviations between the fiducial and true cosmology can later be accounted for through model comparisons or geometric corrections, such as the Alcock-Paczyński effect (see ??.

For the measurement of this correlation, a weighted covariance estimator is used [?], [116], [78]:

$$\xi_A = \frac{\sum_{i,j \in A} w_i w_j \delta_i \delta_j}{\sum_{i,j \in A} w_i w_j}, \tag{4.8}$$

with A being a 4Mpc/h width bin (or pixel) in $(r_\parallel, r_\perp)$, and w_i, w_j are the weights[4], which are applied to the Lyman-α pixels to account for the diverse signal-to-noise ratios resulting from the varying brightness of quasars and exposure times, ensuring that our measurements of correlations accurately reflect the underlying data (technical details are explained in [127], [75]). The sum is performed over all pixel pairs i, j across different LOS, excluding the ones in the same one to avoid continuum fitting errors that affect the forest. For each bin A, the model correlation ξ_{mod} is evaluated at the weighted mean separation $r_\parallel, r_\perp$ of the Lyman-α pixel pairs in the data.

For BAO analysis with the Lyman-α forest, the autocorrelation is generally measured from $[0, 200]$Mpc/h in different amount of bins, parallel and perpendicular to the LOS. Figure 4.2 illustrates the Lyman-α autocorrelation function as a function of μ, r, where $\mu = r_\parallel/r$, ranging from $\mu \in [0.95, 1]$ (closest to the LOS) to $\mu \in [0, 0.5]$ (furthest from the LOS), and $r^2 = r_\parallel^2 + r_\perp^2$.

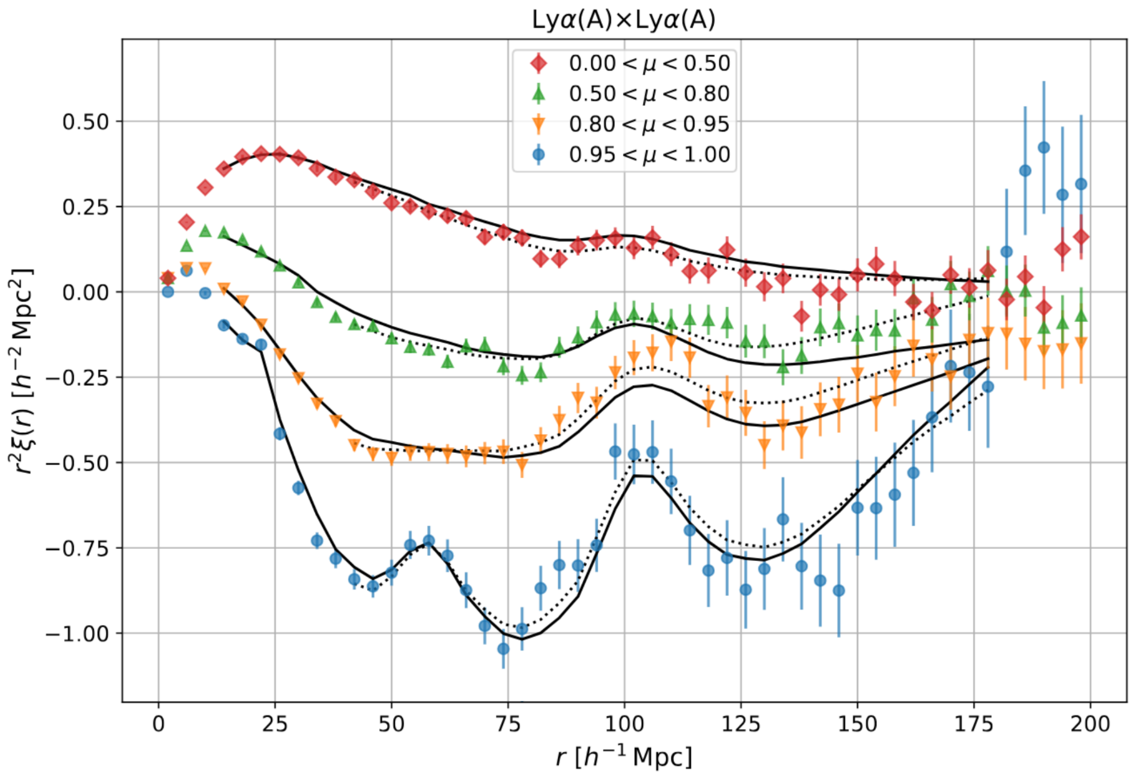


Figure 4.2: Lyman-α auto-correlation function computed with data from the first year of operations of the DESI survey (points), with the best-fit-models (black solid lines). The different colors represent the four analyzed wedges, where the blue points are the correlation closest to the LOS and the red one is the farthest. [75]

[4]a modified Lyman-α weights, to be used instead of the inverse of the variance σ_q^2, explained in detail in eq. 3.1 in [127] and eq. 3.2 in [75]

4.2.2 Cross-correlation

The cross-correlation function measures the relationship between two different tracers (or groups of datasets) as a function of a separation r. In the context of IGM, the correlations are generally IGM fluctuations (often Lyman-α, but for the development of this work any of the absorber forests defined previously) correlated with QSO across different scales. Similarly to the auto-correlation, the estimator for the cross-correlation is defined as [127], [75]:

$$\xi_A = \frac{\sum_{i,j\in A} w_i w_j \delta_i}{\sum_{i,j\in A} w_i w_j} \tag{4.9}$$

for a pixel of a given absorption i and quasar j, and for the weights w_j corrected for QSO bias evolution given by:

$$w_j = \left(\frac{1+z_j}{1+z_{\rm fid}}\right)^{\gamma_q - 1}. \tag{4.10}$$

where $\gamma_q = 1.44 \pm 0.08$ [84] and $z_{\rm fid} = 2.25$[127], and the sum runs over quasar pixel pairs, excluding pixels from their respective background quasar. The cross-correlation is also computed in terms of the LOS and transverse separation, with $r_\perp \in [0, 200]\mathrm{Mpc}/h$ and $r_\parallel \in [-200, 200]\mathrm{Mpc}/h$.

Figure 4.3 shows the shape of the cross-correlation computed with the Lyman-α forest, as a function of μ, r. It can be seen that the cross-correlation is asymmetric under the permutation of the two tracers due to several reasons: spurious excess correlation due to contamination from other transitions appears at either $r_\parallel > 0$ or $r_\parallel < 0$, that μ is averaged over $\mu \in [-1, 1]$, since the cross-correlation has negative values of $r_\parallel$, systematic redshift errors (see eq. 4.9 of [127]), and the redshift evolution of the bias of each tracer. It is usually defined as *negative separations* when the IGM pixel is between the observer and the QSO, and *positive separations* when the quasar is between the observer and the IGM pixel.

It can also be seen that the shape of the cross-correlation seems to be reversed with respect to the matter correlation function. This is because the bias of the absorbers is negative (see clarification about bias in 4.3), and the one of QSO is positive, resulting in a negative product in general.

section/table/figure ?

Wedges and slices in the correlation function

The correlation functions between the forests of different absorbers are often represented in different angular bins relative to the LOS. These are called **wedges**. These are useful for studying different features, such as metal contamination. The correlations depend on two components of separation: $r_\parallel$, along the LOS, and $r_\perp$, transverse component. Instead of analyzing the full correlation function, one can integrate it over specific angular ranges (or wedges) in μ, where $\mu = r_\parallel / r$ is the cosine of the angle between the separation vector and the line of sight. By analyzing wedges, one can isolate different physical effects, improving constraints on BAO measurements while studying and mitigating systematics.

In the case of cross-correlations, **slices** refer to fixed-separation cuts through the correlation function. These slices provide a way to visualize the clustering signal and other features more clearly. A fixed $r_\perp$ slice shows how the correlation function varies along the line of sight at a given transverse separation. A fixed $r_\parallel$ slice reveals how the clustering signal behaves in the transverse direction, highlighting the large-scale structure perpendicular to the line of sight.

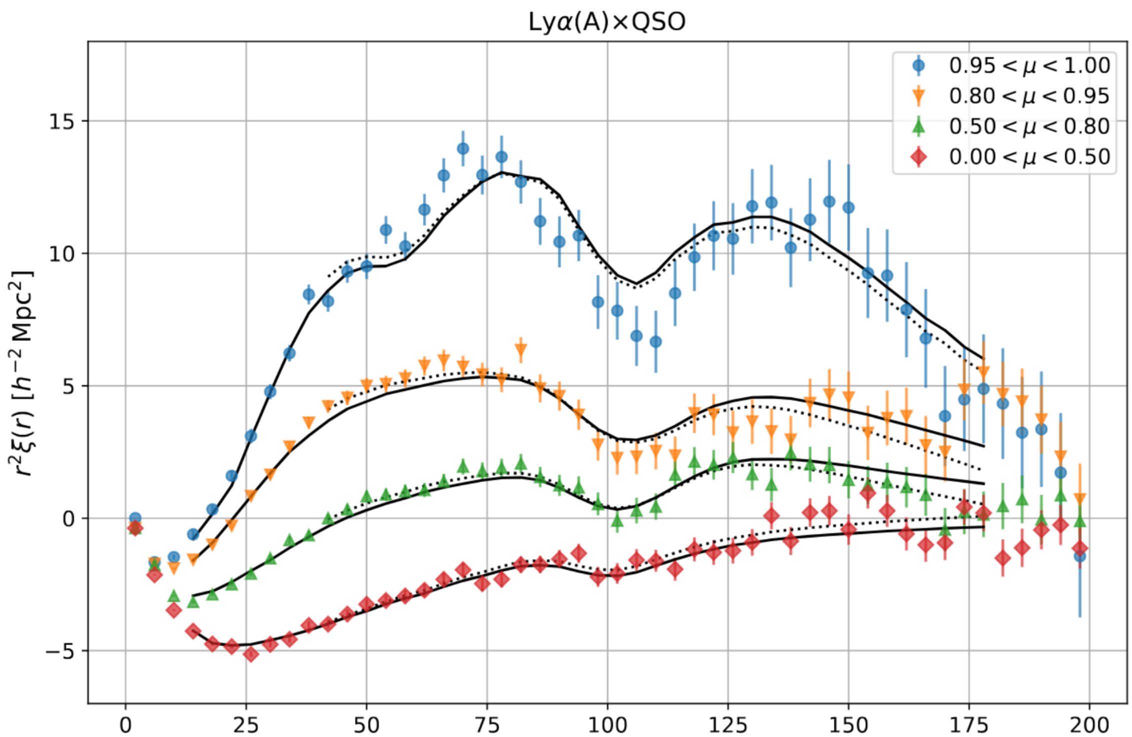


Figure 4.3: Lyman-α and quasars cross-correlation function computed with data from the first year of operations of the DESI survey (points), with the best-fit-models (black solid lines). The different colors represent the four analyzed wedges, where the blue points are the correlation closest to the LOS and the red one is the farthest. [75]

4.3 Flux Power Spectrum $P_F(k)$

According to what was explained in 1.4.1, the Fourier transform of a given $\xi(r)$ is a power spectrum $P(k)$. This means that, just as there is a matter $P(k)$ to a $\xi(r)$, there is a flux $P_F(k)$ to a flux $\xi_F(r)$. Since the flux traces the density field, the flux power spectrum $P_F(k)$ provides constraints on the total matter power spectrum, considered in the linear regime $P_{lin}(k)$. They are related by the equation:

$$P_F(k) = b_F(1 + \beta_\mu^2)^2 P_{lin}(k)D(k) \tag{4.11}$$

where b_F is the flux bias, a linear bias parameter that relates fluctuations in the transmitted flux to the underlying linear matter density fluctuations described by $P_{lin}(k)$. This type of bias arises because observable tracers of the LSS (such as galaxies, quasars, and IGM traced by different forests) do not perfectly follow the distribution of dark matter, but instead trace it in a biased way. In the linear regime, where density fluctuations are small, this bias is assumed to be scale-independent and can be modeled as a constant multiplicative factor. Specifically, the flux transmitted through the IGM is anti-correlated with the matter density: higher matter density leads to stronger absorption and thus lower flux. This results in a negative flux bias, meaning that the flux fluctuations trace the matter fluctuations with opposite sign. The accurate measurement of b_F is essential, since it allows the inference of the underlying matter power spectrum $P_{lin}(k)$ from IGM statistics, enabling cosmological parameter constraints from Lyman-α forest data. The term $(1-\beta\mu^2)^2$ originates from linear RSD (Kaiser effect). It enhances clustering along the line of sight due to peculiar velocities of the absorbing gas. Here, β is the RSD parameter, related to the linear growth rate of structure and describing the effect of peculiar velocities on the observed clustering, μ is the cosine of the angle between the wavevector k. The last term, $D(k)$ introduces the non-linear corrections, which are corrections from small-scale nonlinear effects, including thermal broadening, gas pressure effects, and velocity dispersion of the absorbing gas.

Among these terms, it is important to highlight the importance of b_F, since it is the quantity that is measured with absorber forests in the development of this research work.

Chapter 5

The Lyman-α Forests of DESI as Tracers of the LSS

Over the years, several extensive analyses have been performed, refining our understanding about the Lyman-α forest, making it a key tracer of the Universe at high redshift. From the measurements of an early BOSS cite and cite, to the latest DESI data, the studies performed with the Lyman-α forest have proved to be robust estimators of the BAO detection and estimation of cosmological parameters, which has paved the way for high precision galaxy cosmology at redshifts $z > 2$.

This chapter reviews the advancements done since the start of the cosmological analyses with the IGM traced by the Lyman-α forest, highlighting the key results and methodologies that underscore its importance within the framework of cosmological analyses.

As explained in chapter 2, DESI is a groundbreaking project designed to map the Universe with unprecedented precision by measuring BAO over a range of redshifts. Over the course of its five-year survey, DESI aims to measure the redshift of over 40 million galaxies and quasars, covering 14,000 deg^2 of the sky. The project has already achieved significant milestones, including assembling a dataset of its first year of operations with redshifts for approximately 13 million galaxies, 1.5 million quasars, and 4 million stars over more than 9,500 deg^2, and is currently performing the first BAO and cosmological analyses assembling a dataset with the data from the third year of operations.

A key aspect of DESI is using the Lyman-α forest as a tracer of the LSS of the Universe. These absorption features are sensitive to the distribution of matter on cosmological scales, making them a powerful tool for tracing the underlying dark matter distribution and measuring the expansion history of the Universe at redshifts above $z > 2$.

In its first year of data, DESI has used this dataset to perform seven BAO measurements across different redshifts. Among these, a significant result was achieved using the Lyman-α forest, where DESI performed an auto-correlation analysis at a redshift of $z = 2.33$. Additionally, a cross-correlation between the Lyman-α forest and the positions of quasars enabled an even more precise measurement of BAO at these high redshifts. While chapter 4 delves in the methodology followed to make these measurements, this chapter discusses the results of these analyses, highlighting how DESI uses the Lyman-α forest as an LSS tracer, contributing to our understanding of the expansion of the Universe at redshifts inaccessible galaxy surveying.

As mentioned before, the results of the DESI survey presented in [75] are built on the success and constant evolution and improvement of previous cosmological surveys such as the Baryon Oscillation

Spectroscopic Survey (BOSS)cite and the extended BOSS (eBOSS) cite have successfully utilized the Lyman-α forest as a powerful probe of LSS. BOSS pioneered the use of the Lyman-α forest to measure the distribution of matter at high redshifts, resulting in the first detection of BAO at $z \sim 2.3$ confirm and cite. eBOSS extended these measurements, improving the precision of BAO and providing cross-correlations between the Lyman-α forest and quasar positions. These groundbreaking efforts laid the foundation for current surveys like DESI to refine high-redshift BAO measurements and further explore the universe's expansion history.

The evolution of these analyses and results is presented in the following paragraphs:

Since the first detection of the BAO in 2012, with BOSS DR9 [117], the tools for the analysis of cosmo-statistics of the IGM have evolved significantly, converging in the analysis presented in DESI DR1. [117], [45], [121], [122], [123], [78], and [75] have collectively traced the evolution of Lyman-α forest cosmology from its pioneering use in BOSS over ten years ago, to the latest developments in DESI.

1. Busca et al [117] present the first detection of BAO using the Lyman-α forest from the ninth data release of SDSS-III BOSS [130]. It pioneered the use of the Lyman-α forest for cosmology, demonstrating the potential of this high-redshift probe to map the LSS. The analysis centered on auto-correlation of the Lyman-α forest at redshift $z \sim 2.5$, providing the first cosmological constraints on the Hubble parameter and angular diameter distance at such high redshifts. This set the stage for future work by showing that the Lyman-α forest could successfully measure BAO at early cosmic times.

2. Bautista et al. [120] presents a BAO measurement at redshift 2.3 using data from the twelfth data release of SDSS-III BOSS [131]. The analysis utilizes the Lyman-α forests of more than 150,000 quasars in the redshift range $2.1 < z < 3.5$. It focuses on the Lyman-α auto-correlation, and presents a significant detection of the BAO peak, contributing to the determination of the cosmological parameters associated with the expansion rate of the Universe at high redshift. This high-redshift BAO measurement provides unique constraints on the Hubble parameter $H(z)$ and the angular diameter distance $D_A(z)$, enhancing our understanding of dark energy and the large-scale structure of the universe at earlier cosmic epochs. It also elaborates on the construction of synthetic datasets for the validation of the analyses.

3. Du Mas des Bourboux et al. [121] focuses, complementary, on the cross-correlation between the Lyman-α forest and quasars from the twelfth data release of SDSS-III BOSS [131]. This paper refined the methodology for measuring BAO at high redshifts. By incorporating the cross-correlation into the analysis, it achieved even more accurate results for the BAO signal at $z = 2.4$. Both works were done with significant improvements in data processing, statistical treatment, and the development of synthetic datasets to study, characterize, and mitigate systematic errors.

4. De Sainte Agathe et al. [122] provide BAO measurements through the auto-correlation of Lyman-α at $z = 2.34$ from the twelfth data release of SDSS-IV eBOSS DR14 data [132]. Over 170,000 Lyman-α forests were analyzed to determine the parameters of the BAO peak, with results that are consistent with the ΛCDM model. Additionally, Lyman-α absorption in the Lyman-β region (nowadays defined as *region B*, see 4.1) of 56,154 spectra were used for the first time.

5. Blomqvist et al. [123] present a parallel measurement of BAO at $z = 2.35$ by analyzing cross-correlations between Lyman-α absorption and quasar positions from the twelfth data release of SDSS-IV eBOSS DR14 [132] data. The study utilizes over 266,000 quasars, expanding the range of Lyman-α forest data to include the Lyman-β region (nowadays defined as *region B*, see 4.1) as well, which increases precision by reducing noise and systematics.

7. Du Mas des Bourboux et al. [78] present BAO measurements from both: the Lyman-α

forest auto-correlation and Lyman-α-quasar cross-correlations, utilizing data from the sixteenth data release of SDSS-IV eBOSS [133]. This analysis is performed with a sample of 210,005 quasars with redshifts $z > 2.1$ and 341,468 quasars at $z > 1.77$. The study is made with an updated version of the code PICCA[129] that makes it capable not only of expanding the spectral range to include the Lyman-β region (nowadays defined as *region B*, see 4.1), but also improving the S/N and reducing statistical uncertainties. Combining auto-correlation and cross-correlation yields distances at $z = 2.33$, consistent with ΛCDM. These high-precision measurements validate Lyman-α forest data as an effective BAO tracer, setting a foundation for further studies like DESI to probe large-scale structure at high redshifts.

8. The DESI Collaboration [75] present the the BAO measurements with all the tracers described in Chapter 2, conforming the data release of its first year of operations. In particular, the measurements with the 1.5 million quasar sample. This was done by applying similar but more robust techniques as in BOSS and eBOSS, performed by the codes PICCA [129] and VEGA [134] (see 6.3), but with a dataset that is six times larger. DESI builds on the success of previous surveys by increasing the sample size of quasars and Lyman-α forest spectra, improving the precision of BAO measurements at $z = 2.33$ using both auto-correlation of the Lyman-α forest and cross-correlation with quasar positions. With this sample, DESI provides an unprecedented level of precision in high-redshift BAO measurements, enabling even tighter constraints on the Hubble parameter and angular diameter distance compared to BOSS and eBOSS (as presented in **??**). To illustrate the comparison between the previous sample to the one conforming the DESI DR1 dataset, the right panel of figure 2.7 compares the redshift distribution of the quasars that conform the DESI DR1 and SDSS DR16 datasets, together with the Lyman-α pixels.

Drawing from the achievements and constant evolution and improvement of the IGM analyses probed by the Lyman-α forest in region A and B and, following the proposal of considering the C IV forest as a possible BAO tracer [85] together with several analyses of the C IV and Mg II forests [42], [84], [86] the following chapters will use the same techniques, but with using the Lyman-β (in region B), Si IV, C IV, and Mg II forests as LSS tracers and analyze their potential use as BAO tracers as well with the largest and most precise dataset up-to-date.

Chapter 6

The Lyman β Forest as Tracer of the Large Scale Structure

While the Lyman-α forest has been extensively used to trace the distribution of matter at high redshifts, the Lyman-β forest remains a less explored but promising alternative. By studying the correlations in forests of multiple LOS, it can be used as an independent tracer of the cosmic web, offering complementary constraints on the matter power spectrum and the growth of structure.

One of the key motivations for investigating the Lyman-β forest is its potential to be used as a BAO tracer. Since the restframe wavelengths of the forest is different than the Lyman-α, it allows the study of additional modes of the LSS, potentially reducing statistical uncertainties and systematic biases. Additionally, joint analyses of both forests could improve constraints on cosmological parameters by increasing the effective volume of the intergalactic medium sampled.

In this chapter, we explore the theoretical foundations and observational prospects of using the Lyman-β forest as a tracer of LSS, and present the first cross-correlations between the Lyman-β forest in the B region with quasars from the DESI DR1 sample.

6.1 The Lyman-β forest

Section 3.2.1 describes the phenomenology that generates the Lyman transitions and forests. Summarizing, when an electron transitions from energy level $n = 3$ to $n = 1$, it emits the Lyman-β line, at a wavelength of $\lambda_{Ly\beta} = 1025.18Å$. The Lyman-$\beta$ forest (from $\lambda = 920Å$ to $\lambda = 1020Å$) is the result of the absorption of light of a quasar at the Lyman-β wavelength in IGM overdensities distributed at a redshift range of $2.91 < z < 8.63$. Figure 6.1 shows a quasar spectrum at a redshift of 3.16, observed by DESI, where the Lyman-β emission line and the Lyman-β forest are highlighted in blue, together with Lyman-α emission line and the Lyman-α forest highlighted in red.

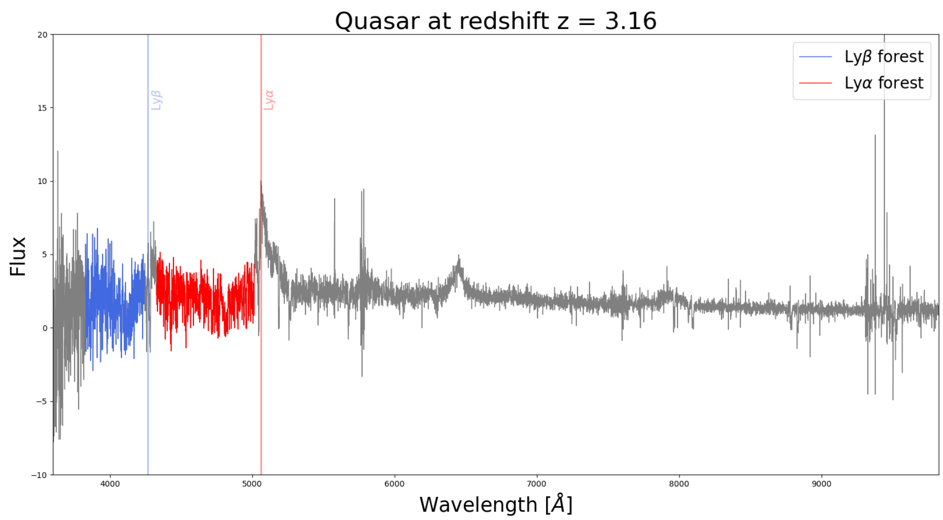


Figure 6.1: Spectrum of a quasar at a redshift of 3.16, where the Lyman-β emission line a rest-frame wavelength of $\lambda_{Ly\beta} = 1025.18\mathring{A}$, and the Lyman-$\beta$ forest are shown in blue, and the Lyman-α emission line a rest-frame wavelength of $\lambda_{Ly\alpha} = 1215.67\mathring{A}$ and the Lyman-α forest are shown in red.

6.2 Measurements of Cross-Correlations of Quasars and the Lyman β Forest

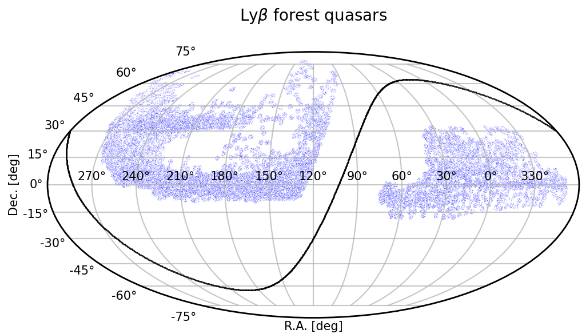
Sections 3.2.1 and 6.1 explored the nature of the Lyman-α and Lyman-β transitions and forests. This section presents the computation of cross-correlations between the density fluctuations $\delta_{Ly\beta}$ and quasar positions.

Quasar Catalog for the Lyman-β forest analysis

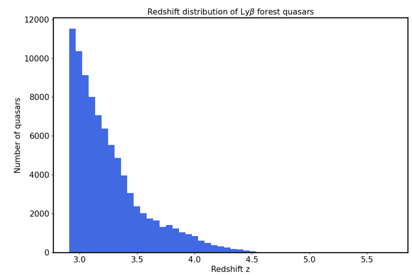
Figure 6.2 shows the spacial distribution and redshift of the quasars with Lyman-β forests, used to compute cross correlations with the aforementioned absorption.

Correlation function

The cross-correlation function of the Lyman-β forest with quasars at a scale of $-60\text{Mpc}/h < r < 60\text{Mpc}/h$ is presented in figure 6.3. It is displayed in four panels, a, b, c, and d, each representing a different slice (see 4.2.2), different ranges of $r_\perp$ (often referred to as *slices*), which cover $0\text{Mpc}/h < r_\perp < 2\text{Mpc}/h$, $4\text{Mpc}/h < r_\perp < 6\text{Mpc}/h$, $8\text{Mpc}/h < r_\perp < 10\text{Mpc}/h$, and $12\text{Mpc}/h < r_\perp < 14\text{Mpc}/h$.

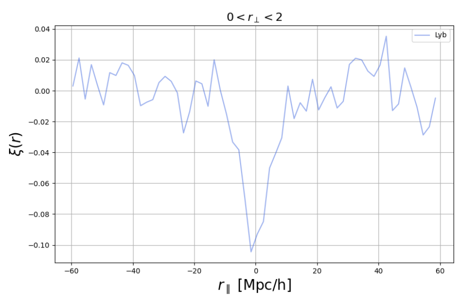


(a)

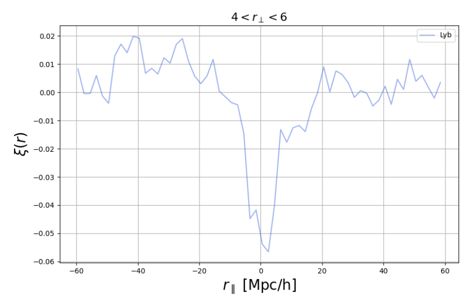


(b)

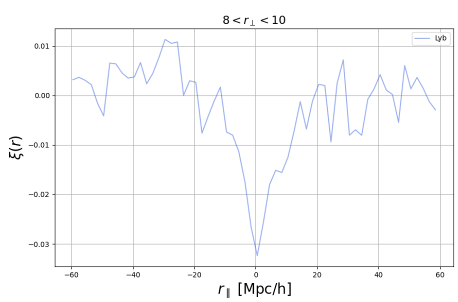
Figure 6.2: (a) Footprint of the quasar catalog that conforms the sample of Lyman-β forests of DESI DR1 used for this analysis, and (b) histogram of z distribution of such sample.



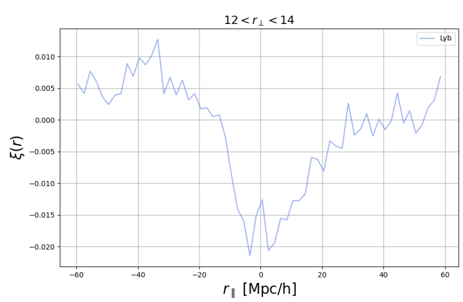
(a)



(b)



(c)



(d)

Figure 6.3: Cross-correlation function of the Lyman-β forest with quasars at a scale of $-60\text{Mpc}/h < r < 60\text{Mpc}/h$. The four panels represent a different slice: (a) $0\text{Mpc}/h < r_\perp < 2\text{Mpc}/h$, (b) $4\text{Mpc}/h < r_\perp < 6\text{Mpc}/h$, (c) $8\text{Mpc}/h < r_\perp < 10\text{Mpc}/h$, and (d) $12\text{Mpc}/h < r_\perp < 14\text{Mpc}/h$.

6.3 Best Fit Parameters

This section shows the best fit model to the quasar and Lyman-β computed cross-correlation. This fitting process was carried out using the software VEGA, a tool specifically developed for modeling the 3D correlation functions between quasars and intergalactic absorbers, such as the Lyman-α forest. VEGA constructs both the theoretical model and the corresponding likelihood function, enabling parameter estimation either through likelihood minimization with IMINUIT or posterior sampling using POLYCHORD. It is designed to interface with data products from PICCA, and is primarily used to extract cosmological information (especially BAO signals) from quasar–absorption cross-correlation measurements [134], [135], [78]. We estimate the value of the bias b, letting this as the only free parameter during the fitting process, while all the rest[1] remain fixed to avoid degeneracies. Figure 6.4 shows the best fitted model (in red) to the computed cross-correlation (in blue). The computed value for the bias, given as $b\eta = f/b$ with b the linear bias and f the growth factor fixed at $f =$, is $b\eta_{Ly\beta} = -0.0344 \pm 0.0012$.

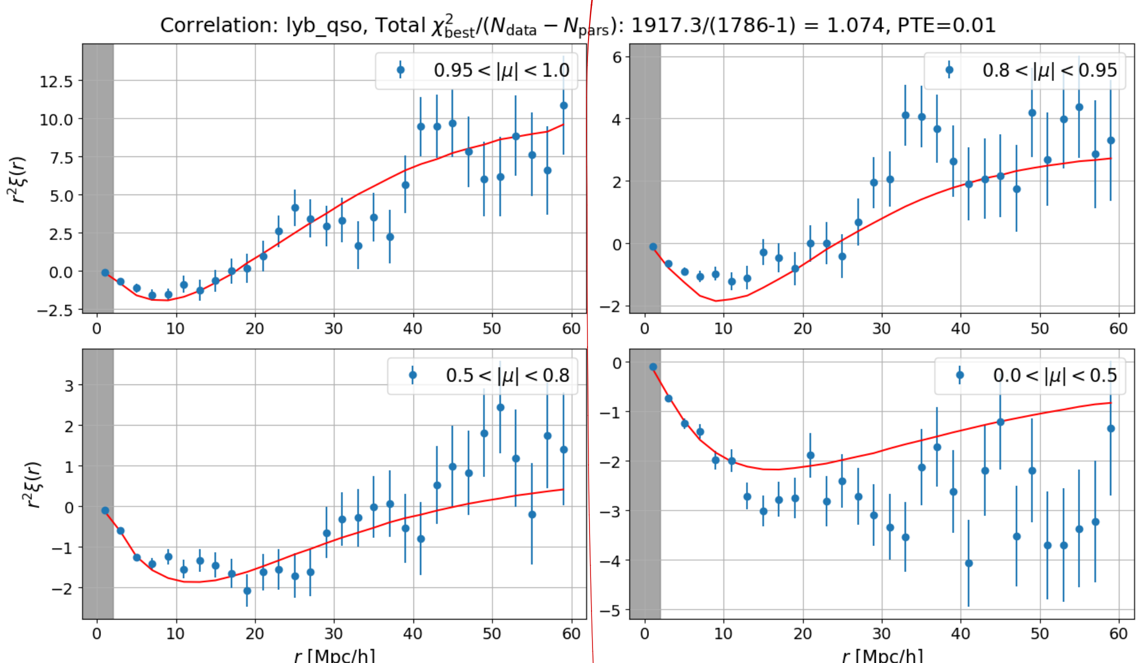


Figure 6.4: Cross correlation of the Lyman-β forest with quasars, shown in the wedges, from top left panel to bottom right panel: $0.95 < |\mu| < 1.0$, $0.8 < |\mu| < 0.95$, $0.5 < |\mu| < 0.8$, and $0 < |\mu| < 0.5$. The bias found for the Lyman-β forest is $b\eta_{Ly\beta} - 0.0344 \pm 0.0012$

ya se encontraron los best-fit ... and then?

que concluyes?

info nueva, que sigue?

p' q' sirve?

[1] like α, $a_\parallel$, $a_\perp$, β, etc.

Chapter 7

The SiIV, CIV, and MgII Forests as Large Scale Structure Tracers

The BAO scale has been measured with several tracers that probe the history of the expansion of the Universe. At lower redshifts ($z < 1.5$) it has been measured with galaxies and quasars. At higher redshifts ($z > 2$) the number of these tracers is much lower, which limits their efficacy for high-precision measurements. As a result, measurements are conducted using the IGM traced by different absorbers. Since hydrogen is the most abundant element conforming the IGM, the main absorber for the BAO measurement is Lyman-α. Nonetheless, the Lyman-α forest is measurable at $z > 2$, while galaxies and clustering quasars are good targets to measure the BAO scale at redshifts below $1.5 < z$?cite [?] Complete SDSS, see Calum's[3]. As a consequence, there is an interval of z with less tracers. The analysis of metal forests allows to study structure and its evolution in this interval of z. Figure 7.1 shows in the far right panel how the metallicity of gas evolves with redshift, which makes it clear that the study of metallicity of the IGM at different evolutionary stages of the Universe is important for the understanding of the evolution history of the Universe and structure.

7.1 Absorber forests

There are several different atomic transitions used to trace these density fluctuations. In general, the continuum of absorptions of a given transition, tracing the overall fluctuations of matter density is called a forest. The forests are generated by a set of absorptions of the the atomic transitions occurring at different z (the z at which the different overdensities of IGM, where these transitions occur, are located). Here we analyze three different forests of particular interest for cosmology: Si IV, C IV, and Mg II. In an analogous way as in the Lyman forests, other atomic transitions can generate their corresponding forests. Depending on the intensity of these transitions and the S/N in which they are detected, they can be used as LSS tracers. Table 7.1 shows the wavelength range of each of them, and figure 7.2 illustrates a spectrum of a quasar at $z = 2.63$ from DESI DR1 that contains all the aforementioned forests.

As opposed to the nature of the Lyman-α and Lyman-β forests, that are produced by singlets, the forests explored here are composed by doublets. As a consequence, the measured 3D cross-correlation functions (defined and described in section 4.2) are the superposition of two correlations

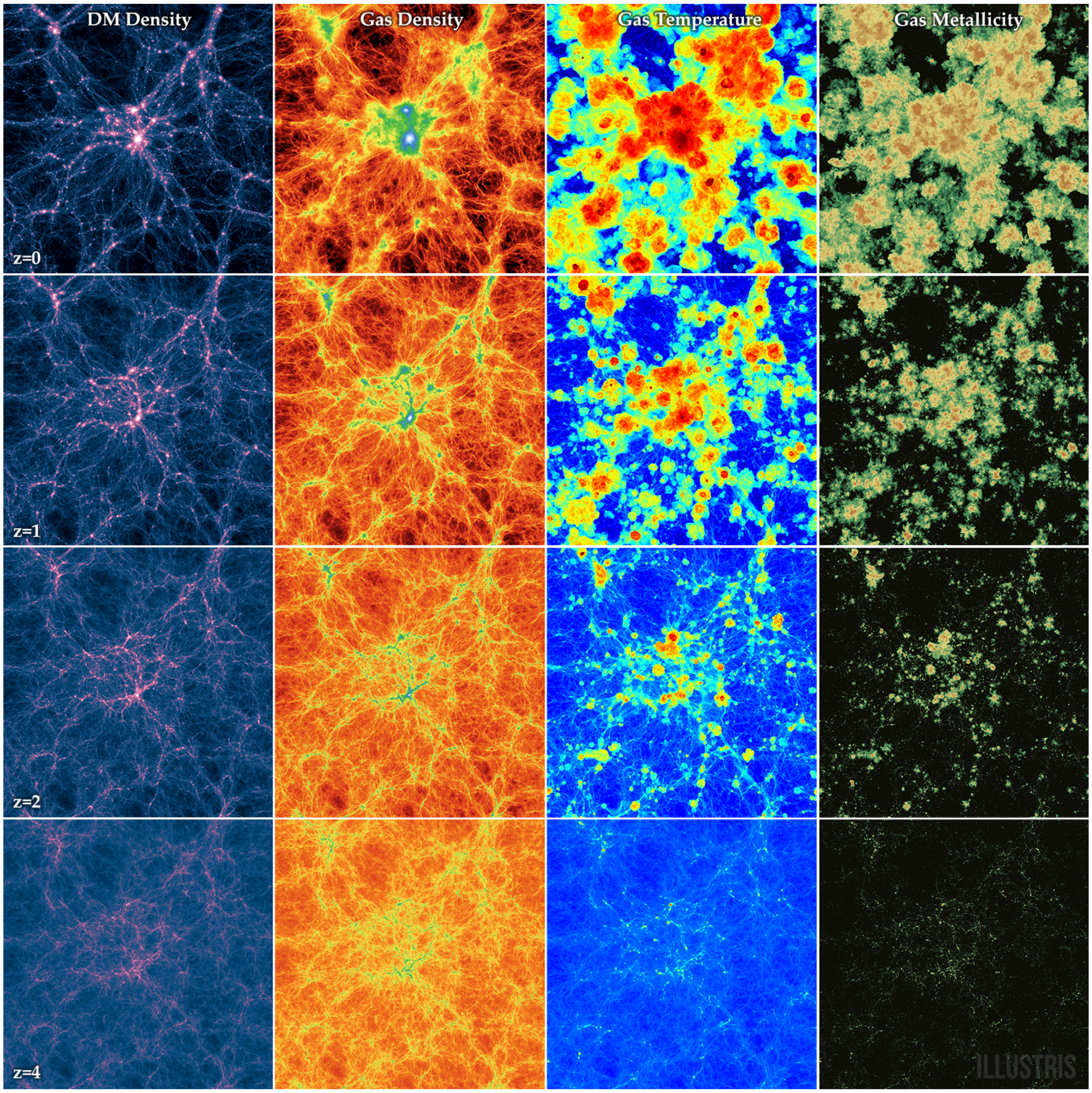


Figure 7.1: Redshift evolution of a box slice of the Illustris simulation, from $z = 4$ to $z = 0$, showing four projections: dark matter density, gas density, gas temperature, and gas metallicity [136].

separated by a distance Δr_m, where m represents the different forests, and with a different bias. For the best-fit analyses of correlations with doublet nature transitions, an *effective wavelength*, which is an intensity weighed average $\lambda_{\rm eff}$ is defined. The effective wavelengths $\lambda_{\rm eff_{SiIV}}$, $\lambda_{\rm eff_{CIV}}$, and $\lambda_{\rm eff_{MgII}}$ are shown in table 7.1

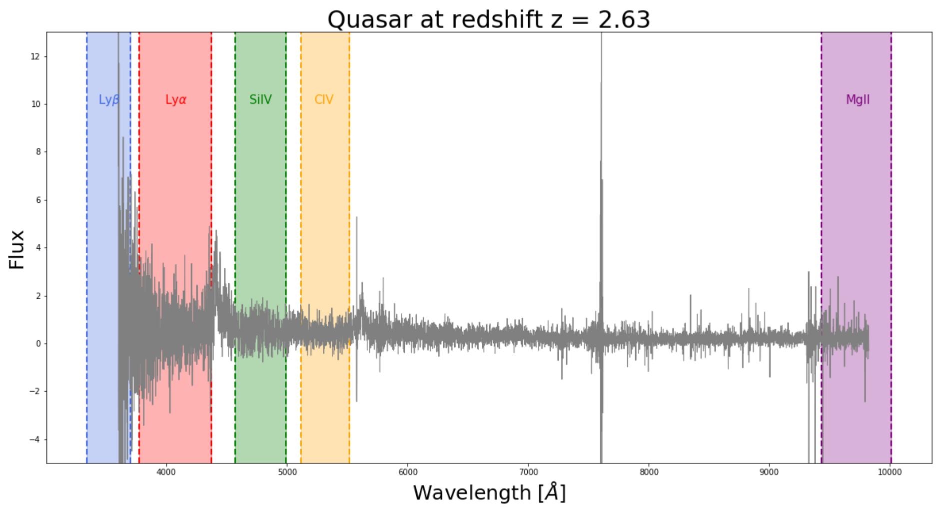


Figure 7.2: DESI DR1 quasar at $z = 2.63$ with the absorber forests in the following colors: Lyman-β in blue, Lyman-α in red, Si IV in green, C IV in yellow, and Mg II in purple.

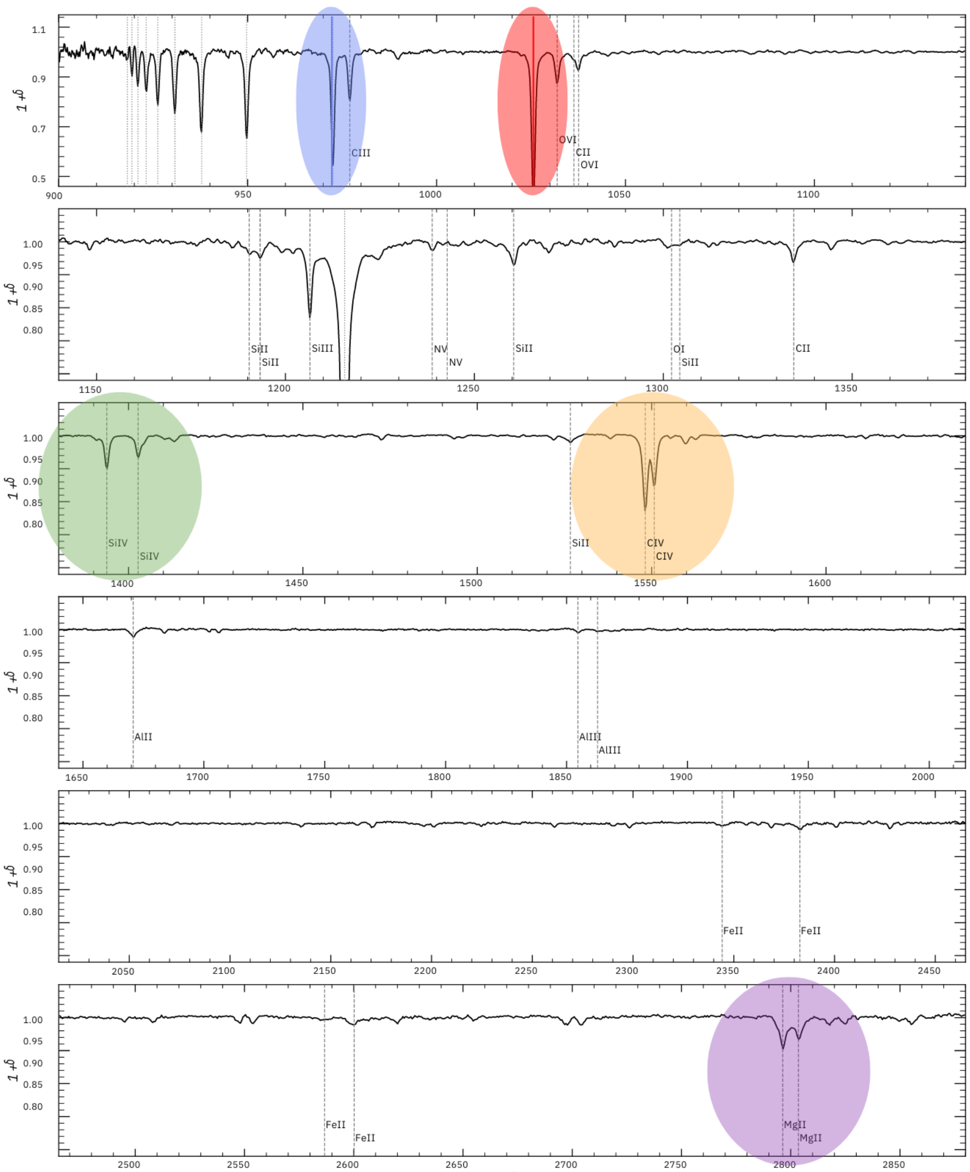
Forest	$\lambda_{\min}[\AA]$	$\lambda_{\max}[\AA]$	$\lambda_a[\AA]$	$\lambda_b[\AA]$	$\lambda_{\text{eff}}[\AA]$
Ly-β°*	920	1020	1025.72	-	-
Ly-α*	1040	1205	1216	-	-
Si IV°*	1260	1375	1402.77	1393.76	1396.76
CIV°*	1410	1520	1550.78	1548.20	1549.06
Mg II*	2600	2760	2803.53	2796.35	2798.75

Table 7.1: Metal absorption forests analyzed in this study. The first column lists the atomic transitions that define each forest. The second and third columns give the minimum and maximum rest-frame wavelengths of the corresponding forest regions. The fourth and fifth columns show the rest-frame wavelengths of the relevant absorption features—either singlets or doublets—associated with each forest. The fifth column is the effective wavelength, an intensity weighed average of the doublet that acts as a singlet for the fitting of the model in the correlations. The definitions of the forests are defined in [137] (*), and [84] (°).

Forest	Scale r [Mpc/h]
Ly-β	$-60 < r < 60$
Si IV	$-60 < r < 60$
C IV	$-40 < r < 40$
Mg II	$-60 < r < 60$

Table 7.2: Scales in Mpc/h in which the correlation functions were computed.

not mentioned



not mentioned

Figure 7.3: Stacked quasar spectra observed by eBOSS [138]. The highlighted areas show the absorption that cause the different forests: circled in blue is the Lyman-β absorption, in orange the Lyman-α, in green the SIV, in yellow the C IV, and in red Mg II. The doublet nature of the latter three transitions should be easily noted.

7.1.1 CIV (triply ionized carbon) forest

The C IV forest (at rest frame wavelengths $\lambda = 1420\mathring{A}$ to $\lambda = 1520\mathring{A}$), illustrated in figure 7.5, is formed by the doublet caused by the electronic transition from the level $2s^2S - 2p^2p^0$, where photons are emitted at wavelengths of $\lambda_a = 1548.2\mathring{A}$, and $\lambda_b = 1550.8\mathring{A}$[1], as shown in figure 7.4 and it is one of the most populated and the strongest of the metal forests observed in quasar spectra. This makes it a good target to analyze the LSS in the Universe at redshift $z > 1.4$, as it also offers a valuable tool to conduct cosmological analysis at these redshifts with IGM absorption features [85]. Carbon is considerably less abundant than hydrogen in the Universe. As a consequence, the observation and study of the C IV forest present additional challenges, compared to the Lyman-α forest, such as the need of higher S/N for its detection. Nevertheless, since the C IV forest can be observed across a broader redshift range ($1.5 < z < 5.4$), it makes it a relevant an potentially rich LSS tracer. Given that the presence of metals is a consequence of processes intrinsic to galactic evolution, the C IV not only serves as an effective LSS tracer but also provides insights into evolved structures within the observed forest range.

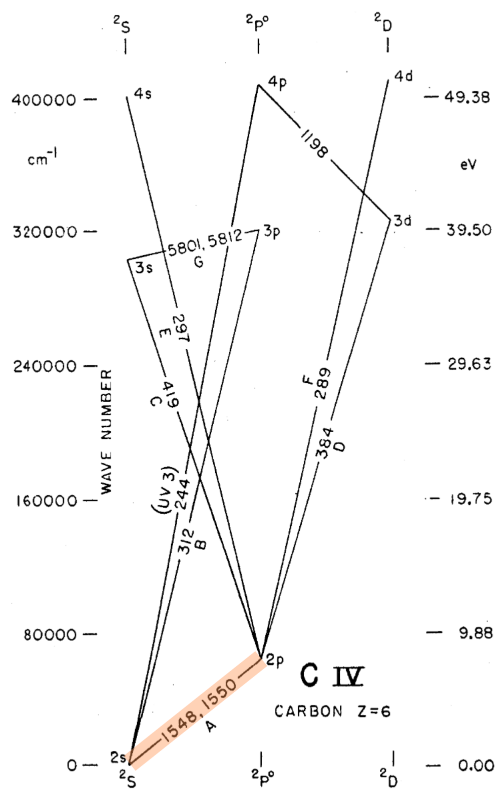


Figure 7.4: Transitions of C IV. The C IV forest analyzed in this work is formed by the doublet caused by the electronic transition from the level $2s^2S - 2p^2p^0$, where photons are emitted at wavelengths of $1548.2\mathring{A}$, and $1550.8\mathring{A}$, highlighted in orange. [139]

[1]Also denoted by $\lambda_{\mathrm{CIV}(1548)} = 1548.20\mathring{A}$ and $\lambda_{\mathrm{CIV}(1551)} = 1550.78\mathring{A}$

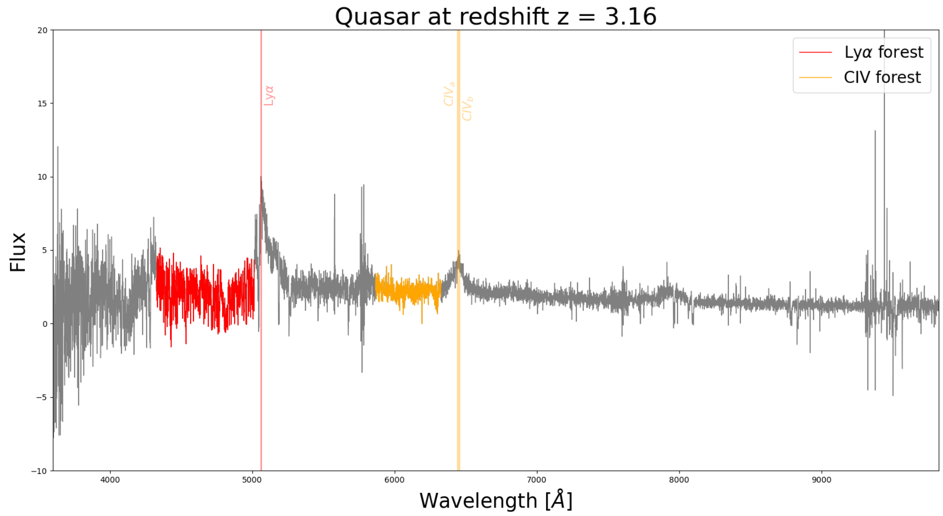


Figure 7.5: Spectrum of a quasar at a redshift of $z = 3.16$ observed by DESI. The C IV emisison lines at rest-frame wavelengths of $\lambda_{\mathrm{CIV}_a} = 1548.2\AA$ and $\lambda_{\mathrm{CIV}_b} = 1550.8\AA$ and the C IV forest are highlighted in orange, and the $\lambda_{\mathrm{Ly}\alpha} = 1215.67\AA$ emisison line and forest are shown in red.

7.1.2 SiIV (Triply ionized silicon) forest

As described in [42], the Si IV forest (at rest-frame wavelengths of $\lambda = 1260\AA$ to $\lambda = 1375\AA$) is made by a doublet composed of $\lambda_{\mathrm{SiIV}_a} = 1393.76\AA$ and $\lambda_{\mathrm{SiIV}_b} = 1402.77\AA$, which takes place when electrons transition from $3s^2S - 3p^2P^0$[2]. It can be observed in the redshift range of $1.85 < z < 6.14$, making it a potentially valuable source of information, despite the challenges provided by lower S/N than C IV. Figure 7.6 shows a spectrum of a quasar at a redshift of $z = 3.16$ observed by DESI. The region of the spectrum highlighted in green corresponds to the Si IV forest.

[2]This transition is analogous to Mg II. The change in energy levels and the structure of figure 7.7 is valid for Si IV, with the corresponding values found in [139]

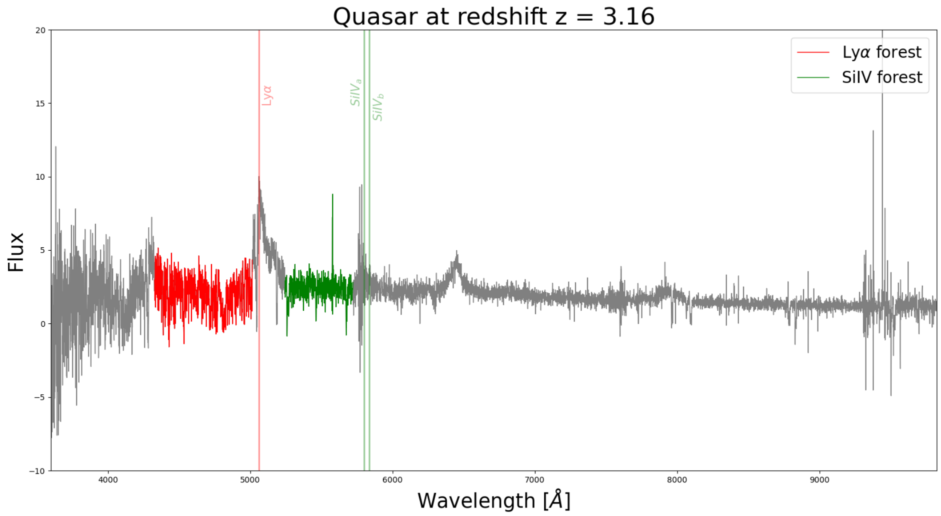


Figure 7.6: Spectrum of a quasar at a redshift of $z = 3.16$ observed by DESI. The Si IV emisison lines at rest-frame wavelengths of $\lambda_{\mathrm{SiIV}_a} = 1393.76\mathring{A}$ and $\lambda_{\mathrm{SiIV}_b} = 1402.77\mathring{A}$ and the Si IV forest are highlighted in green, while the $\lambda_{\mathrm{Ly}\alpha} = 1215.67\mathring{A}$ emission line and the Lyman-α forest are shown in red.

7.1.3 MgII (Singly ionized magnesium) forest

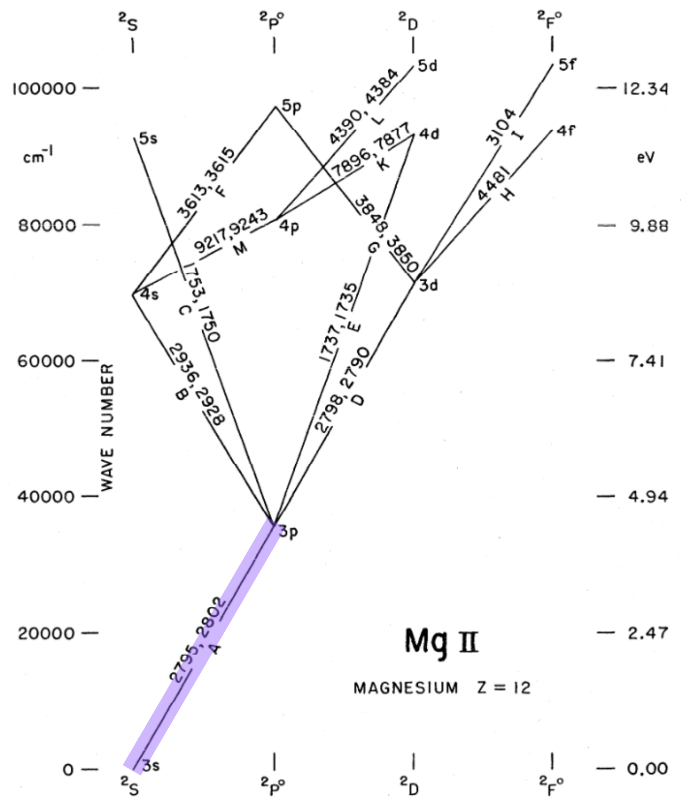


Figure 7.7: Transitions of Mg II. The Mg II forest analyzed in this work is formed by the doublet caused by the electronic transition from the level $3s^2S - 3p^2p^0$, where photons are emitted at wavelengths of $\lambda_{\mathrm{MgII}_a} = 2796.35\mathring{A}$ and $\lambda_{\mathrm{MgII}_b} = 2803.53\mathring{A}$, highlighted in violet

The Mg II forest, at rest-frame wavelengths of $\lambda = 2600\mathring{A}$ to $\lambda = 2760\mathring{A}$, is made by a doublet composed of $\lambda_{\mathrm{MgII}_a} = 2796.35\mathring{A}$ and $\lambda_{\mathrm{MgII}_b} = 2803.53\mathring{A}$, emitted when electrons transition from $3s^2S - 3p^2p^0$, as shown in figure 7.7. It can be observed between redshifts of $0.38 < z < 2.55$, making it a complementary tracer at redshifts than cannot be observed by other forests. It is also a region where, as explained in [84], there is no confusion with other doublets (namely C IV and Si IV). Figure 7.8 shows the spectrum of a quasar at a redshift of $z = 2.21$ observed by DESI. The region of the spectrum highlighted in violet is the Mg II forest.

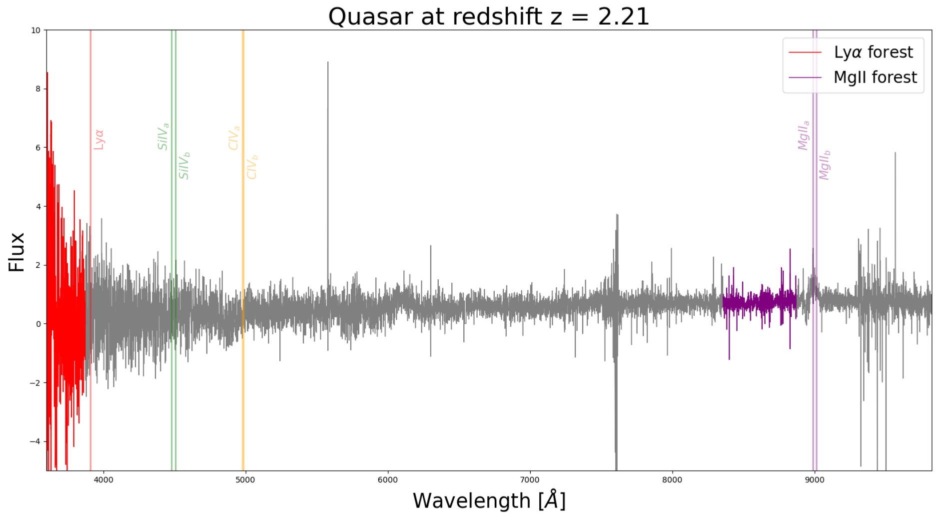


Figure 7.8: Spectrum of a quasar at a redshift of $z = 2.21$ observed by DESI. The Mg II emisison lines at rest-frame wavelengths of $\lambda_{\text{MgII}_a} = 2796.35\mathring{A}$ and $\lambda_{\text{MgII}_b} = 2803.53\mathring{A}$ and the Mg II forest are highlighted in purple, while the $\lambda_{\text{Ly}\alpha} = 1215.67\mathring{A}$ emission line and the Lyman-α forest are shown in red. The Si IV and C IV doublets are shown in green and orange lines respectively.

7.2 Measurements of Cross-Correlations of Quasars and Absorber Forests

Following the methodology explained in 4.2, the three cross-correlations between the absorber (C IV, Si IV, and Mg II) forests and quasar samples were computed. The spatial (footprint and z) distribution of the quasars used for each measurement and the respective correlations are shown in the following subsections. The three different correlations are presented in four panels each. These panels present different slices (ranges of $r_\perp$, see 4.2.2). Since the data analysis pipeline transforms wavelengths to redshifts assuming a given absorption (generally Lyman-α, but potentially defined depending on the given absorption for a specific analysis), such line and a second one[3] occupying the same physical position are absorbed at different apparent redshifts, resulting in an apparent distance separation given by:

$$r_\parallel = (1+z)D_H(z)\Delta\lambda/\lambda_{main}, \tag{7.1}$$

where D_H is the Hubble distance, $\Delta\lambda = \lambda_{sec} - \lambda_{main}$ is the wavelength separation of the secondary transition (λ_{sec}) with respect to the main one [140]. All the forests analyzed in this chapter are doublets, so the *main* wavelength corresponds to the most prominent absorption and the secondary one is the following. Given a redshift, the wavelength difference of these doublets $\Delta\lambda$ transforms into a distance $r_\parallel$ that defines the resolution (binning) in which the correlations are computed. The

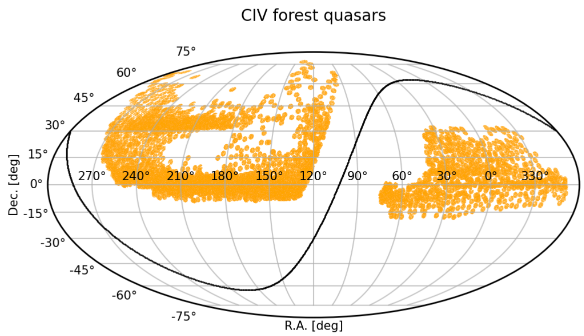
[3]Which could be a metal contaminating the forest, or a doublet

redshift used to make this conversion is the effective redshift z_{eff}, the weighted average redshift of each of the datasets (catalogs described in sections 7.2.1, 7.2.2 and 7.2.3).

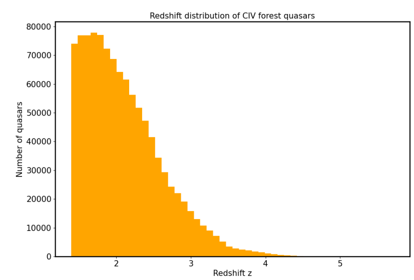
7.2.1 CIV forest analysis

Quasar Catalog sample for the CIV forest analysis

Figure 7.9 shows the footprint and redshift distribution of the quasars with C IV forests, used to compute cross-correlations. This sample will hereafter be referred to as *C IV quasar*[4] *catalog*



(a)



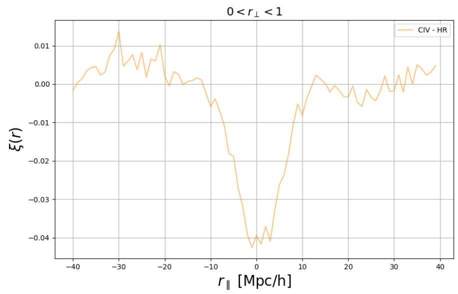
(b)

Figure 7.9: (a) Footprint of the quasar catalog that conforms the sample of C IV forests of DESI DR1 used for this analysis, (b) histogram of z distribution of such sample.

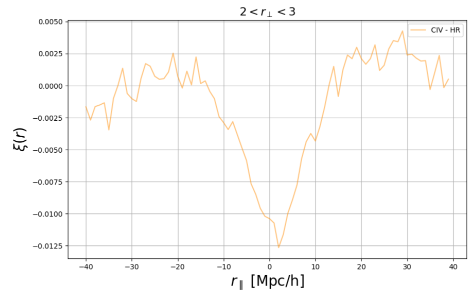
Correlation function of CIV forests and quasars

Figure 7.10 presents the cross-correlation functions between the C IV forest density fluctuations and the C IV quasar catalog. In order to capture the doublet nature of this transition in the correlation function, which is at $r_{\parallel} = 4.85$ Mpc/h at $z_{\mathrm{eff}} = 2.09$, it had to be computed with a high resolution at scales of -40 Mpc/$h < r < 40$ Mpc/h. It is presented in four panels, a, b, c, and d, each representing a different slice (see 4.2.2), which cover 0 Mpc/$h < r_{\perp} < 1$ Mpc/h, 2 Mpc/$h < r_{\perp} < 3$ Mpc/h, 4 Mpc/$h < r_{\perp} < 5$ Mpc/h, and 6 Mpc/$h < r_{\perp} < 7$ Mpc/h.

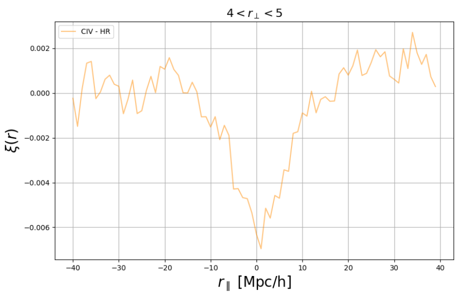
[4] Or C IV-QSO



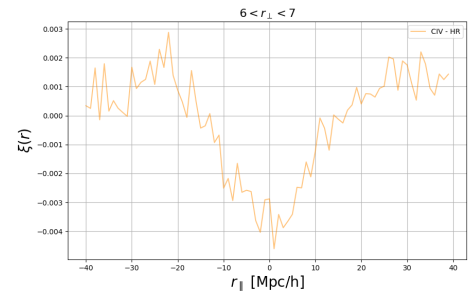
(a)



(b)



(c)



(d)

Figure 7.10: Cross-correlation function of the C IV forest with quasars at a scale of -40 Mpc/h < r < 40 Mpc/h. The four panels represent a different slice: (a) 0 Mpc/$h < r_\perp <$ 1 Mpc/h, (b) 2 Mpc/$h < r_\perp <$ 3 Mpc/h, (c) 4 Mpc/$h < r_\perp <$ 5 Mpc/h, and (d) 6 Mpc/$h < r_\perp <$ 7 Mpc/h.

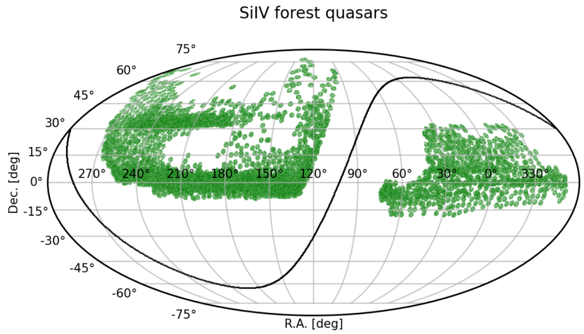
7.2.2 SiIV forest analysis

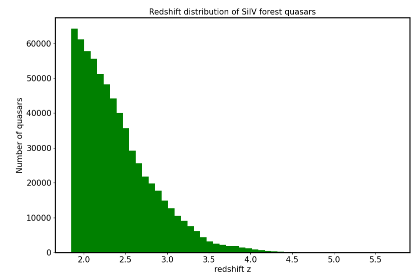
Quasar Catalog sample for the SiIV forest analysis

Figure 7.11 shows the footprint and redshift distribution of the quasars with Si IV forests, used to compute cross-correlations with the aforementioned absorption.

Correlation function of SiIV forests and quasars

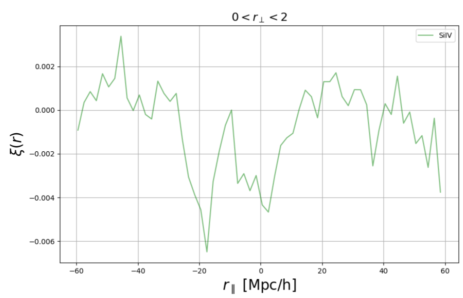
Figure 7.12 presents the cross-correlation functions between the Si IV forest density fluctuations and the Si IV quasar catalog. Since the doublet nature is detected at $r_\parallel = 18.6$ Mpc/h at a $z_{\rm eff} = 2.23$, the correlation was computed at scales of -60 Mpc/$h < r <$ 60 Mpc/h with a standard resolution, and is presented in four panels, a, b, c, and d, each representing a different slice (see 4.2.2) covering 0 Mpc/$h < r_\perp <$ 2 Mpc/h, 4 Mpc/$h < r_\perp <$ 6 Mpc/h, 8 Mpc/$h < r_\perp <$ 10 Mpc/h, and 12 Mpc/$h < r_\perp <$ 14 Mpc/h.

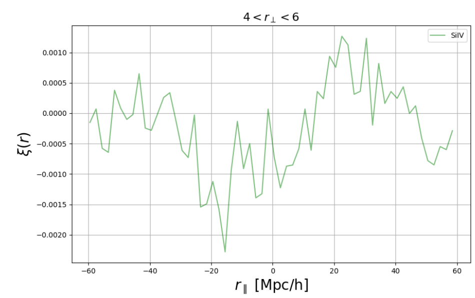




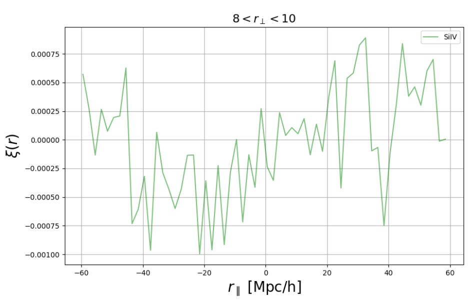
(a) (b)

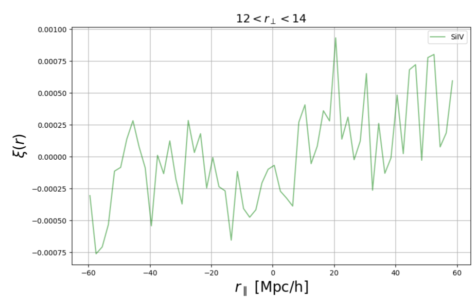
Figure 7.11: (a) Footprint of the quasar catalog that conforms the sample of Si IV forests of DESI DR1 used for this analysis, (b) histogram of z distribution of such sample.





(a) (b)





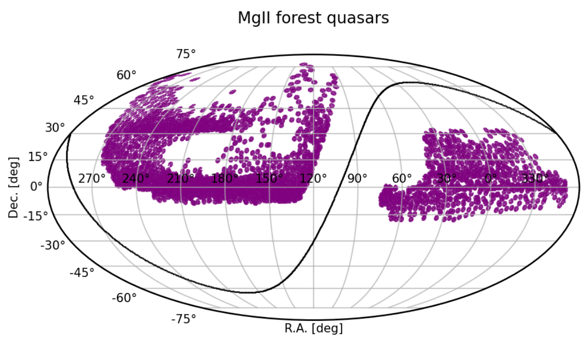
(c) (d)

Figure 7.12: Cross-correlation function of the Si IV forest with quasars at a scale of -60 Mpc$/h < r < 60$ Mpc$/h$. The four panels represent a different slice: (a) 0 Mpc$/h < r_\perp < 2$ Mpc$/h$, (b) 4 Mpc$/h < r_\perp < 6$ Mpc$/h$, (c) 8 Mpc$/h < r_\perp < 10$ Mpc$/h$, and (d) 12 Mpc$/h < r_\perp < 14$ Mpc$/h$.

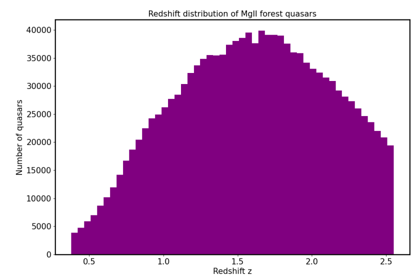
7.2.3 MgII forest analysis

Quasar Catalog sample for the MgII forest analysis

Figure 7.13 shows the footprint and redshift distribution of the quasars with Mg II forests, used to compute cross-correlations with the aforementioned absorption.



(a)



(b)

Figure 7.13: (a) Footprint of the quasar catalog that conforms the sample of Mg II forests of DESI DR1 used for this analysis, (b) histogram of z distribution of such sample.

Correlation function of MgII forests and quasars

Figure 7.14 presents the cross-correlation functions between the Mg II forest density fluctuations and the Mg II quasar catalog. Since the doublet nature is detected at $r_{\parallel} = 7.94$ Mpc/h at a $z_{\rm eff} = 1.71$, the correlation was computed at scales of -60 Mpc/$h < r < 60$ Mpc/h with a standard resolution, and is presented in four panels, a, b, c, and d, each representing a different slice (see 4.2.2) covering 0 Mpc/$h < r_{\perp} < 2$ Mpc/h, 4 Mpc/$h < r_{\perp} < 6$ Mpc/h, 8 Mpc/$h < r_{\perp} < 10$ Mpc/h, and 12 Mpc/$h < r_{\perp} < 14$ Mpc/h.

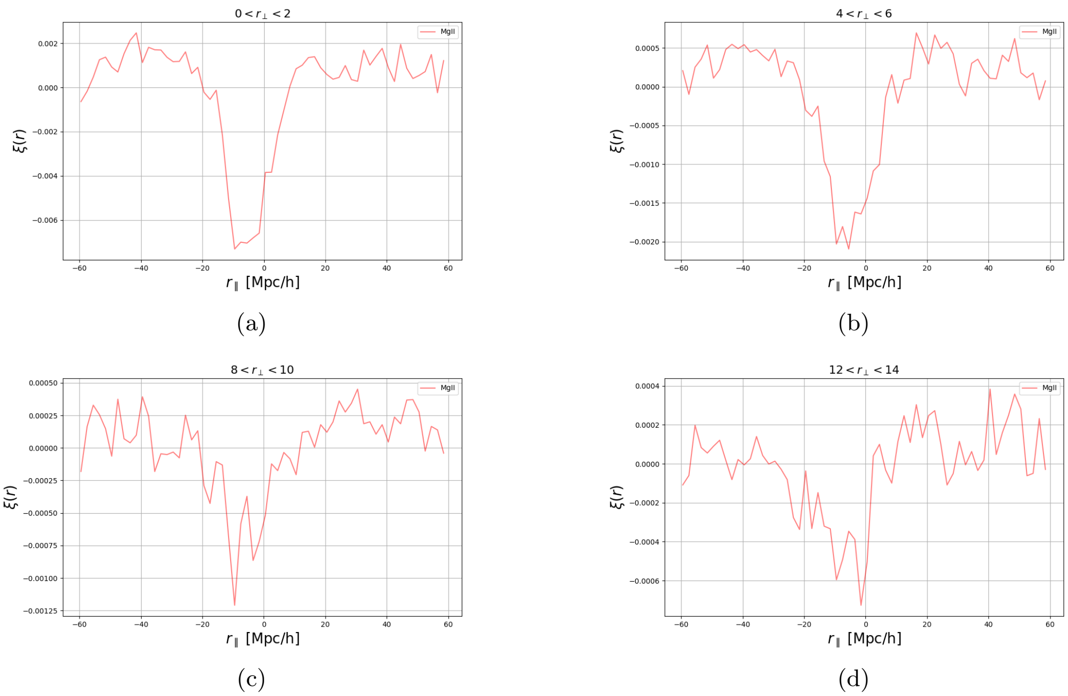


Figure 7.14: Cross-correlation function of the Mg II forest with quasars at a scale of -60 Mpc/$h < r < 60$ Mpc/h. The four panels represent a different slice: (a) 0 Mpc/$h < r_\perp < 2$ Mpc/h, (b) 4 Mpc/$h < r_\perp < 6$ Mpc/h, (c) 8 Mpc/$h < r_\perp < 10$ Mpc/h, and (d) 12 Mpc/$h < r_\perp < 14$ Mpc/h.

7.3 Best Fit Parameters

explicar

The obtention of the best-fit model for the cross-correlations between the absorber (Si IV, C IV, Mg II) forests and quasars is ongoing work. The approach follows the procedure used for the Lyman-β forest in 6.3, where the bias parameter $b\eta$ is computed using the VEGA code. However, adapting this methodology to these absorbers requires further refinements in the fitting process, as well as additional computational and human effort.

As mentioned before, the doublet nature of these transitions causes the correlations to behave as the superposition of two signals separated by a fixed distance. The modeling of these features within the fitter involves treating the main line as the absorber that defines the forest, while the secondary one is considered a contaminant. The implementation of this approach is currently in progress, and results are expected in the fall of 2025.

Se calculan funciones de correlacion, pero hay algo inconcluso, solo fue calcularlas?

Existe alguna publicación?

explicar / separar tus contribuciones

Chapter 8

Discussion and Conclusions

cita [ ]

Potential of the Lyman-β and CIV forests as BAO tracers

Even though the analysis performed for this research covers a scale up to 60Mpc/h, the size and resolution of the datasets used for the Lyman-β and C IV forest clustering for DR1 analysis are promising about their potential use for the detection of the BAO features. The Lyman-β cross-correlation functions presented in chapter 6 are the first ever measured correlations using the Lyman-β forest using it as a tracer of LSS of the IGM. To ~~better understand~~ how the Lyman-α and Lyman-β forests trace the underlying matter distribution, their respective linear biases are compared. The cross-correlation between Lyman-β forest QSO bias obtained in this work, reported in 6.4, is $b_\beta = -0.0344 \pm 0.0012$, while the Lyman-$\alpha$ x QSO bias is $b_\alpha = -0.099^{+0.015}_{-0.013}$ [75]. These values are reported in 8.1. The ratio of these two values is

$$\frac{b_\beta}{b_\alpha} \approx 0.35. \tag{8.1}$$

This indicates that the Lyman-β forest is a significantly weaker LSS tracer than the Lyman-α forest. This is consistent with the strength difference of both transitions, and the generally lower S/N in the B region. Nonetheless, it would be interesting to compare this value using synthetic datasets, where systematics and astrophysical effects influencing the measurement of correlations with both tracers can be better controlled. The smaller absolute value of the bias suggests that Lyman-β absorption features are less strongly correlated with the underlying matter density and may originate from regions with lower average overdensity. A lower bias implies a lower clustering amplitude. Contrary to what one might initially conclude, this does not make the Lyman-β forest less competitive. On the contrary, it highlights its potential as a complementary tracer to Lyman-α in multi-tracer analyses. In addition to offering independent information, the Lyman-β forest can probe redshift ranges where Lyman-α alone becomes less effective. As analyses using this tracer become more detailed, it will be important to consider that the Lyman-β forest is more prone to contamination from Lyman-α absorption of lower-redshift systems. Therefore, dedicated efforts to characterize these contaminants—along with others such as O VI—will be crucial in enabling its reliable use as a cosmological tracer.

Part of the investigation done during the PhD program consisted on testing this for the C IV forest in a forecasted synthetic dataset of the final DESI survey, made by the codes CoLoRe

Parameter	Combined	Absorber x Absorber	Absorber x QSO
b_α	$-0.1078^{+0.0045}_{-0.0054}$	-0.1078 ± 0.003	$-0.099^{+0.015}_{-0.013}$
b_β	-	-	-0.0344 ± 0.0012

Table 8.1: Comparison of the biases of Lyman-α (b_a) and Lyman-β (b_β). The first row are the best-fit values (mean of the posterior) and uncertainties (68% credible intervals) for the Lyman-α bias of DESI DR1 [75]. The second one is the value of the bias of Lyman-β x QSO obtained in this work, reported in 6.4.

explicar

and LyaCoLoRe [141], [142], [143], [128]. This analysis began during the commisionning stage of the survey, with an early version of PICCA that has required several modifications, making it an ongoing research developed parallelly to the one presented in this thesis. Nonetheless, the amount of data and the resolution used for the binning of DR1 data is starting to present several computational challenges. As a first stage, the strategy was to prioritize the characterization and understanding of the C IV doublet, resolving it in the correlations and modelling it properly in the fitting before stepping into the challenges of the BAO scales found during the work with synthetic datasets. The following steps are to extend the scale of the analysis to cover the BAO scale, as well as modify accordingly PICCA to perform the analysis efficiently with the highest possible resolution. It is important to note that this work faces several interesting challenges moving forward. One key assumption so far is that the relation between optical depth τ and matter overdensity δ_m follows the form given in equation 4.2. While this relation is well motivated for hydrogen, it does not necessarily hold for other elements. Although we adopt this approximation for the additional transitions analyzed, it should be emphasized that this represents a first approach. Future refinements are both possible and planned, as part of ongoing efforts to improve the physical accuracy of the model.

Modeling the doublet

As explained in section 7.1, since Si IV, C IV, and Mg II forests are doublets and this impacts significantly the shape of the correlation functions, this behavior has to be well understood and characterized. This affects two lines of the analysis: one is in the computation of the correlation function, the second one is in the bias fitting. While the doublets do not affect the computation of the correlation directly, but their shape, it is convenient to define a reference wavelength for the analysis. This is to choose the reference absorption feature, which can be the main, the secondary or an effective (weighted average) wavelength. This will only affect the center of the correlations, potentially shifting it slightly toward negative or positive values of r. In the case of the fitting process, the definition of the reference wavelength has a more transcendental impact. The current analysis defines the main absorption of the doublet as the reference, and the secondary one is treated as a contaminant. This will allow the computation of both biases (as free parameters together or separately). Alternative, one could use a different criteria which could be more convenient according to the approach to the analysis (like choosing as the main line the one in the blue side of the spectrum, or the weighted-average wavelength).

Improvement of Catalogs

The first attempt to build a catalog of each tracer followed the methodology in [137][1]. The analysis was made with the datasets obtained filtering by redshift and wavelength the quasars, so that the computations could be performed with the respective forests. Since these catalogs were constructed, the amount of data taken by the DESI instrument has increased significantly. This has led to a more exhaustive discussion among a larger group of people to re-define the new catalogs for the current and future analyses. While these new criteria do not impact state-of-the-art analyses, their refinement with a larger dataset could introduce some effects. However, these are not expected to significantly alter the results in this line of research.

Cosmological model

As explained in section 2.4.2, under *Alternative Models*, DESI DR1 results tend towards a $w_0 w_a$ cosmology. The results derived from the Lyman-α forest do not strongly favor either the $w_0 w_a$ or the ΛCDM model, since both models fit the Lyman-α forest data similarly well. This is illustrated in figure 2.8. Most of the forests analyzed in this research meet the same criteria, since the effect of having one model or the other one will affect the observations at much smaller redshifts. Nonetheless, the analysis of Mg II (particularly, but could be generalized to any forest that could help constraining the cosmological model), where $z_{\mathrm{eff}_{\mathrm{MgII}}} = 1.71$, could potentially be affected by the cosmological model, which could make it a very interesting forest to analyze with future datasets. This would require two approaches: observational data analysis with computational tools capable of performing their study, and synthetic datasets created with both, the standard cosmological model, and a different cosmological model (namely $w_0 w_a$) as priors in order to confront them and have a better control of the framework so that it can be re-studied. The same would apply if an alternative model were to gain broader acceptance by the survey, or if the ΛCDM model were to be discarded.

Future

Undoubtedly, this work is considered to have a lot of potential serving as a door to new research within the area of cosmology with absorbers. There are several lines of research considered to be very interesting that are planned to be followed:

- ⋆ Perform a thorough analysis on the Lyman-β forest, both as a LSS tracer using correlations at scales of $|r| < 60$ Mpc/h for clustering purposes and at scales that include the BAO feature. This analysis will most likely include the study of contaminants (namely Lyman-α and O VI absorption), following the work of [144], [145], [128], [78], and [140]. This will most likely be done with the DESI DR2 dataset currently being analyzed.

- ⋆ Extend the analysis of C IV, Si IV, and Mg II forests to scales that include the BAO feature with DESI DR1 and DR2 datasets.

[1] This work presents and validates a catalog of Lyman-α forest flux-transmission fluctuations for 3D analyses, based on 88,511 quasars from the Survey Validation (SV) phase of DESI and the first two months of the main survey (M2), as part of the Early Data Release (EDR).

- ⋆ Once again, following the work of [144], [145], [128], [78], and [140], study the effect of C IV, Si IV, Mg II, and O VI in the Lyman-α forest flux correlation function, as well as their biases in BAO and other studies.

- ⋆ Measure the bias of all the forests mentioned at different redshifts to indirectly infer the abundance of metal elements and gain insights into the evolution of cosmic structure and chemical enrichment across epochs. Such analysis can enhance our understanding of the metal enrichment of the IGM, as well as the physical processes responsible for it - such as AGN feedback, structure formation, star formation, and galaxy formation and evolution in the early Universe. Achieving this will likely require the development of additional tools, including realistic IGM simulations to better capture the dynamics and properties of the enriched IGM. Furthermore, complementary studies with galaxy observations are essential. These include cross-correlations between metal absorbers and galaxy distributions, analyses of galaxy properties such as metallicity and star formation rate, circumgalactic medium (CGM) studies using quasar LOS, and investigations of feedback signatures in galaxy spectra. Such approaches can help establish direct connections between galaxy evolution and IGM enrichment, providing a more comprehensive picture of the processes shaping the high-redshift Universe. As explained in Chapter 4, the Gunn–Peterson effect provides an approximation to relate the distribution of dark matter to neutral hydrogen in the IGM. In particular, under the FGPA (see equation 4.2), the optical depth of hydrogen absorption is modeled as a power-law function of the underlying matter density. This framework is valid as a first-order approach on large scales and at redshifts $z \approx 2 - 5$, where the hydrogen in the IGM is highly ionized but still produces measurable absorption in quasar spectra. While this approximation has proven effective in linking gas and matter distributions via the Lyman-α forest, its accuracy diminishes at smaller scales and in regions where baryonic physics, temperature fluctuations, and reionization effects become significant. The inclusion of additional absorption lines from metal ions like as C IV, Si IV, and Mg II can provide complementary information that could help to constrain deviations from the FGPA and refine our understanding of the matter–gas connection.

- ⋆ In previous chapters it was established that the model used for this analysis was the ΛCDM framework. Nonetheless, throughout the development of this work the cosmological results of DESI DR1 and DR2 show a preference towards a dynamical dark energy model, specifically the $w_0 w_a$ parameterization. Figure 2.8 was used to illustrate that the Lyman-α forest results, at $z < 2$, did not seem to favor either of these models. Nonetheless, this may not hold at lower redshifts, where additional metal absorption forests become visible. Employing these forests for cosmological analysis would require more detailed studies, as the growth of fluctuations may exhibit different behaviors across redshifts. This highlights the need for improved simulations that include metals not merely as contaminants, but as tracers of the underlying LSS. Such an approach could yield valuable insights into the redshift evolution of the dark energy equation of state, offering independent constraints from distinct, yet complementary, tracers across a broad range of cosmic epochs.

- ⋆ Estimation of cosmological parameters from all the forests analyzed in this research, scheduled for Fall 2025.

Bibliography

[1] K. Arun, S. Gudennavar and C. Sivaram, Dark matter, dark energy, and alternate models: A review, Advances in Space Research **60** (2017) 166–186.

[2] CERN, Breaking new ground in the search for dark matter, 2024.

[3] The Dark Energy Survey Collaboration, The Dark Energy Survey, arXiv e-prints (2005) astro [astro-ph/0510346].

[4] T. M. C. Abbott, M. Aguena, A. Alarcon, S. Allam, O. Alves, A. Amon et al., Dark Energy Survey Year 3 results: Cosmological constraints from galaxy clustering and weak lensing, **105** (2022) 023520 [2105.13549].

[5] DESI collaboration, A. e. a. Aghamousa, The DESI Experiment Part I: Science,Targeting, and Survey Design, 1611.00036.

[6] DESI collaboration, A. e. a. Aghamousa, The DESI Experiment Part II: Instrument Design, 1611.00037.

[7] Ž. Ivezić, S. M. Kahn, J. A. Tyson, B. Abel, E. Acosta, R. Allsman et al., LSST: From Science Drivers to Reference Design and Anticipated Data Products, **873** (2019) 111 [0805.2366].

[8] R. Laureijs, J. Amiaux, S. Arduini, J. L. Augères, J. Brinchmann, R. Cole et al., Euclid Definition Study Report, arXiv e-prints (2011) arXiv:1110.3193 [1110.3193].

[9] D. Spergel, N. Gehrels, C. Baltay, D. Bennett, J. Breckinridge, M. Donahue et al., Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report, arXiv e-prints (2015) arXiv:1503.03757 [1503.03757].

[10] L. Wenzl, C. Doux, C. Heinrich, R. Bean, B. Jain, O. Doré et al., Cosmology with the roman space telescope – synergies with cmb lensing, Monthly Notices of the Royal Astronomical Society **512** (2022) 5311–5328.

[11] NOIRLab, ¿qué tan rápido se está expandiendo el universo?, 2021.

[12] N. Drake, Detectadas las posibles primeras estrellas del universo, National Geographic (2018) .

[13] D. Collaboration, Cosmology and dark energy, 2025.

[14] A. Muñoz-Gutiérrez, A. de la Macorra and Y. Olguín Trejo, Introducción a la cosmología, 2021.

[15] E. W. Kolb and M. S. Turner, The Early Universe, vol. 69. 1990.

[16] H. Mo, F. C. van den Bosch and S. White, Galaxy Formation and Evolution. May, 2010.

[17] M. Vargas-Magaña, J. E. Bautista, J.-C. Hamilton, N. G. Busca, Aubourg, A. Labatie et al., An optimized correlation function estimator for galaxy surveys, Astronomy and Astrophysics **554** (2013) A131.

[18] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini et al., Planck2018 results: Vi. cosmological parameters, Astronomy and Astrophysics **641** (2020) A6.

[19] COBE collaboration, G. F. e. a. Smoot, Structure in the COBE differential microwave radiometer first year maps, Astrophys. J. **396** (1992) L1.

[20] Boomerang collaboration, C. B. Netterfield et al., A measurement by Boomerang of multiple peaks in the angular power spectrum of the cosmic microwave bac Astrophys. J. **571** (2002) 604 [astro-ph/0104460].

[21] B. Rabii et al., MAXIMA: A Balloon - borne Cosmic Microwave Background anisotropy experiment, Rev. Sci. Instrum. **77** (2006) 071101 [astro-ph/0309414].

[22] M. C. Runyan et al., The Arcminute Cosmology Bolometer Array Receiver, Astrophys. J. Suppl. **149** (2003) 265 [astro-ph/0303515].

[23] WMAP collaboration, C. L. Bennett et al., The Microwave Anisotropy Probe (MAP) mission, Astrophys. J. **583** (2003) 1 [astro-ph/0301158].

[24] Planck collaboration, J. Tauber, M. Bersanelli, J. M. Lamarre, G. Efstathiou, C. Lawrence, F. Bouchet et al., The Scientific programme of Planck, astro-ph/0604069.

[25] European Space Agency, From an almost perfect universe to the best of both worlds, 2018.

[26] M. P. I. for Astrophysics, Millennium-ii simulation, 2009.

[27] European Space Agency, The cosmic budget of 'ordinary' matter, 2018.

[28] K.-G. Lee, A. Krolewski, M. White, D. Schlegel, P. E. Nugent, J. F. Hennawi et al., First Data Release of the COSMOS Lyα Mapping and Tomography Observations: 3D Lyα Forest Tomograpl The Astrophysical Journal Supplement Series **237** (2018) 31 [1710.02894].

[29] COSMOS collaboration, P. Hess, A. Rouillard, A. Kouloumvakos, P. C. Liewer, J. Zhang, S. Dhakal et al., First Data Release of the COSMOS Lyα Mapping and Tomography Observations: 3D Lyα Forest Tomograpl Astrophys. J. Suppl. **237** (2018) 31 [1710.02894].

[30] T. Müller, Clamato – igm tomography, 2017.

[31] S. Courteau et al., Galaxy Masses, Rev. Mod. Phys. **86** (2014) 47 [1309.3276].

[32] E. Consortium, Mapping the dark universe with gravitational weak lensing, n.d.

[33] D. Collaboration, The desi science mission, n.d.

[34] M. White, Mini-lecture on redshift space distorsions, n.d.

[35] J. Peacock, "Large-scale structure and matter in the universe." Presentation at the Royal Society, January, 2003.

[36] A. J. S. Hamilton, Linear Redshift Distortions: A Review, p. 185–275. Springer Netherlands, 1998. 10.1007/978-94-011-4960-$0_1$7.

[37] E. V. Linder, Cosmic growth history and expansion history, Physical Review D **72** (2005) .

[38] P. McDonald and D. J. Eisenstein, Dark energy and curvature from a future baryonic acoustic oscillation survey using the lyman-α forest, Phys. Rev. D **76** (2007) 063009.

[39] D. J. Eisenstein, H.-J. Seo and M. White, On the robustness of the acoustic scale in the low-redshift clustering of matter, The Astrophysical Journal **664** (2007) 660.

[40] C. for Astrophysics and Supercomputing, "Evolution of perturbations through different stages of the universe." http://astronomy.swin.edu.au/cms/imagedb/albums/scaled_cache/cmbFinal_sam_small-280x396.jpg.

[41] T. A. J. et al., "Desi: A new window on the universe." YouTube video, 2020.

[42] M. e. a. Blomqvist, The triply-ionized carbon forest from eBOSS: cosmological correlations with quasars in SDSS-IV DR14, JCAP **1805** (2018) 029 [1801.01852].

[43] D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton, R. C. Nichol et al., Detection of the baryon acoustic peak in the large-scale correlation function of sdss luminous red galaxies, The Astrophysical Journal **633** (2005) 560–574.

[44] S. Cole, W. J. Percival, J. A. Peacock, P. Norberg, C. M. Baugh, C. S. Frenk et al., The 2df galaxy redshift survey: power-spectrum analysis of the final data set and cosmological implications, Monthly Notices of the Royal Astronomical Society **362** (2005) 505 [https://academic.oup.com/mnras/article-pdf/362/2/505/6155670/362-2-505.pdf].

[45] J. E. Bautista et al., Measurement of baryon acoustic oscillation correlations at $z = 2.3$ with SDSS DR12 Lyα-Forests, Astron. Astrophys. **603** (2017) A12 [1702.00176].

[46] H. e. a. du Mas des Bourboux, Baryon acoustic oscillations from the complete SDSS-III Lyα-quasar cross-correlation function at $z = 2.4$, Astron. Astrophys. **608** (2017) A130 [1708.02225].