English summary

Many people often feel overwhelmed when they think about the Universe and the obvious questions that accompany this thought. How can we blame them? Astrophysics and cosmology are definitely as fascinating as complicated, wild, and "frightening". As scientists, the least we can do is make these concepts more accessible to everybody, to help fight the fear of the unknown. We can start this story by saying that the study of our Universe is very broad; so wide that there are probably as many aspects that could be studied as the number of stars living out there. The macro-areas we can identify, progressing from *our* surroundings to the edges of the Universe, include: Earth & near-Earth astrophysics, Solar System & exoplanets, stars & their evolution, interstellar & intergalactic medium, galaxies & their evolution, cosmology & large-scale structure, and early Universe & fundamental physics. All of these topics play a unique role in building, piece by piece, the big puzzle of our knowledge, and we look at each of them depending on the question we are trying to answer.

For the last four years, I have been wondering about the early Universe, in particular the *Epoch of Reionization* (EoR), when the first luminous sources started to emit ionizing radiation and, hence, the gas permeating the Universe transitioned from a neutral to an ionized state. It is a well-known fact that supermassive black holes (SMBHs) of a billion solar masses were already in place during this cosmic era, less than a billion years after the Big Bang. These objects are observed as *quasars* at redshifts $z \gtrsim 6$, accreting systems that shine as luminous beacons through the fog of the early intergalactic medium (IGM). Their existence poses a series of open questions: how such massive black holes could grow so rapidly, how long they remain active as luminous quasars, and what their light can reveal about the timing and nature of IGM reionization.

Thus, quasars are not only extraordinary laboratories for studying black hole physics, but also unique probes of the IGM. The ultraviolet radiation they emit carves out ionized regions in the surrounding gas, producing distinct observational signatures visible in their spectra. Among these are the so-called (*i*) *proximity zones*, enhanced ionization regions immediately blueward of the quasar Lyα emission line, and the extended absorption features known as (*ii*) *damping wings*, that may appear redward of the Lyα line if the IGM retains a significant neutral hydrogen fraction. Careful interpretation of these features allows us to place constraints on both the properties of quasars, such as their lifetimes and ionizing luminosities, and on the global ionization state of the early Universe.

The central aim of this thesis is to extract physical insights from high-quality spectra of quasars at $z \gtrsim 5.7$ by means of various types of analysis of their Lyα transmission and spectral features in general. To this end, we have first collected, reduced, and published data from a large sample of sources; then, we have used and expanded this catalog based on the scientific goal to achieve.

Chapter 1 provides the astrophysical background necessary to understand the content of this work. In particular, we provide some guidelines on what the EoR and active galactic nuclei (AGN) are, focusing on high-z quasars' features and related challenges, and concluding with the problem of the rapid SMBH growth. We introduce the main observational tools used to study both IGM and quasars, and describe the open questions that motivate this work.

Chapter 2 presents our spectroscopic dataset of 45 quasars spanning $6.50 < z < 7.65$, with luminosities in the range $-28.82 \leq M_{1450} \leq -24.13$. The spectra were collected with both echelle and long-slit instruments and uniformly processed with the open-source pipeline `PypeIt`, ensuring consistent calibration and reduction. Alongside a detailed description of the sample selection, reduction, and data quality (i.e., median signal-to-noise ratio in the J, H, and K bands), the chapter introduces composite rest-frame UV spectra constructed from this dataset, and compares the main one of them with other high- and low-redshift templates.

The resulting catalog and composites are publicly released and provide a reference framework for high-z quasar continua and emission-line properties, crucial for applications such as proximity zone studies. Moreover, the comparison with other templates reveals that the overall spectral shapes and emission-line strengths remain consistent across cosmic time, suggesting limited evolution in the broad UV properties of quasars.

Chapter 3 presents uniform measurements of proximity zone sizes ($R_{\rm p}$) for 59 quasars in the range $5.76 < z < 7.55$ (where 35 of them come from the sample described in Chapter 2, and the remaining from other two public catalogs), including 15 objects analyzed for the first time. By extending the sample to nearly 100 quasars using literature measurements, we explore how $R_{\rm p}$ varies with both luminosity and redshift. The luminosity dependence follows expectations from simple photo-ionization arguments, whereas the inferred redshift trend is steeper than previously found in the literature ($R_{\rm p} \propto 10^{-0.4M_{1450}/2.87} \cdot (1+z)^{-2.44}$).

A striking outcome is the significant scatter in $R_{\rm p}$ at fixed z and M_{1450}. We attribute this to the interplay between quasar lifetime variations, effects in line of sight density fluctuations and host environments, as well as differences in the neutral hydrogen fraction and the incidence of optically thick absorbers. We find that 13 quasars stand out with unusually small $R_{\rm p}$ values. In all but one case, the presence of proximate absorbers is unlikely, suggesting that extremely young sources ($t_{\rm Q} \lesssim 10^4$ yr) could be a natural explanation for such small proximity zones.

Chapter 4 introduces a forward-modeling framework to confront observed Lyα transmission probability distribution functions (PDFs) with predictions from simulated spectra that account for uncertainties in the quasar continuum and instrumental noise. This approach is applied to 21 quasars from the E-XQR-30 sample at $5.76 < z < 6.59$, and PDFs are computed in two rest-frame ranges: the Lyα forest ($1170 - 1190$ Å) and the proximity zone ($1190 - 1218$ Å). From IGM absorption models, we generate thousands of mock realizations for each object and compare them to the data using statistical tests.

The analysis reveals that real transmission PDFs within proximity zones are generally well reproduced by the models, indicating that current treatments of radiative transfer and continuum fitting capture the physics in quasar-dominated regions. In contrast, several redshift bins in the Lyα forest show significant mismatches with the fiducial homogeneous IGM assumptions. These discrepancies point toward additional physical ingredients, such as patchy reionization, residual heating, or ultraviolet background fluctuations, that are not captured by simple models. The results highlight the diagnostic power of transmission PDFs and underscore the need for more sophisticated simulations to interpret the IGM near the end of reionization.

Taken together, the works presented in this thesis offer new understandings into the interaction between quasars and the IGM during the EoR. By building a high-quality spectroscopic dataset and composite spectrum, systematically measuring quasar proximity zones, and applying forward-modeling techniques to Lyα transmission statistics, we have both expanded the available observational constraints and highlighted the limitations of simplified IGM models. These results emphasize the role of quasars as unique probes of cosmic history, while also underscoring the need for increasingly realistic simulations to capture the complexity of the early Universe. In this way, the thesis makes a small but significant contribution to the larger enigma of how the first luminous structures transformed our cosmos.