Studying pressure effect using Lyman-alpha forests of close quasar pairs

Marinichenko, Mariia

Studying pressure effect using Lyman-α forests of close quasar pairs

THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE
in
PHYSICS

Author :	Mariia Marinichenko
Student ID :	3747107
Supervisor :	Oleksii Boyarsky
Second corrector :	Matthieu Schaller

Leiden, The Netherlands, February 16, 2024

Studying pressure effect using Lyman-α forests of close quasar pairs

Mariia Marinichenko

Instituut-Lorentz, Leiden University
P.O. Box 9500, 2300 RA Leiden, The Netherlands

February 16, 2024

Abstract

The intergalactic medium (IGM) contains most of the baryonic matter of the Universe and serves as a suitable environment for probing the thermal history of the Universe. The crucial moment in IGM evolution is the Epoch of Reionization, corresponding to the transition from neutral to ionized IGM. However, due to the observational limitations, this period is still not well understood. In this thesis, we focus on constraining IGM thermal history by using Lyman-α forests data. This method is applicable in a wide range of temperatures, densities, and ionization fractions of cosmic gas at $z \sim 2-5$. Observations show that the longitudinal flux power spectrum of Lyman-α forest exhibits a cut-off at small scales. This phenomenon is caused by thermal Doppler broadening, peculiar velocities along the line of sight (LOS), Hydrogen pressure smoothing, and warm dark matter. The first two effects act only along LOS, while the last two affect all spatial directions. To separate the one-dimensional and three-dimensional effects, we used the method of close quasar pairs, which is based on studying the correlations between Lyman-α forests of close quasar pairs. We used the Kolmogorov-Smirnov test to analyze the differences between distributions of phase difference θ, which characterizes correlations between Lyman-α forests. The calculations were performed for various thermal histories, parameters characterizing IGM, LOS separations $r_\perp$, wavenumbers k, and accounting for different effects (Doppler broadening and peculiar velocities). Our results indicate that this method can distinguish various thermal

histories regardless of the IGM thermal state and one-dimensional effects. Moreover, at separations of the order of pressure broadening, there is a prominent feature caused by different influences of pressure smoothing at large and small scales. In addition, this simple and powerful approach has the potential to distinguish scenarios with warm dark matter.

Contents

Version of February 16, 2024 – Created March 5, 2024 - 09:37

Chapter 1

Introduction

The ΛCDM model is a well-established cosmological paradigm, supported by numerous experiments and observational probes. It postulates that the Universe is homogeneous at large scales, isotropic, spatially flat, and filled with dark energy (cosmological constant Λ), cold dark matter (CDM), and baryonic matter.

The intergalactic medium (IGM) contains $\sim 80\%$ of the baryonic matter [1–6] and provides important information about the evolution of the Universe. It is useful for probing small-scale structures, cosmic microwave background (CMB) anisotropies, and physics influencing galaxy formation [7]. In addition, IGM can assist in distinguishing between warm and cold dark matter [8].

The main ingredient of IGM, neutral Hydrogen, was created during the recombination epoch ($z \sim 1100$). Its interactions with photons maintained thermal equilibrium up to $z \sim 100$ when it started to cool rapidly. At $z \sim 20$, stellar formation ends the cosmological "dark ages" [9–11] and starts the epoch of Reionization (EoR) when the IGM becomes ionized and heated up to the temperatures of $T \sim 10^4$ K. The history of this epoch is a crucial element in understanding the evolution of IGM. It contains information about the first ionization sources and their role in the formation of complex cosmic structures. The EoR is followed by processes of adiabatic cooling and inverse Compton scattering that left an imprint on the Hydrogen structures, providing insights into the poorly constrained reionization history [12–14].

The most efficient way to probe IGM is via absorption lines. There are two types of lines suitable for this purpose [15]: the 21-cm line and the Lyman-α forest. The former arises from the transition between two states with different spins of the ground level of neutral Hydrogen [16]. This is

Version of February 16, 2024 – Created March 5, 2024 - 09:37

a prospective tool, that attracted a lot of attention during the last decades [17]. However, its full potential is currently constrained by the ongoing development of observational technology. In contrast, the Lyman-α forest is an established and powerful method for studying IGM [18]. It is a collection of absorption lines in the spectra of distant bright objects originating from the transition of the electron from the ground to the first excited state of the neutral Hydrogen. The Lyman-α forest can probe a large range of redshifts and thermal parameters characterizing Hydrogen, including temperature, density, and ionization fraction, providing robust statistical analysis due to the abundance of absorbers [18]. Additionally, because of the clear physics of absorbers, the Lyman-α forest is an optimal choice for studying the pressure smoothing scale in the IGM.

Nevertheless, even the most sophisticated methods require a new generation of instruments, such as shown in Figure 1.1 the James Webb Space Telescope (JWST) and the Square Kilometre Array (SKA). The former is one of the most powerful tools for probing the reionization history [19]. Equipped with a near-infrared camera and spectrograph, JWST conducts ultra-deep imaging and spectroscopic surveys of first galaxies and quasars, studying the evolution of their number, luminosity, and metallicity. Those measurements will answer the question about the nature of reionization sources. In addition, the mission also performs the diagnostics of the Lyman-α forest to determine the timing of reionization. Another ambitious project is SKA [20, 21], the largest radio telescope in the world. Thousands of dipole antennas will occupy an area of one square kilometer in Australia and South Africa, providing a baseline of up to 150 km. The telescope will operate in a frequency range from 100 MHz to 25 GHz, and the SKA-low part will study the redshifted 21-cm neutral Hydrogen line with high sensitivity and resolution in frequency and angular domains. The survey will make the tomography of the 21-cm signal to map the shape of the ionized regions, facilitating in finding and studying the first sources of reionization. Besides that, SKA will measure the power spectrum of the 21-cm line to examine cosmic gas density, ionization, and temperature, which will help to constrain the timing of reionization. Both surveys will explore the unobserved before high redshift regions (up to $z \sim 30$) to illuminate the intricacies of reionization history.

Despite the noticeable attention from the scientific community, the EoR remains one of the most mysterious topics in modern cosmology. The reason is in limited observational probes that are affected by various uncertainties and systematics. For example, CMB polarization measurements provide only the integrated optical depth to EoR, while 21-cm line measurements are contaminated by foregrounds. However, the combination

(a) The James Webb Space Telescope. Credit: NASA.

(b) The Square Kilometre Array. Credit: SKAO

Figure 1.1: *New generation surveys for constraining the epoch of reionization.*

of different observations enables constraining of the reionization with high precision [22, 23]. Moreover, it is possible to extract useful information even with one type of observation using more sophisticated methods, such as close quasar pairs, which is the focus of this thesis.

1.1 Epoch of Reionization

Before the recombination, the Universe was a hot, dense plasma of Hydrogen and Helium ions accompanied by electrons. The adiabatic expansion effectively cooled the plasma leading to the formation of the first Hydrogen and Helium atoms and initiating the cosmic "dark ages". At $z \sim 1000$ the Universe became sufficiently cold, initiating the recombination of electrons and protons into Hydrogen atoms, making the Universe neutral and transparent to photons. The last scattering of these photons left an imprint on the observed anisotropies in the CMB.

After recombination, the gravitational collapse led to the formation of dense structures. The first stars emerged at redshifts $z \sim 30$ [24]. Their radiation created bubbles of ionized cosmic gas, indicating the end of the cosmic dark ages and the beginning of the Epoch of Reionization (Figure 1.2).

Sources of Reionization

The first sources of light, initiating the reionization, are the Population III (POPIII) stars. They were to a great extent different from the modern ones, namely short-lived, massive ($M \gtrsim 100 M_\odot$), and made of Hydrogen and Helium [26]. Nuclear reactions in their stellar cores produced metals,

Version of February 16, 2024 – Created March 5, 2024 - 09:37

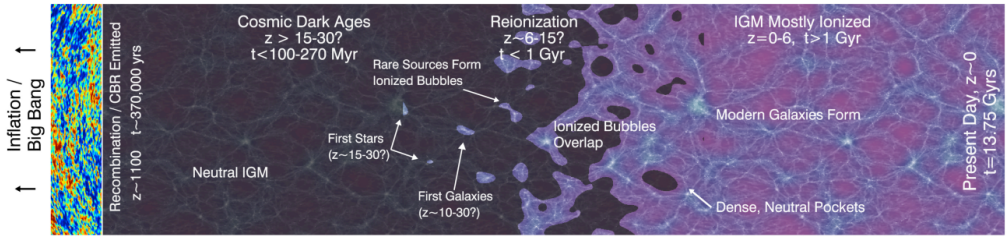


Figure 1.2: *Schematic representation of the ionization history of the Universe. Credit: Robertson et al. (2010)[25].*

providing the material for a new star type (Population II). Those stars with metallicity $Z \sim 10^{-4}$ comprised the first galaxies and had a significant impact on the reionization history.

At the end of their life, POPIII stars collapsed into intermediate-mass black holes (IMBHs) with mass M in the range $10^2 M_\odot \gtrsim M \gtrsim 10^5 M_\odot$. Accreting the surrounding gas, they became miniquasars, a type of active galactic nuclei (AGNs) [27]. Radiating in a wide range of wavelengths, miniquasars are important reionization sources. They are among the most luminous objects in the Universe, which differentiates them from other AGNs. So far, almost a million quasars have been detected by various sky surveys. The full list can be found in [28].

Despite the insufficient knowledge about primary radiation sources, it is evident that the reionization was driven by both stars and miniquasars. The main argument in favor of miniquasars is their ability to heat Hydrogen to the observed temperatures of $\sim 10^4$ K. While first generations of stars shine in the ultraviolet (UV) part of the spectra, quasars can emit X-ray photons with energies enough not only for Hydrogen ionization but also for its heating [29]. However, stellar reionization better agrees with the observed amount of soft X-ray and infrared (IR) radiation [30–32].

In addition, more exotic reionization sources were proposed, such as decaying and annihilating dark matter, evaporating black holes, and decaying cosmic strings [33].

Redshift of Reionization

Observations show that at $z \sim 6$, all distant quasars have a trough in their spectra blueward Lyman-α line. An example spectrum is represented in Figure 1.3. This phenomenon is caused by the significant amount of neutral Hydrogen in the IGM, which absorbed quasar radiation and became ionized. At $z \lesssim 6$, this effect can be expressed in terms of optical depth

τ_{GP} [34]:

$$\tau_{GP}(z) = 4.5 \cdot 10^5 \left(\frac{\Omega_m h^2}{0.13}\right)^{-1/2} \left(\frac{\Omega_b h^2}{0.02}\right) \left(\frac{1+z}{7}\right)^{3/2} x_{HI}(1+\delta), \quad (1.1)$$

where Ω_m and Ω_b are the total matter and baryonic density parameters, and x_{HI} is the neutral Hydrogen fraction. This equation shows that $x_{HI} \gtrsim 10^{-4}$ is sufficient for a considerable absorption.

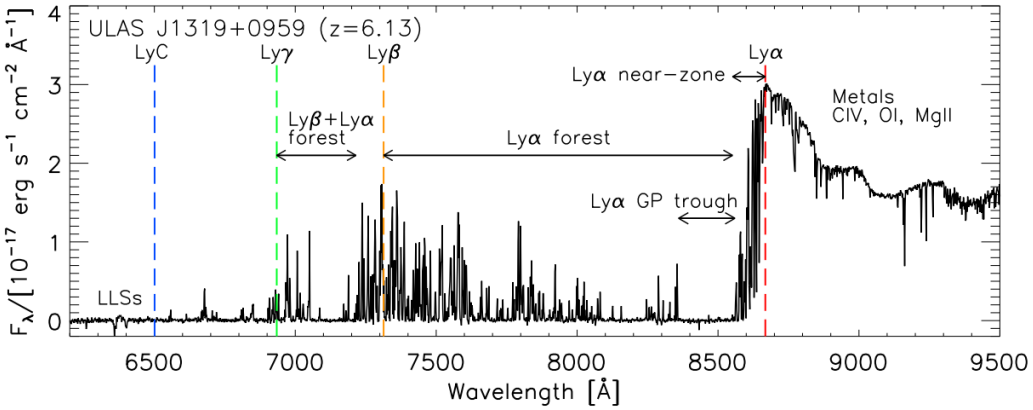


Figure 1.3: *Spectrum of quasar ULAS J1319+0959, located at $z = 6.13$, with prominent Gunn-Peterson trough. The spectrum obtained by X-Shooter spectrograph on the Very Large Telescope. Credit: Becker et al. (2015) [35].*

The "Gunn-Peterson trough" [36] indicates the end of Hydrogen reionization at $z \sim 6$ [37, 38]. The recent Plank data [39] constrains this time to $7.8 \leq z \leq 8.8$ with maximum redshift range $\Delta z = 2.8$ depending on the reionization model. In addition, [39] demonstrates that the Universe was less than 10% ionized at $z \sim 10$, discarding models of late reionization. However, the range of possible reionization histories is broad even under the existing constraints, which is illustrated by Figure 1.4.

After the Hydrogen ionization, the IGM underwent a period of adiabatic cooling ended by AGN activity at $z \sim 3.5$ [15]. Quasars emit energetic photons capable of effectively ionizing Helium from HeII to HeIII (54.4 eV) and increasing IGM temperature to $\sim 10^4$ K. The epoch of Helium reionization ended at $z \sim 2.7$, which is imprinted in troughs in quasar spectra [41]. Completely ionized IGM then cooled down allowing photons to travel freely across the Universe.

Version of February 16, 2024 – Created March 5, 2024 - 09:37

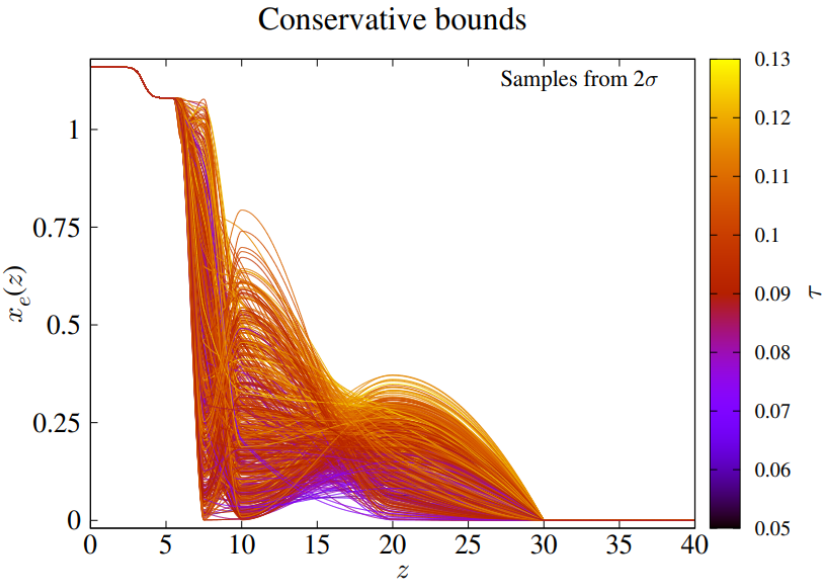


Figure 1.4: *Free electron fraction redshift evolution $x_e(z)$ corresponding to thermal histories within 95.4% confidence interval of the conservative Planck CMB data. Colors indicate the optical depth to reionization τ. Credit: Hazra et al.(2017) [40].*

1.2 Hydrogen pressure

At large scales, Hydrogen distribution follows dark matter distribution. However, this is not the case at small scales, where baryon densities are high and pressure plays an important role. It smooths Hydrogen distribution at some characteristic scale resulting in deviations from the dark matter one.

To characterize the Hydrogen pressure, it is essential to understand the dynamics of linear perturbations in the dark matter-baryon fluid. Within the Newtonian framework, a self-gravitating non-relativistic fluid in a non-expending environment is described by three fundamental equations, namely energy conservation, Euler equation, and Poisson equation:

$$\begin{cases} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \\ \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} + \frac{\nabla p}{\rho} + \nabla \phi = 0, \\ \Delta \Phi = 4\pi G \rho, \end{cases} \tag{1.2}$$

where $\mathbf{v}(\mathbf{x}, t)$ is the fluid velocity, $P(\mathbf{x}, t)$ is the pressure, $\rho(\mathbf{x}, t)$ is the density, and $\Phi(\mathbf{x}, t)$ is the gravitational potential.

To complete the system, the equation of state $\frac{\partial p}{\partial \rho} = c_s^2$ is required. Here, $c_s^2 = \gamma \frac{T}{\mu}$ is the speed of sound in the fluid, γ is the adiabatic index, and μ is a mean molecular mass of the gas.

The system (1.2) can be rewritten in terms of small perturbations $\delta \ll 1$ around a homogeneous background:

$$\rho(\mathbf{x}, t) = \rho_0(t) + \delta\rho(\mathbf{x}, t) = \rho_0(t)\big(1 + \delta(\mathbf{x}, t)\big), \tag{1.3}$$
$$\mathbf{v}(\mathbf{x}, t) = \mathbf{v}_0(\mathbf{x}, t) + \delta\mathbf{v}(\mathbf{x}, t), \tag{1.4}$$
$$P(\mathbf{x}, t) = P_0(t) + \delta P(\mathbf{x}, t), \tag{1.5}$$
$$\Phi(\mathbf{x}, t) = \Phi_0(\mathbf{x}, t) + \delta\Phi(\mathbf{x}, t). \tag{1.6}$$

In an expanding Universe, the background velocity is a Hubble velocity $\mathbf{v}_0(\mathbf{x}, t) = \mathbf{x} \cdot H(t)$ and unperturbed potential $\Phi_0(\mathbf{x}, t) = \frac{2\pi G}{3}\rho_0 \mathbf{x}^2$, where $a(t)$ is the scale factor, and $H = \frac{\dot{a}}{a}$ is the Hubble parameter. Substituting these expressions into the system (1.2) and neglecting terms of second or higher order, we obtain the linearized equations:

$$\begin{cases} \frac{\partial \delta}{\partial t} + H\mathbf{x} \cdot \nabla\delta + \nabla \cdot \mathbf{v} = 0, \\ \frac{\partial}{\partial t}\delta\mathbf{v} + H\delta\mathbf{v} + H(\mathbf{x} \cdot \nabla)\delta\mathbf{v} + c_s^2 \nabla\delta + \nabla\delta\Phi = 0, \\ \Delta\delta\Phi = 4\pi G\delta\rho. \end{cases} \tag{1.7}$$

To eliminate terms accounting for uniform expansion, we can write the system (1.7) in terms of comoving coordinates:

$$\mathbf{r}(t) = \frac{\mathbf{x}(t)}{a(t)}, \qquad \mathbf{u}(\mathbf{r}, t) = \frac{\mathbf{v}(a\mathbf{r}, t)}{a(t)}. \tag{1.8}$$

Furthermore, in the linear regime, it is convenient to make a Fourier transform:

$$\delta(\mathbf{r}, t) = \int d^3k \, e^{i\mathbf{k}\cdot\mathbf{r}} \, \delta(\mathbf{k}, t), \tag{1.9}$$

where $\mathbf{k}$ is the comoving wavevector. The resulting oscillations corresponding to different wavevectors are independent and can be analyzed separately. After eliminating $\mathbf{u}$ and Φ from the equations via algebraic transformations, the equation of the evolution of linear density perturbation is

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} = \delta\left(4\pi G\rho_0 - \frac{c_s^2 k^2}{a^2}\right), \tag{1.10}$$

where $\dot{\delta} = \frac{d\delta}{dt}$ is the derivative with respect to proper time t. The space and time dependence can be factorized, providing the exponent solution

Version of February 16, 2024 – Created March 5, 2024 - 09:37

in the static background ($\dot{a} = 0$):

$$\delta(t) = \delta_0 \exp\left[\pm\left(4\pi G\rho_0 - c_s^2 k^2\right)t\right] = \delta_0 \exp\left[\pm 4\pi G\rho_0\left(1 - \frac{\lambda_J^2}{\lambda^2}\right)t\right], \tag{1.11}$$

where $\lambda_J = \sqrt{\frac{\pi}{G\rho_0}}$ is the Jeans length, the scale at which pressure and gravitational forces are balanced. Modes with length $\lambda > \lambda_J$ correspond to gravity dominance over the pressure, leading to collapsed objects. Conversely, modes with $\lambda < \lambda_J$ are sound waves.

In the case of an expanding multicomponent Universe, each component λ is characterized by an equation of state:

$$\ddot{\delta}_\lambda + 2\frac{\dot{a}}{a}\dot{\delta}_\lambda = \sum_\alpha 4\pi G\rho_{0,\alpha}\delta_\alpha - \frac{c_{s,\lambda}^2 k^2}{a^2}\delta_\lambda, \tag{1.12}$$

where the summation α goes over all components.

During the matter-dominated epoch, only contributions from dark matter and baryons are important. Other components can be neglected. The evolution of perturbations in this fluid is described by the system [12]:

$$\begin{cases} \ddot{\delta}_{DM} + 2\frac{\dot{a}}{a}\dot{\delta}_{DM} = 4\pi G\rho_0\left(f_{DM}\delta_{DM} + f_b\delta_b\right), \\ \ddot{\delta}_b + 2\frac{\dot{a}}{a}\dot{\delta}_b = 4\pi G\rho_0\left(f_{DM}\delta_{DM} + f_b\delta_b\right) - \frac{c_s^2 k^2}{a^2}\delta_b, \end{cases} \tag{1.13}$$

where f_{DM} and f_b are the dark matter and baryon mass fractions of the average mass density of the Universe ρ_0, and δ_{DM} and δ_b are the Fourier components of density fluctuations of dark matter and baryons. System (1.13) shows that the differences between Hydrogen and dark matter fluctuations are caused by pressure smoothing.

Assuming a constant in time Jeans scale ($T \propto \frac{1}{a}$ at all times) and negligible baryonic contribution ($f_b = 0$), the solution of the system (1.13) satisfies

$$\delta_b(k,t) = \frac{\delta_{DM}(k,t)}{1 + k^2/k_J^2}. \tag{1.14}$$

This expression shows that on large scales ($k \to 0$), baryon density perturbations follow those of dark matter, while on small scales ($k \to \infty$), they are filtered by the factor k_J^2/k^2.

Filtering scale

In reality, the temperature T does not evolve as $\propto \frac{1}{a}$ at all times, meaning that the pressure smoothing depends on the entire thermal history of the

Version of February 16, 2024 – Created March 5, 2024 - 09:37

Universe. Counteracting gravitational collapse, pressure defines the subsequent evolution of the halo. The characteristic scale is called **the filtering scale** ${k_F}^{-1}$ and defined as [12]:

$$\frac{1}{k_F^2(t)} = \frac{1}{D_+(t)} \int_0^t dt'\ a^2(t') \frac{\ddot{D}_+(t') + 2H(t')\dot{D}_+(t')}{k_J^2(t)} \int_{t'}^t \frac{dt''}{a^2(t'')}, \tag{1.15}$$

where $D_+(t)$ is the growth factor of the dark matter density fluctuations [42]. It is important to notice, that this formula is applicable only for small scales($k \lesssim h\,\mathrm{cMpc}^{-1}$).

At $z \gtrsim 3$, the Universe is matter-dominated ($\Omega_m \approx 1$), implying that $D_+(t) = a(t)$ and $a(t) = t^{2/3}$. Then the Eq. (1.15) can be simplified to

$$\frac{1}{k_F^2(a)} = \frac{1}{a} \int_0^a \frac{da'}{k_J^2(a')} \left[1 - \left(\frac{a'}{a}\right)^{\frac{1}{2}}\right]. \tag{1.16}$$

After using the mean value theorem,

$$k_F^2(t) = k_J^2(t^*), \tag{1.17}$$

where $0 < t^* < t$. As a result, the filtering scale at a time t coincides with the Jeans scale at some earlier time t^*.

After the recombination photons continued to interact with cosmic gas by inverse Compton scattering, resulting in the equality of the filtering scale and Jeans scales. At $z \gtrsim 100$, this heating mechanism lost efficiency, and the gas temperature dropped rapidly as $T \propto a^{-2}$, causing the decrease of Jeans scale. Consequently, the filtering scale exceeded the Jeans scale up to the reionization when it experienced significant changes due to the thermal processes of that epoch. At $z < z_{reion}$, the temperature decreased as $T \propto a^{-\alpha}$ with $0 < \alpha < 1$, meaning that the filtering scale became smaller than the Jeans scale. This phenomenon is explained by the delayed response of baryonic matter to thermal processes, showing that the gas requires time to adjust to pressure changes.

In a matter-dominated Universe, this reaction time can be estimated as:

$$t_{react} \sim \frac{\lambda_F}{c_s} = \sqrt{\frac{\pi}{G\rho_0}} \sim \frac{10^{11}}{(1+z)^{3/2}}\ \mathrm{yr}. \tag{1.18}$$

This time is larger than the age of the Universe, which explains why the filtering scale grows so slowly.

It has to be noted that all the derivations above are done with the assumption of linearity ($\delta \lesssim 1$). However, the Lyman-α forest probes

$z \sim 2-5$, where scales of the order of the pressure smoothing are highly nonlinear. The processes of galaxy formation influence them. There is no analytical expression for the k_F in a nonlinear case. In addition, in a quasi-linear regime, the pressure smoothing depends on the local density, which is not captured by the Eq. (1.15) [43].

During the post-reionization epoch, the filtering scale ranges from approximately 50 ckpc to 70 ckpc [12, 14]. However, the nonlinear pressure smoothing scale is larger, about 100 ckpc at a redshift $z \sim 3$, and can grow up to 200 ckpc as the IGM temperature approaches 10^5 K [14].

The power law approximation (1.14) is valid only for the small k, the exact general expression is

$$\frac{\delta_b(k,t)}{\delta_{DM}(k,t)} = e^{-\frac{k^2}{k_F^2}}. \tag{1.19}$$

Only for the fixed exponent of T power law evolution, there is an k^{-2} asymptote in high k [12].

1.3 Desity of neutral Hydrogen

During and after the EoR, Hydrogen ionization is induced by two processes: photoionization and collisional ionization. The rate of collisional ionization R_{col} is:

$$R_{col} = \langle \sigma_{col} v_{rel} \rangle_T n_{HI} n_H = \alpha_{col} n_{HI} n_H, \tag{1.20}$$

where σ_{col} is cross-section of collisional ionization, v_{rel} is the relative velocity of particles, n_{HI} and n_H are neutral and total Hydrogen densities, α_{col} is the velocity-averaged collisional ionization cross section. At temperatures below Hydrogen ionization $T_{ion} = 1.6 \times 10^5$ K, his process can be neglected.

The rate of photoionization is:

$$R_\gamma = \int_{E_\gamma > 13.6\,\mathrm{eV}} \sigma(E_\gamma)\, c\, n_{HI} \frac{dn_\gamma}{dE_\gamma} dE_\gamma = \Gamma_\gamma n_{HI}, \tag{1.21}$$

where $\sigma(E_\gamma)$ is the cross-section of ionization, $\frac{dn_\gamma}{dE_\gamma}$ is the energy density of photons, and Γ_γ is the rate of ionization, which depends on the distribution of ionizing photons (star activity).

The process inverse to ionization is recombination, which has the rate:

$$R_R = \langle \sigma_R v_{rel} \rangle_T n_e n_p \approx \alpha_R n_H n_e, \tag{1.22}$$

Version of February 16, 2024 – Created March 5, 2024 - 09:37

where α_R is velocity-averaged recombination cross-section. The α_R depends on temperature as $\alpha_R \propto T^{-0.7}$ [14, 44].

Close to the reionization, the evolution of neutral Hydrogen fraction can be described by

$$\frac{dn_{HI}}{dt} = R_R - R_\gamma. \tag{1.23}$$

The solution is:

$$n_{HI}(t) = \frac{\alpha_R n_H n_e}{\Gamma_\gamma} + C\, n_H e^{-\Gamma_\gamma t}, \tag{1.24}$$

where the constant C can be inferred from the initial conditions.

In the ionization equilibrium,

$$n_{HI}(t) = \frac{\alpha_R n_H n_e}{\Gamma_\gamma} = \frac{\alpha_R n_H^2}{\Gamma_\gamma} \frac{1 - Y_p/2}{1 - Y_p}, \tag{1.25}$$

where Y_p is the Helium mass fraction. This formula is valid only for highly ionized minimum ($n_{HI} \ll n_H$).

1.4 Thermal state of IGM after the reionization

To understand the EoR, we need to look at the temperature-density relation for neutral Hydrogen. The evolution of a photoionized IGM is influenced by various factors including changes in density over time, properties of ionizing radiation, and cosmological parameters. Due to the lack of knowledge about the actual thermal history of the Universe, the common approach is to examine the effects of a range of reionization histories.

A simplified but comprehensive approach is to integrate the Zel'dovich approximation with equations of thermodynamic and chemical evolution. It results in the temperature-density relation.

The evolution of gas temperature $T(\mathbf{r}, t)$ is defined as [44]:

$$\frac{dT}{dt} = -2HT + \frac{2T}{3(1+\delta)}\frac{d\delta}{dt} - \frac{T}{\sum_i \tilde{X}_i}\frac{d\sum_i \tilde{X}_i}{dt} + \frac{2}{3k_B n_b}\frac{dQ}{dt}, \tag{1.26}$$

where n_i is the proper number density of the specie i, $\frac{dQ}{dt}$ is the heat produced in unit volume by the surrounding gas particles, m_p is the proton mass, and $\tilde{X}_i = \frac{n_i m_p}{(1+\delta)\rho_{0,b}}$. The latter can be calculated as:

$$\frac{dQ}{dt} = \sum_i n_i \int_{\nu_i}^{\infty} 4\pi J(\nu)\sigma_i(h\nu - h\nu_i)\frac{d\nu}{h\nu} = n_i \epsilon_i, \tag{1.27}$$

Version of February 16, 2024 – Created March 5, 2024 - 09:37

where J_ν is the specific intensity of the ionizing radiation at frequency ν, ν_i is the ionization frequency the specie i, σ_i is the ionization cross-section, and ϵ_i is the heat gain per particle of specie i.

The Eq.(1.26) accounts for adiabatic cooling and heating (first two terms), change of energy per particle due to changes in the number of particles (third term), and external heating.

Quantities $\tilde{X}_i$ can be calculated from the abundance equations accounting for recombination and ionization:

$$\frac{d\tilde{X}_i}{dt} = \sum_{j,k} \tilde{X}_j \tilde{X}_k R_{jk} \left[\frac{\rho_{0,b}(1+\delta)}{m_p}\right] - \tilde{X}_i \Gamma_i, \tag{1.28}$$

where R_{jk} is the recombination rate of the species j and k resulting in i, and Γ_i is the photoionization rate of the specie i:

$$\Gamma_i = \int_{\nu_i}^{\infty} 4\pi J_\nu \sigma_i \frac{d\nu}{h\nu}. \tag{1.29}$$

The dynamic of the fluid element is completely described by Eq. (1.26) and Eq. (1.28) combined with Zel'dovich approximation:

$$1 + \delta = (1 - D_+(t)\lambda)^{-1}, \tag{1.30}$$

where λ is an empirical constant.

In a system where ionization equilibrium is maintained (photoionization and recombination rates are equal), and where the ionization radiation field is uniform and constant in time and space, **the temperature-density relation** has a power-law form [44]:

$$T = T_0(1+\delta)^{\gamma-1}, \tag{1.31}$$

where $T_0 = T_0(z)$ is the baryon temperature at the average mass density, and $(1+\delta)^{\gamma(z)-1}$ are the fluctuations around the mean.

The Eq.(1.31) is of great significance for studying the thermal evolution of IGM. It has high precision and works not only for instant reionization models but also for models with preheating. In addition, this relation is independent of the ionization radiation spectra (as long as there is enough flux to keep the Universe highly ionized).

The temperature-density relation has two parameters (T_0 and γ) that depend on redshift.

The slope $\gamma(z)$ increases with redshift and earlier beginning of reionization. It can take values from $\gamma = 1$ (right after the reionization) to

$\gamma = 1.62$ (the asymptote for the early reionization). Importantly, this slope remains relatively insensitive to the precise values of reionization temperature and cosmological parameters.

The temperature T_0 decreases as reionization happens earlier. It is influenced by the reionization radiation spectrum and is expected to be around 2×10^4 K immediately after reionization [44]. When almost all Hydrogen is ionized, the IGM temperature is the result of two processes: photoionization heating and changes in the number of species. In this case, there is an analytical solution [44]:

$$T_{reion} = \frac{\epsilon_{HI}}{3k_B\Gamma_{HI}}, \tag{1.32}$$

which is valid in the limit of large t.

The mean temperature T_0 has an uncertainty about 50% due to the uncertainties in the reionization history. It causes 30% uncertainty in Hydrogen column density, affecting the interpretation of observations.

Chapter 2

Lyman-α forest and close quasar pairs

After reionization, the IGM is transparent to background radiation. However, a small Hydrogen fraction remains neutral and creates absorption lines in spectra of high-redshift quasars, which radiate in the entire electromagnetic spectrum, as shown in Figure 2.1. The most prominent feature is the Lyman-α line, which has a rest-frame wavelength of $\lambda_0 = 1215.67$ Å.

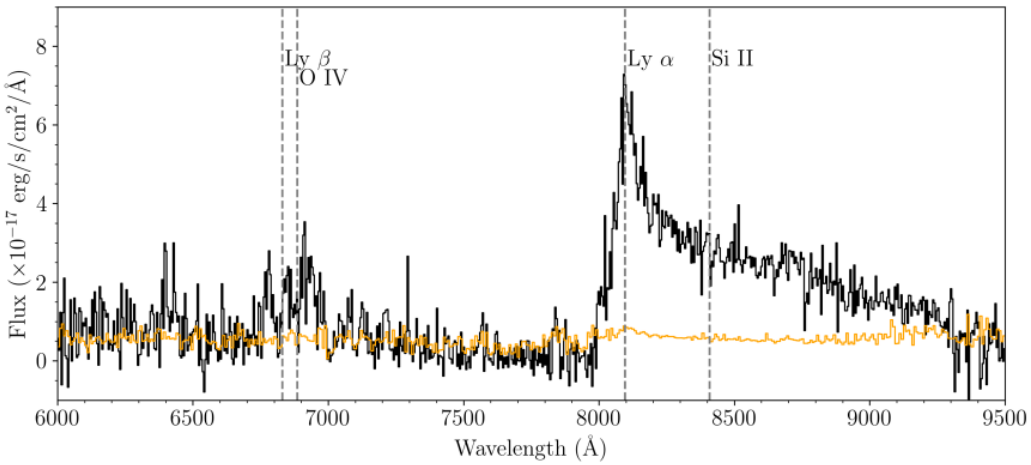


Figure 2.1: *1D spectrum of P144+50 with prominent emission lines measured by the Isaac Newton Telescope. The quasar is located at redshift $z = 5.66$. Credit: J. D. Wagenveld et al. [45].*

The collection of those absorption lines in quasar spectra corresponding to the Lyman-α transition within extended structures of column den-

sity $n_{HI} \leq 10^{17}\text{cm}^{-2}$ is called **the Lyman-α forest**. Distribution of neutral Hydrogen and background radiation define properties of these absorption lines [14].

The Lyman-α forest serves as a powerful tool for probing the IGM physics at redshifts $z \sim 2-5$, giving insights into the thermal history of the Universe. This redshift range is chosen due to the observability of redshifted Lyman-α lines using ground-based optical and IR telescopes. Additionally, at this range, there are numerous bright quasars with spectra of high signal-to-noise ratio.

The largest surveys targeting quasars are the extended Baryon Oscillation Spectroscopic Survey (eBOSS, [46]) of the Sloan Digital Sky Survey (SDSS, [47]) and the Dark Energy Spectroscopic Instrument (DESI, [48]). In total, almost a million quasars are known and collected in the Million Quasars Catalogue [28] by E. W. Flesch. The most distant quasars are located at $z \sim 10$, while the majority are concentrated at $2 < z < 3$, as illustrated in Figure 2.2.

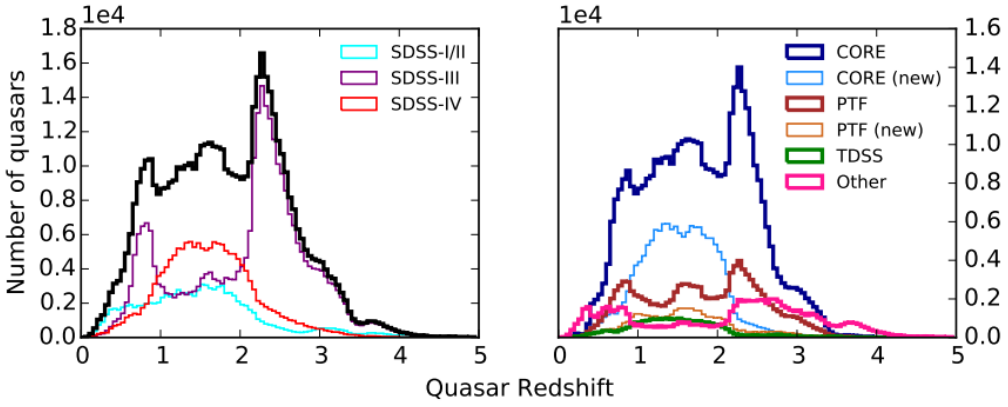


Figure 2.2: *Redshift distribution of SDSS quasars in the range* $0 < z < 5$. *Credit: SDSS collaboration [47].*

2.1 Optical Depth

The Lyman-α absorption is described by optical depth, which in the simplest scenario of a static environment is given by:

$$\tau_0(\lambda) = \int_0^r \sigma\left(\frac{c}{\lambda}\right) n_{HI}(r')dr', \tag{2.1}$$

Version of February 16, 2024 – Created March 5, 2024 - 09:37

where λ is the wavelength of the absorption, r is the distance traveled through the absorber, $\sigma_{HI}(\nu)$ is the frequency $\nu = c/\lambda$ dependent Lyman-α cross-section, $n_{HI}(r)$ is the density of neutral Hydrogen at the distance r along LOS, and $c = 3 \cdot 10^8$ m/s is the speed of light.

The Lyman-α cross-section is defined as:

$$\sigma_{HI}(\nu) = \frac{\pi e^2}{m_e c} f \phi(\nu - \nu_\alpha), \tag{2.2}$$

where e is the charge of an electron, m_e is electron mass, $f = 0.416$ is the strength of an oscillator. Function $\phi(\nu - \nu_\alpha)$ determines the shape of the cross-section, which can be delta-function or Lorentz profile.

However, the formula above is valid only for a static Universe, which is not the real case. To account for the expansion, one must rewrite the length element $dr = -\frac{cdz}{(1+z)H(z)}$. The Lyman-$\alpha$ absorption optical depth in this case is expressed as:

$$\tau(\lambda) = \int_0^{z_s} \sigma_{HI}\left(\frac{(1+z)c}{\lambda}\right) n_{HI}(z) \frac{cdz}{(1+z)H(z)}, \tag{2.3}$$

where z_s is the position of the source.

However, Eq. (2.3) does not include temperature Doppler broadening and peculiar velocities, requiring modifications to account for these factors.

Gas particles with temperature T follow Maxwell's distribution:

$$f(v) = \sqrt{\frac{m}{2\pi k_B T}} e^{-\frac{mv^2}{2kT}} = \frac{1}{b\sqrt{\pi}} e^{-\frac{v^2}{b^2}}, \tag{2.4}$$

where v is the one-dimensional particle velocity, m mass of gas particles, and $b = \sqrt{\frac{2k_B T}{m}}$ is the thermal dispersion.

Then the Lyman-α Doppler broadened cross-section is defined as an integral of $\sigma_{HI}(\nu(v))$ with the weight $f(v)$:

$$\sigma_T(\frac{c}{\lambda}) = \int_{-\infty}^{\infty} \sigma_{HI}\left(\frac{c}{\lambda}\sqrt{\frac{c-v}{c+v}}\right) \frac{1}{b\sqrt{\pi}} e^{-\frac{(v-v_\parallel)^2}{b^2}} dv, \tag{2.5}$$

where we also included shift due to the peculiar velocities $v_\parallel$.

The final expression for optical depth in the expanding Universe, which accounts for thermal Doppler broadening and peculiar velocities, is given by [49, 50]:

$$\tau(\lambda) = \int_0^z \frac{n_{HI}(z) c \; dz}{(1+z)H(z)} \int_{-\infty}^{\infty} \sigma_{HI}\left(\frac{c}{\lambda}\sqrt{\frac{c-v}{c+v}}\right) \frac{1}{b(z)\sqrt{\pi}} e^{-\frac{(v-v_\parallel(z))^2}{b^2(z)}} dv. \tag{2.6}$$

Version of February 16, 2024 – Created March 5, 2024 - 09:37

This result is known as the '**fluctuating Gunn-Peterson approximation**' (FGPA) [51]. Essentially, the resulting absorption resembles the Gunn-Peterson trough, fluctuating along its path containing regions of varying densities and having a power-law dependence on density $\tau \propto \rho^{\alpha}$. The exponent α can be calculated from the right-hand side of Eq. (2.6):

$$\tau \propto n_{HI} \propto n_H^2 \alpha_R \propto \rho^2 T^{-0.7} \propto \rho^{2-0.7(\gamma-1)} T_0^{-0.7(\gamma-1)}. \tag{2.7}$$

Expressions (2.6) and (2.7) illustrate that the characteristics of the absorption line (width and depth) are affected by the size and density of the Hydrogen cloud, peculiar velocity, gas temperature, and velocity gradient.

2.2 Flux Power Spectrum

The transmission flux is the fraction of light that comes from the source to the observer

$$F(v) = e^{-\tau(v)}, \tag{2.8}$$

where $\tau(v)$ is the optical depth depending on Hubble velocity

$$v = c \ln\left(\frac{\lambda}{\lambda_0(1+z)}\right) = \frac{H(z)}{1+z} y \tag{2.9}$$

at the comoving position y along the LOS.

Several statistics based on the longitudinal (1D) spectra are used to quantitatively characterize Lyman-α forest [14]. Namely, the flux power spectrum (FPS), probability distribution function (PDF) of wavelet amplitudes, PDF of transmitted flux, etc.

However, the most widely used quantity in the Lyman-α forest method is FPS $P_F(k)$, a Fourier transform of flux auto-correlation function:

$$P_F(k) = V\langle|\tilde{\delta}_F(k)|^2\rangle, \qquad \tilde{\delta}_F(k) = \frac{1}{V}\int_0^V dv\, e^{-ikv} \underbrace{\frac{F(v) - \langle F\rangle}{\langle F\rangle}}_{\delta_F(v)}, \tag{2.10}$$

where $\langle\cdot\rangle$ is the ensemble average, and $k = \frac{2\pi}{v}$.

Observations show, that FPS has a cut-off on small scales $\lambda_c \approx 30\text{km s}^{-1}$ [52–54], illustrated at Figure 2.3.

Four effects are influencing the cut-off [8]:

- thermal Doppler broadening;

Version of February 16, 2024 – Created March 5, 2024 - 09:37

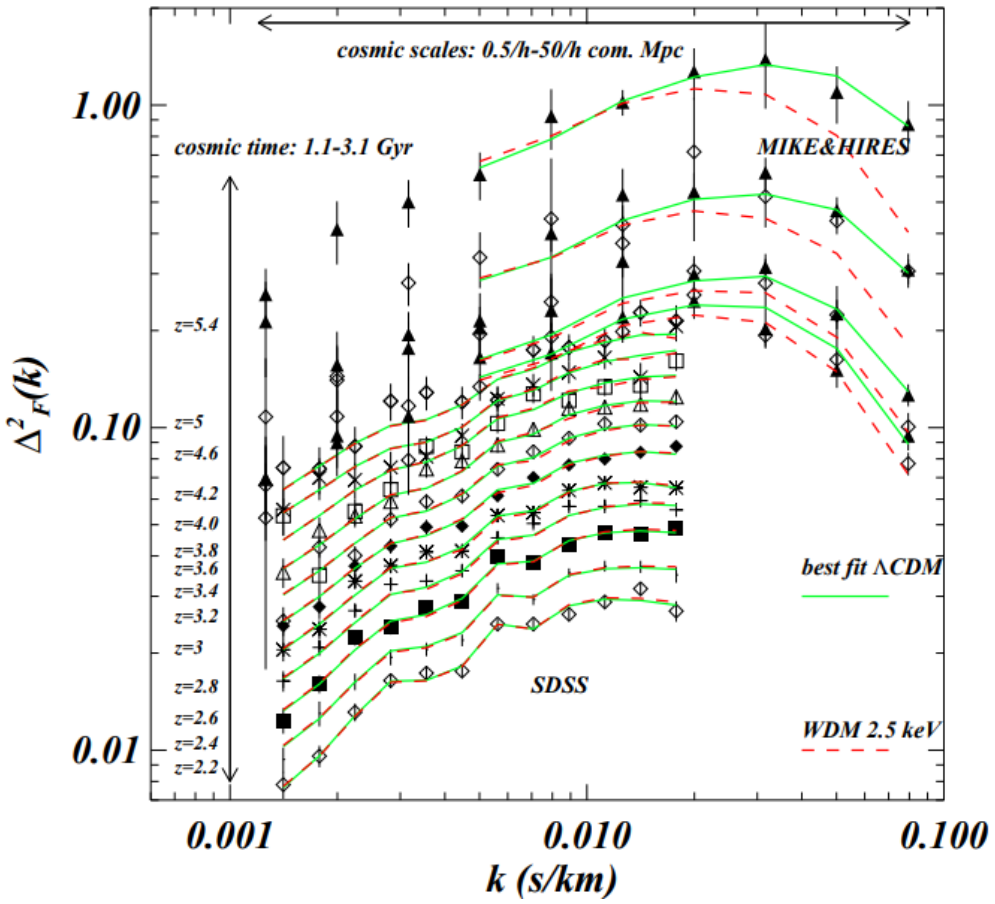


Figure 2.3: *FPS in dimensionless units from SDSS+HIRES+MIKE data. Credit: M. Viel et al. [53].*

- shift due to peculiar velocities;
- pressure smoothing;
- warm dark matter free-streaming.

In this study, we consider all the effects except the last one. The first two are one-dimensional (act only along LOS), while the third one is three-dimensional (acts in all spatial dimensions).

As a result, $k_c^{-2} \approx k_F^{-2} + k_b^{-2}$ [12], where k_b^{-2} accounts for the cut-off due to the thermal Doppler broadening. Worth noticing that peculiar velocities do not impact the position of the cut-off, as they do not contribute to the broadening of absorption lines.

To infer pressure filtering scale λ_F from the observed cut-off scale λ_c,

we need knowledge of the Doppler broadening λ_b. However, a significant complication comes from the uncertainty in measurements of the gas temperature, upon which λ_b depends [14, 50]. Different measurements produce contradicting values of γ, some of them even give $\gamma < 1$ [55], which corresponds to an inverted temperature-density relation inconsistent with the most appealing reionization models. As a result, different thermal histories (pressure effects) combined with different IGM temperatures can lead to the same observed value of the cut-off. Unfortunately, the FPS can not break this degeneracy.

To understand the problem, it is useful to estimate k_c, k_b, and k_F. The Doppler broadening can be calculated as [8]:

$$k_b \approx \frac{\sqrt{2}}{\Delta v_b} = \sqrt{\frac{m_H}{k_B T}} = 0.11 \left(\frac{T}{10^4 \mathrm{K}}\right)^{-1/2} \frac{\mathrm{s}}{\mathrm{km}}, \tag{2.11}$$

where T is the IGM temperature, $m_H = 1.67 \times 10^{-24}$g is the mass of Hydrogen atom, and $k_B = 1.38 \times 10^{-23} \mathrm{J}^{-1} \cdot \mathrm{K}$ is the Boltzmann constant.

For a simple estimation of the contribution from pressure smoothing, we can use the expression for the Jeans length instead of Eq.(1.15) [8]:

$$k_p \approx \frac{\sqrt{2}}{\Delta v_p} = \sqrt{2} \frac{2\pi}{H_0 \lambda_{J,0}} = 0.076 \left(\frac{T}{10^4 \mathrm{K}}\right)^{-1/2} \frac{\mathrm{s}}{\mathrm{km}}, \tag{2.12}$$

where $\sqrt{2}$ comes from the fact that the power spectrum is the square of $\tilde{\delta}_F(k)$.

In addition, it is important to consider the cut-off due to finite resolution. In observational data, this cut-off arises from the finite spectrometer resolution $\mathcal{R} = \Delta\lambda/\lambda = c/\Delta v_s$:

$$k_s = \frac{\sqrt{2}}{\Delta v_s} = 0.21 \frac{6.6\,\mathrm{km\,s^{-1}}}{\Delta v_s} \frac{\mathrm{s}}{\mathrm{km}} \tag{2.13}$$

For numerical simulations, the finite spacing between particles in the simulation box contributes to the cut-off:

$$k_{sim} = \frac{(1+z)}{H(z)} \frac{N^{1/3}}{L}, \tag{2.14}$$

where N is the number of particles in the simulation box of volume L^3. For instance, at $z = 5$, with $L = 20\,h^{-1}$ Mpc and $N = 512^3$ the numerical value of this quantity is approximately $0.27\,\mathrm{s\,km^{-1}}$. This value defines the lower

Version of February 16, 2024 – Created March 5, 2024 - 09:37

boundary on the simulation resolution required for clear observation of cut-offs coming from other effects.

The cut-off scales introduced above have close numerical values for realistic parameter settings. It shows the challenge of distinguishing contributions from different effects and interpreting the results. Needless to say, the situation becomes even more intricate with WDM.

The solution to this problem is to collect multiple parallel close LOS and probe the transverse direction. This eliminates !D effects acting along LOS (Doppler broadening and peculiar velocities) and keeps only 3D ones. This method is known as **the method of close quasar pairs**.

2.3 Quasar pairs

Close quasar pairs present an opportunity to explore the transverse direction, which is sensitive to pressure smoothing.

The method proposed in Rorai et al. (2013) [56] has the aim to determine λ_F. It is based on the fact that pressure smoothing is a 3D effect, acting equivalently in all dimensions, while thermal Doppler broadening and peculiar velocities are 1D effects, acting only along the LOS. Measuring correlations between close quasar pairs allows probing the transverse direction, which is affected only by 3D effects.

Depending on the transverse separation, different effects can be studied, which is discussed in [57]. In this study, we focus on pairs with separations $r_\perp \lesssim 300$ ckpc or $\theta = \frac{r_\perp}{H(z=3)} \lesssim 40''$, where the pressure effect is dominant. At much smaller scales, quasar pairs are highly correlated, providing limited information. At larger scales, the situation is the opposite - spectra are uncorrelated. The example of close quasar pairs spectra is presented in Figure 2.4.

Furthermore, the investigation of multiple close quasars (triplets and more) enables probing Hydrogen spatial distribution and estimating the anisotropy of the ionization source [57].

Understanding the origin of quasar pairs is important for their searches. Quasar pairs mostly form by chance, demonstrating their probabilistic nature. However, they can also emerge during the merger of two galaxies hosting supermassive black holes of similar masses and having separation $\Delta d \lesssim 10$ ckpc [59]. Most known quasars are located at $z \sim 3$, while the most distant one is observed at $z = 5.66$ and has transverse separation of $\Delta d = 7.3$ ckpc [59].

A collection of quasar pairs suitable for studying pressure smoothing (within the redshift range $2 \lesssim z \lesssim 4$ and transverse separations $r_\perp < 450$

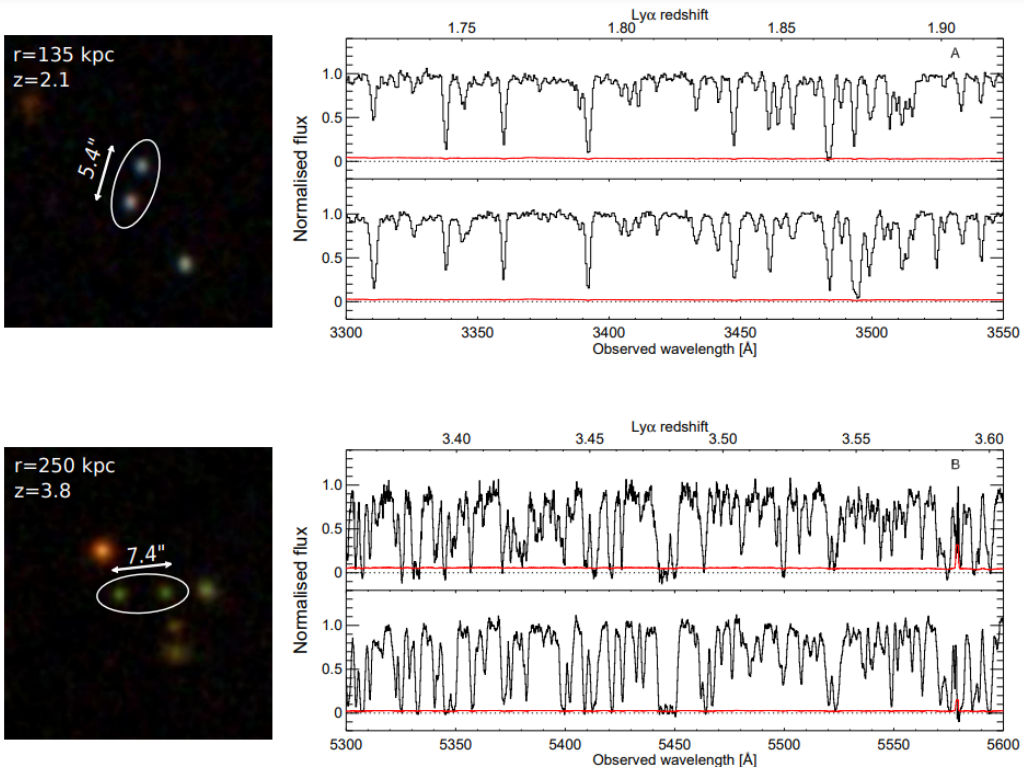


Figure 2.4: *Spectra of the quasar pairs SDSS J073522.43+295710.1, SDSS J073522.55+295705.0 (upper panel) and SDSS J102116.98+111227.6, SDSS J102116.47+111227.8 (lower panel). Credit: Rorai et al. (2013)[58].*

ckpc) is provided in [58]. None of them are gravitational lenses, and the overlap of their spectra is less than 30 % for preventing overlap between the Lyman-α and Lyman-β forests from each quasar. In total, there are 25 pairs available for analysis, which is sufficient for determining λ_F with 5% precision [56].

2.4 Phase difference and its properties

The main function to characterize quasar spectra is the FPS $P(k)$, which in the isotropic case is related to the 3D power spectrum by the following relation:

$$P_{3D}(k) = \frac{1}{2\pi}\frac{1}{k}\frac{dP(k)}{dk}. \tag{2.15}$$

However, this relation breaks down due to thermal Doppler broadening and redshift-space distortions [56].

Version of February 16, 2024 – Created March 5, 2024 - 09:37

For a quasar pair with a transverse separation $r_\perp$, the counterpart to the FPS is the cross-power spectrum $\pi(k, r_\perp)$:

$$\pi(k, r_\perp) = \Re[\delta\tilde{F}_1(k)\delta\tilde{F}_2^*(k)]. \tag{2.16}$$

This expression demonstrates that as separations become smaller $r_\perp \to 0$, then $\delta\tilde{F}_1(k) \to \delta\tilde{F}_2(k)$, and the cross-power spectrum coincides with FPS. Thus, $\pi(k, r_\perp)$ has properties of FPS in the longitudinal direction and of the correlation function in the transverse direction. This statement is illustrated by Figure 2.5.

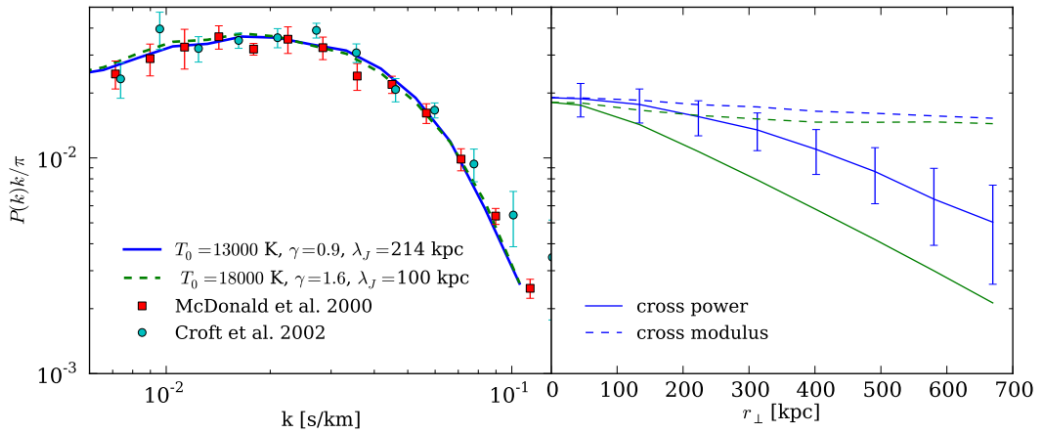


Figure 2.5: *The illustration of degeneracies between parameters (T_0, γ, λ_F). The left part represents 1D dimensionless FPS, while the right one shows a dimensionless cross-power spectrum and corresponding moduli. Credit: Rorai et al. (2013) [56].*

To eliminate dependence on poorly constrained parameters characterizing IGM (T_0 and γ) while keeping sensitivity to λ_F, we can use the following relation:

$$\delta\tilde{F}(k) = \rho(k)e^{i\theta(k)}. \tag{2.17}$$

Then the cross-power spectrum can be expressed as

$$\pi_{12}(k) = \rho_1(k)\rho_2(k)\cos(\theta(k)), \tag{2.18}$$

where $\theta(k) = \theta_1(k) - \theta_2(k)$.

Taking into account that $P(k) = \langle\rho^2(k)\rangle$, it is evident that all the information about T_0 and γ is encapsulated in the moduli $\rho_1(k)$ and $\rho_2(k)$.

Consider the influence of T_0 and γ on δF to understand the dependence on IGM thermal parameters. Eq.(2.7) suggests that the dependence on γ is weak in the constrained range $0.8 < \gamma < 1.6$. The effect of T_0 on δF

is mitigated by fixing the mean flux $\langle F \rangle$, in other words by selecting the appropriate photoionization rate Γ [14, 56].

Moreover, despite $T = T(v)$, thermal Doppler broadening is a close equivalent of convolution with a Gaussian kernel, which is a multiplication in Fourier space. Because phases are preserved by convolution with a symmetric kernel, the phases of $\pi(k, r_\perp)$ are influenced by λ_F rather than by T_0 and γ. This property allows studying the filtering scale by the phase difference θ:

$$\theta(k) = \arccos\left(\frac{\Re[\delta\tilde{F}_1(k)\delta\tilde{F}_2^*(k)]}{\sqrt{|\delta\tilde{F}_1(k)|^2|\delta\tilde{F}_2(k)|^2}}\right). \tag{2.19}$$

This observable $\theta(k)$ does not include the moduli of the Fourier transform of the flux, indicating its independence from T_0 and γ. However, it contains information about λ_F, which is the focus of this thesis.

Therefore, $\theta(k, r_\perp)$ characterizes the correlations between the spectra of two quasars. This random observable depends on two variables (wavenumber k and transverse separation $r_\perp$) and has values in the range $[-\pi, \pi]$. For further analysis, it is important to understand the properties of probability distribution of $\theta(k, r_\perp)$.

Consider two parallel LOS from two distant quasars, which are shown in Figure 2.6. These lines are separated by $r_\perp$ and probe regions filled with neutral Hydrogen. The Hydrogen cloud forms an angle φ with a perpendicular to the LOS, implying that the density peaks are separated by $L = r_\perp \tan\varphi$ in the longitudinal direction. For a given k mode, the phase difference between two consecutive peaks is $\theta' = kL = kr_\perp \tan\varphi$.

The condition $p(\theta')d\theta' = p(\varphi)d\varphi$, combined with the isotropy of space $p(\varphi) = \frac{1}{\pi}$, implies that θ' is distributed by:

$$p(\theta') = p(\varphi)\frac{d\varphi}{d\theta'} = \frac{1}{\pi}\frac{d}{d\theta'}\arctan\frac{\theta'}{kr_\perp} = \frac{1}{\pi kr_\perp}\frac{1}{1+(\frac{\theta'}{kr_\perp})^2}. \tag{2.20}$$

Thus, θ' follows a Cauchy distribution. To determine the distribution of θ, we need to consider that for all integer n, phases $\theta' + 2\pi n$ correspond to the same θ in the range $[-\pi, \pi]$. This leads to

$$P(\theta) = \sum_{n\in\mathbb{Z}} p(\theta + 2\pi n). \tag{2.21}$$

After summation, we obtain:

$$P(\theta) = P_{WC}(\theta) = \frac{1}{2\pi}\frac{1-\zeta^2}{1+\zeta^2-2\zeta\cos(\theta)}, \tag{2.22}$$

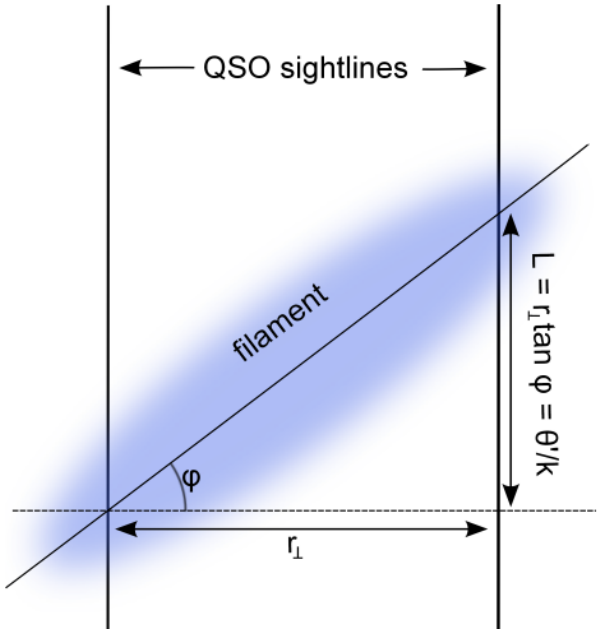


Figure 2.6: *Illustration to the heuristic argument for determining phase difference distribution. Credit: Rorai et al. (2013) [56].*

where ζ is a parameter related to $kr_\perp$ as $\sinh(kr_\perp) = \frac{1-\zeta^2}{2\zeta}$ and taking values from 0 to 1. When $\zeta \to 0$, the distribution becomes uniform, the probability distribution $P(\theta) = \frac{1}{2\pi}$, corresponding to completely uncorrelated spectra. Conversely, when $\zeta \to 1$, it tends towards a Dirac delta $\delta(\theta)$, corresponding to identical spectra. The examples are shown in Figure 2.7.

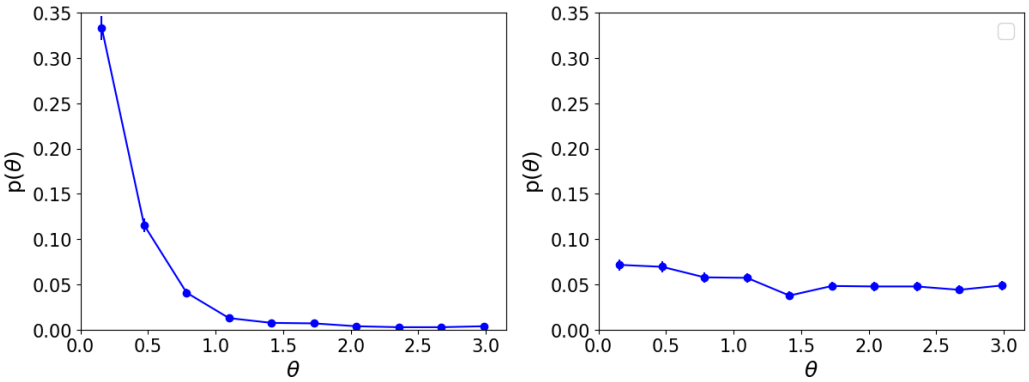
P_{WC} is a good approximation for the distribution of θ [56]. However, for large θ, it yields a higher probability than obtained from simulations. This discrepancy may be caused by neglecting spherical halos in the model.

To determine which values of λ are useful for inferring λ_F, we consider the following quantity:

$$\langle \cos\theta(k, r_\perp) \rangle = \int_{-\infty}^{\infty} P(\theta(k, r_\perp)) \cos\theta(k, r_\perp) d\theta. \tag{2.23}$$

The parameter ζ having values in the range $[0; 1]$. Taking into account that density fluctuations are Gaussian and that for $k \ll \frac{1}{aHr_\perp}$, the moduli $\rho_1(k)$ and $\rho_2(k)$ are nearly equal, we can obtain:

$$\langle \cos\theta(k, r_\perp) \rangle \approx \frac{\pi(k, r_\perp)}{P(k)} = \frac{\int_k^\infty dqq J_0(r_\perp \sqrt{q^2 - k^2}) P_{3D}(q)}{\int_k^\infty dqq P_{3D}(q)}. \tag{2.24}$$



(a) PDF(θ) when $kr_\perp \to 0$. *(b) PDF(θ) when $kr_\perp \to \infty$.*

Figure 2.7: *Examples of PDF(θ) for different transverse separations $r_\perp$ and wavenumbers k.*

The Bessel function J_0 appears due to the transverse separation between LOS. For large q, we have:

$$\langle \cos\theta(k, r_\perp) \rangle \approx \frac{\int_k^{k_\perp} dqq J_0(r_\perp \sqrt{q^2 - k^2}) P_{3D}(q)}{\int_k^{\infty} dqq P_{3D}(q)}, \tag{2.25}$$

where $k_\perp = \frac{x_0}{aHr_\perp}$ and $x_0 = 2.4048$ is the first zero of $J_0(x)$. This expression holds for $k \ll k_\perp$. It shows that 3D modes corresponding to $\lambda_F \sim 100$ ckpc contribute significantly to $\langle \cos\theta(k, r_\perp) \rangle$.

It means, that the relevant separations are those close to the filtering scale λ_F, which is $\sim 30 - 70$ ckpc.

Chapter 3

Simulations

3.1 Hydrodynamic simulations

The standard approach to study IGM properties using the Lyman-α forest is to perform a series of simulations varying parameters of interest and comparing outcomes with the observational data.

For studying EoR, we need cosmological simulations of galaxy formation that capture the non-linear regime of perturbation growth. In this case, it is essential to implement a realistic model of galaxy formation.

In this study, we used high-resolution hydrodynamical simulations of galaxy formation. These simulations contain not only dark matter but also baryonic matter, adding to a galaxy formation model processes such as gas cooling, star formation, dynamics of the interstellar medium, star feedback, supermassive black holes, AGNs, and cosmic rays [60].

Due to the complexity and diversity of effects and dynamics involved, full hydrodynamical simulations are computationally expensive. To address this problem, N-body methods were developed. They replace DM particles of mass m with a large number N of huge (composed) particles characterized by their coordinates $\mathbf{x}_i$ ($i = 1...N$) and velocities $\dot{\mathbf{x}}_i$ ($i = 1...N$). Each such particle has a mass $M = \frac{\rho_0 V}{N}$, where V is the comoving volume of the simulation and ρ_0 is the particle density satisfying:

$$\rho_0 = m \int d\mathbf{x} f(\mathbf{x}, \dot{\mathbf{x}}, t), \tag{3.1}$$

where $f(\mathbf{x}, \dot{\mathbf{x}}, t)$ is the DM distribution function. M is also referred to as the mass resolution of the simulation.

The simulation is based on the widely accepted ΛCDM cosmological model, where dark matter and dark energy account for $\sim 95\%$ of the total

energy density, while the remaining $\sim 5\%$ is baryonic matter. Dark matter is considered cold and collisionless, while dark energy is represented by the cosmological constant Λ. The parameters of the model are set using astronomical observations.

The initial conditions represented by a homogeneous background with Gaussian perturbations, characterized by the matter power spectrum $P_M(k) = Ak^n|T(k)|^2$ with a spectral index $n \approx 1$. The positions of dark matter particles $\mathbf{x}$ are set by evolving them in the repulsive gravity from defined initial positions $\mathbf{q}$ (either Cartesian or 'glass'-like configurations) until reaching equilibrium in comoving coordinates [60]. Particle velocities $\mathbf{v}$ are computed time derivatives of coordinates:

$$\mathbf{x} = \mathbf{q} + D(t)\Psi(\mathbf{q}), \qquad \mathbf{v} = \frac{dD(t)}{dt}\Psi(\mathbf{q}), \tag{3.2}$$

where $D(t)$ is the linear growth factor and $\Psi(\mathbf{q})$ is the curl-free displacement field satisfying $D(t)\nabla \cdot \Psi = -\delta_M$. This method is known as the Zel'dovich approximation [61].

The same approach is applied to initialize baryon particles. The baryon temperature is set equal to the CMB temperature at the redshift of simulation start $z \sim 100$, when density perturbations are still linear. The simulation box is built as a uniform periodic large volume.

The dynamics of dark matter particles, which are the most important component of cosmological simulations, is modeled by the collisionless Boltzmann and Poisson equations:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \dot{\mathbf{x}}\frac{\partial f}{\partial \mathbf{x}} - \frac{\partial \Phi}{\partial \mathbf{x}}\frac{\partial f}{\partial \dot{\mathbf{x}}} = 0, \qquad \Delta\Phi = 4\pi G \int f d\mathbf{v}, \tag{3.3}$$

where Φ is the collective gravitational potential. Those equations are complemented by the Friedmann equations. The majority of simulations use Newtonian gravity and set periodic boundary conditions to satisfy the cosmological principle. Eq. (3.3) are solved by the method of characteristics, which has different realizations [60]:

This method defines forces on a mesh and calculates differential operators using finite difference approximations. The force computation from the Poisson equation is facilitated by the Fast Fourier Transform, resulting in a method complexity of $O(N_g \log N_g)$, where N_g is the number of grid points. Particle positions are interpolated onto the grid, and the gravitational potential is solved using Fast Fourier Transforms, enabling efficient integration of particle trajectories.

Version of February 16, 2024 – Created March 5, 2024 - 09:37

- Particle-Particle (PP) method is the oldest and most straightforward method, where the integration of equations of motion is done by direct summation of forces acting on particles. However, its computational complexity of $O(N^2)$ makes it unsuitable for systems with a large number of particles.

- Particle-Mesh (PM) method defines forces on a mesh and calculates differential operators using finite difference approximation. The gravitational potential is calculated from the Poisson equation using Fast Fourier Transforms, which makes the method complexity equal to $O(N_g \log N_g)$, where N_g is the number of grid points. The equations of motion are integrated numerically using standard schemes like the leapfrog or Verlet method.

- Tree methods distinguish between nearby and distant neighboring particles. They simplify dynamics without producing significant computational errors. A commonly utilized approach is the Barnes-Hut (top-down) algorithm [62], which organizes particles into a hierarchical tree structure by recursively dividing the simulation volume into octants until each node contains a small (predefined) number of particles. Gravitational forces are calculated by summing up the contributions from nearby nodes.

- P^3M methods combine PM for long-range interactions and PP methods for short-range interactions. It gives high accuracy for low computational cost. The difference between this method and the TreePM is shown in Figure 3.1.

- Other complex methods combining approaches mentioned above.

Boltzmann equations for baryonic matter contain collision terms describing various gas properties such as pressure and viscosity. These equations are expressed as:

$$\frac{\partial f}{\partial t} + \dot{\mathbf{x}}\frac{\partial f}{\partial \mathbf{x}} - \frac{\partial \Phi}{\partial \mathbf{x}}\frac{\partial f}{\partial \dot{\mathbf{x}}} = \left[\frac{df}{dt}\right]_{coll}. \tag{3.4}$$

These equations are complemented by continuity, momentum, and energy equations. This system can be solved by either of those three methods: Eulerian, Lagrangian, and Lagrange-Eulerian. In the Eulerian approach, the fluid is considered to move in a static grid, which is an established technique in numerical hydrodynamics. In the Lagrangian method, each particle represents a region of the fluid and has all its properties. It

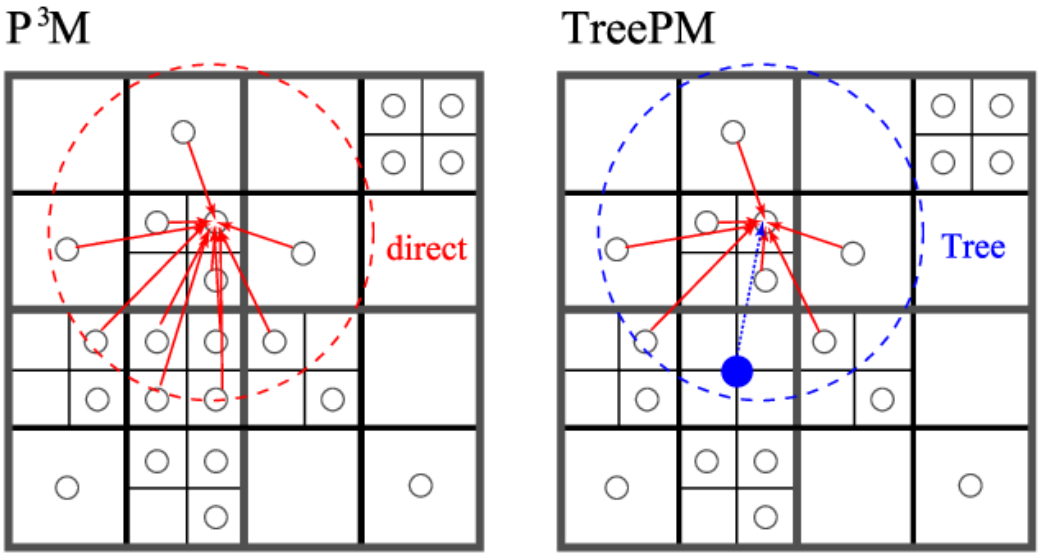


Figure 3.1: *Difference between force calculation in the P^3M and the TreePM methods. Circles represent the cut-off radii, where the short-range forces are calculated. Credit: Ishiyama et al. (2012) [63].*

is done by Smoothed Particle Hydrodynamics (SPH), which divides the fluid into sampling particles and computes fluid characteristics by particle interpolation. The Lagrange-Eulerian methods allow an unconstrained choice of grid velocity [60].

For accurate evaluation of external terms in hydrodynamic equations for baryons, astrophysical processes must be considered. In the study of Lyman-α absorption systems leading to Hydrogen reionization, essential processes are gas cooling, star formation, photoheating, and photoionization. They are typically modeled via semi-analytical methods due to their computational complexity.

The main cooling mechanisms are collisional excitation (most significant at $T \sim 10^4$ K), bremsstrahlung (dominant at $T \sim 10^5$ K), recombination, and collisional ionization. Because all those processes are two-body, cooling is more efficient in denser regions, assisting the formation of compact objects.

Cooling particles are falling into gravitational potential wells, resulting in regions of cold and dense gas. They became centers of star formation, represented by radiating point particles. However, modeling the details of those processes is still a computational challenge.

The most important part of EoR simulations is modeling star emission, which ionizes IGM. It can be done by straightforwardly solving radiative

transfer equations, but this task is extremely computationally expensive. However, a simpler solution was proposed. The idea is to introduce homogeneous and isotropic UV radiation, which is supported by observations. It is called Ultraviolet Background (UVB), which has an intensity that depends only on frequency. This function includes the radiation from all the sources, such as stars and quasars. More sophisticated simulations [64] consider scenarios of patchy reionization [65], where the UVB is not uniform.

Moreover, realistic simulations of reionization must incorporate additional important sources such as quasars and AGNs. They are implemented in two modes: quasar mode and radio mode [60]. The former consists of radiation emitted by a supermassive black hole when it accretes surrounding gas. It then injects significant energy and momentum into the environment, affecting galaxy evolution. The latter mode is implemented by collimated jets of relativistic particles emitted by AGN, which prevent cooling in massive halos and regulate star formation. This process is simulated by considering the impact of radio-emitting sources on surrounding gas, such as synchrotron emission.

3.2 Description of our simulations

We used simulation based on `Gadget-2` code [66], which models the dynamics of dark matter and baryons in an expanding Universe, accounting for gravity and hydrodynamical processes. It is a TreePM SPH code, which uses the Tree method to handle short-range interactions and the PM method to calculate long-range gravitational forces, combined with SPH for modeling gas dynamics. In addition, `Gadget-2` includes energy and entropy conservation in dissipation-free regions.

The initial conditions for simulation were obtained employing the `2LPTic` code [67]. This algorithm is based on second-order Lagrangian perturbation theory (2LPT), which is imprinted in its name. The code takes as input the primordial power spectrum and the cosmological parameters. The former is obtained from the `CAMB` (Code for Anisotropies in the Microwave Background) code [68], which is a software package calculating different cosmological observables, one of which is the matter power spectrum. The cosmological parameters were taken from Planck cosmology [69], summarized in Table 3.1.

The simulation begins at $z = 99$, which is a high enough redshift that the density perturbations are still in the linear regime. The box size is $L = 20$ Mpc h^{-1}, and the number of dark matter particles is $N = 1024^3$.

Version of February 16, 2024 – Created March 5, 2024 - 09:37

Ω_0	0.308 ± 0.012
Ω_Λ	0.692 ± 0.012
$\Omega_b h^2$	0.02226 ± 0.00023
Ω_m	0.308 ± 0.012
H_0	67.81 ± 0.92
n_s	0.9677 ± 0.0060
σ_8	0.8149 ± 0.0093
Y_p	$0.251^{+0.040}_{-0.039}$

Table 3.1: *Cosmological parameters constrained by TT+lowP+lensing Plank observations [69]. Errors indicate 68% confidence level.*

The simulation includes only collisionless cold dark matter that interacts only gravitationally.

Dark matter particles are created in a Lagrangian way, which is a traditional approach in SPH. Baryons are modeled in the form of an ideal gas and initialized on a 'glass'-like grid.

As a hydrodynamical simulation, `Gadget-2` code contains not only gravitational and gas dynamics but also processes crucial for EoR. Those processes include radiative cooling and heating, star formation, and feedback processes. In addition, the algorithm can handle AGNs, magnetohydrodynamics, black hole growth, etc.

The IGM is ionized by homogenous and isotropic UVB, which changes with time and [70]. This approach yields a low neutral Hydrogen fraction (optically thin region). The UVB is implemented by using a cooling function that depends on the redshift and the metallicity of the gas. The input values of this function are the photoionization and photoheating rates of Hydrogen and Helium, the collisional ionization, recombination, and bremsstrahlung. With the UVB background, the cooling and heating equation can be solved.

In this study, we used two thermal histories of reionization: Late and Early from [71]. The Late model assumes a Hydrogen reionization at $z_{reion,HI} = 6.55$, while the Early has $z_{reion,HI} = 9.70$. The Helium reionization occurs at $z_{reion,HeII} = 3.0$ in both scenarios. The models adopt the heat input for Hydrogen reionization $\Delta T_{HI} = 2 \times 10^4$ K and for Helium reionization $\Delta T_{HeII} = 1.5 \times 10^4$. The redshift dependency of the full ionized has an analytical form $\langle x_{HeIII} \rangle = 1.0 - \arctan(z - z_{reion,HI})$, which is consistent with common HeII reionization models. In addition, both histories were created to be consistent with the CMB constraints on the electron scattering optical depth $\tau_e = 0.078 \pm 0.019$ from [72]. Figure 3.2 shows the

Version of February 16, 2024 – Created March 5, 2024 - 09:37

evolution of IGM thermal parameters T_0 and γ for both Early and Late reionization histories, while Figure 3.3 represents temperature-density relations at the redshift $z = 4.6$ of our snapshot.

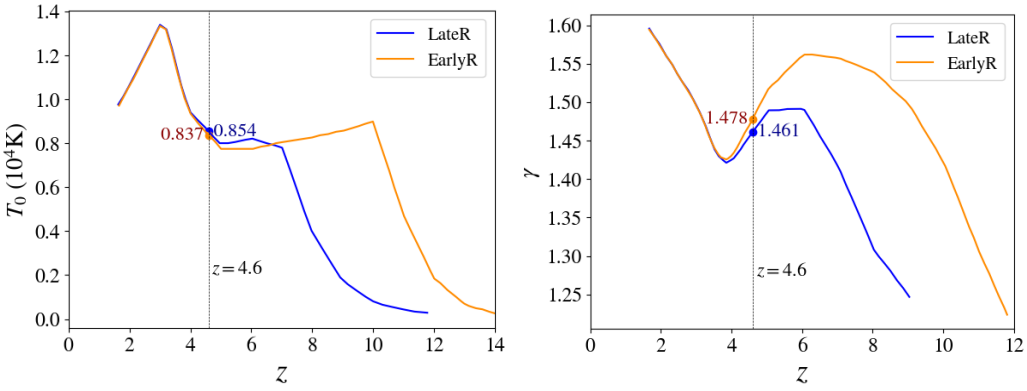


Figure 3.2: *Redshift evolution of thermal parameters characterizing IGM for Early and Late reionization histories. The marked points indicate values of mean IGM temperature T_0 and slope γ of temperature-density relation at the redshift $z = 4.6$ of the snapshot used in this study.*

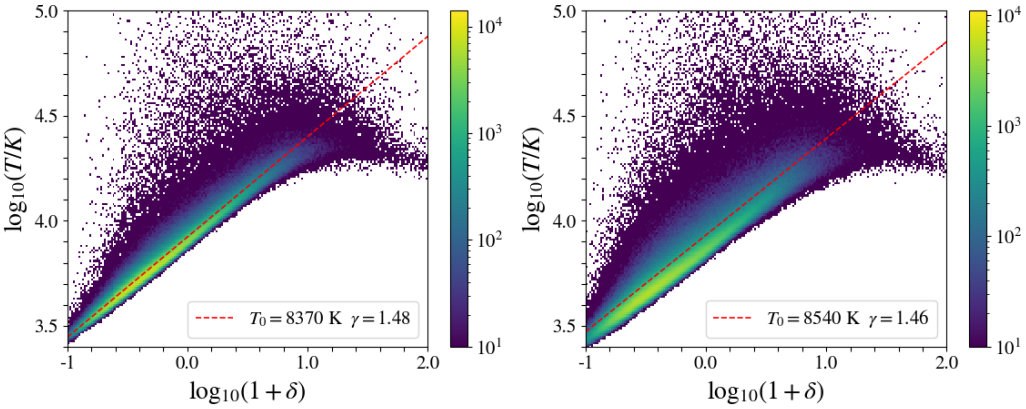


Figure 3.3: *Temperature-density relations for Early (left) and Late (right) reionization histories at $z = 4.6$.*

Figure 3.4 illustrates FPS with only pressure effect for both thermal histories at $z = 4.6$. The Early reionization has a cut-off at $k = 0.1393\frac{\text{s}}{\text{km}}$, while the Late one has a cut-off at $k = 0.1893\frac{\text{s}}{\text{km}}$. Those values can be converted into filtering scales $\lambda_F = 47.6$ ckpc and $\lambda_F = 36.5$ ckpc correspondingly.

The main parameters of simulations are provided in Table 3.2. More detailed information can be found in [71]

Version of February 16, 2024 – Created March 5, 2024 - 09:37

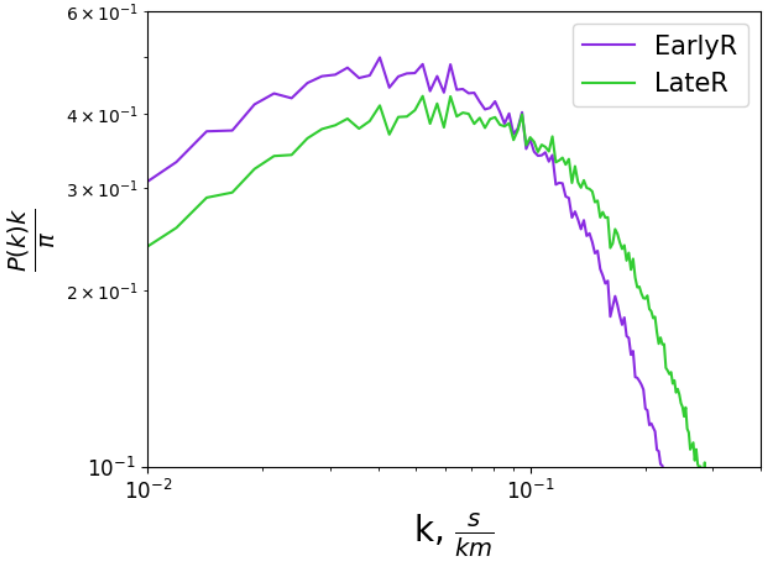


Figure 3.4: *Flux power spectrum for Late and Early reionization histories at* $z = 4.6$ *accounting only for pressure smoothing.*

Simulation	L (Mpc/h)	N	$z_{reion,HI}$	$z_{reion,HeII}$	ΔT_{HI} (K)	ΔT_{HeII} (K)
LateR	20	1024^3	6.55	3.0	2×10^4	1.5×10^4
EarlyR	20	1024^3	9.70	3.0	2×10^4	1.5×10^4

Table 3.2: *Parameters of hydrodynamical simulations utilized in this work. The first column is named according to [71]. All simulations contain CDM and use cosmological parameters shown in Table 3.1.*

3.3 Mock spectra

In this thesis, we study correlations between Lyman-α forests of close quasar pairs to distinguish different thermal histories and get insights into the Hydrogen pressure smoothing scale. For this purpose, we needed to generate a sufficient amount of LOS with various transverse separations for different thermal histories.

To extract the Lyman-α skewers from the simulation, we used the `SPECWIZARD` tool, which utilizes the technique from [73]. It generates a grid of LOS in the simulation box, with a specified number, spacing, and orientation. Each LOS contains gas properties, obtained by interpolation from the nearest SPH particles with a chosen kernel (usually cubic spline). These properties include density, temperature, peculiar velocity, and ionization

fractions of Hydrogen and Helium. The optical density is computed by integrating the cross-sections for Lyman-α scattering of Hydrogen along the LOS, taking into account thermal broadening and peculiar velocities. The resulting flux is equal to the product of the transmission from Equation 2.8 and the continuum level of the background source, such as a quasar.

Simulated LOS are stored as arrays of data points (pixels) corresponding to a specific location along the line of sight. The distance along the LOS is defined by Hubble velocity, divided into bins of size Δv. It means that all the operations with quantities along LOS, such as interpolation and integration, are performed in bins via the Gaussian method [74]. For example, the optical depth in the velocity bin $v(i)$ is computed as:

$$\tau(i) = \sum_X \sum_i \sigma_X \sqrt{\frac{m_X c^2}{2\pi k_B T_j}} \rho_X(j) a \Delta v \exp\left(- \frac{m_X (v(i) - v(j))^2}{2 k_B T_j} \right), \quad (3.5)$$

where the summation represent contributions from different species X and bins $v(j)$.

3.4 Dataset and parametrs

Our dataset contains sets of 1000 LOS at a single redshift of $z = 4.6$, which are shifted along the x-axis. The LOS sets have transverse separations of 50 ckpc, 100 ckpc, 200 ckpc, and 300 ckpc. This range is close to the scale of pressure smoothing, which is a subject of our study.

We have generated these sets for both of the thermal histories described above. Each individual spectrum within these sets contains 1889 pixels, each of a size $\Delta v = 1.4\,\mathrm{km\,s^{-1}}$. However, we restricted our analysis to 210 pixels ($k \leq 0.5\,\mathrm{s\,km^{-1}}$) due to the observational limitations and smoothing of smaller structures. This data allows us to study correlations between the Lyman-α forests of quasar pairs at varying separations and pressure scales. Examples of LOS characteristics from the simulation with Early reionization scenario are shown in Figure 3.5.

Furthermore, we varied the thermal parameters characterizing IGM $\{T_0, \gamma\}$. The collection of used sets can be found in Table (3.3).

In addition, we gradually added various effects into optical depth calculations. Initially, we considered the ideal case with only the pressure effect (p). Then we added temperature broadening (pT), and, lastly, peculiar velocities (pTv). This incremental approach allows tracking the impact of each effect.

Version of February 16, 2024 – Created March 5, 2024 - 09:37

T_0 (K)	γ
7 000	1
7 000	1.4
11 000	1
11 000	1.3
11 000	1.4
11 000	1.6
16 000	1
16 000	1.4
20 000	1
20 000	1.3
20 000	1.4
20 000	1.6

Table 3.3: *Grid of used thermal parameters characterizing IGM.*

Because `SPECWIZARD` calculates flux for each spectrum, we needed to take a step back to achieve our goals. `SPECWIZARD`'s calculations are performed by inserting the parameters $\{T_0, \gamma\}$ from the thermal history at a chosen redshift. However, these were the exact parameters we intended to vary. Consequently, we only utilized overdensities and peculiar velocities from the skewers. We then computed the temperature-density relation with a desired set $\{T_0, \gamma\}$. This relation was subsequently used to calculate the neutral Hydrogen density, which in turn was used to compute the optical depth with various effects. The resulting flux was thus obtained. A detailed description of this procedure can be found in Appendix A.

Upon obtaining the flux, we followed the methodology outlined in [56] and calculated the phase difference using Eq.(2.19). Having the set of those angles, we then perform our analysis.

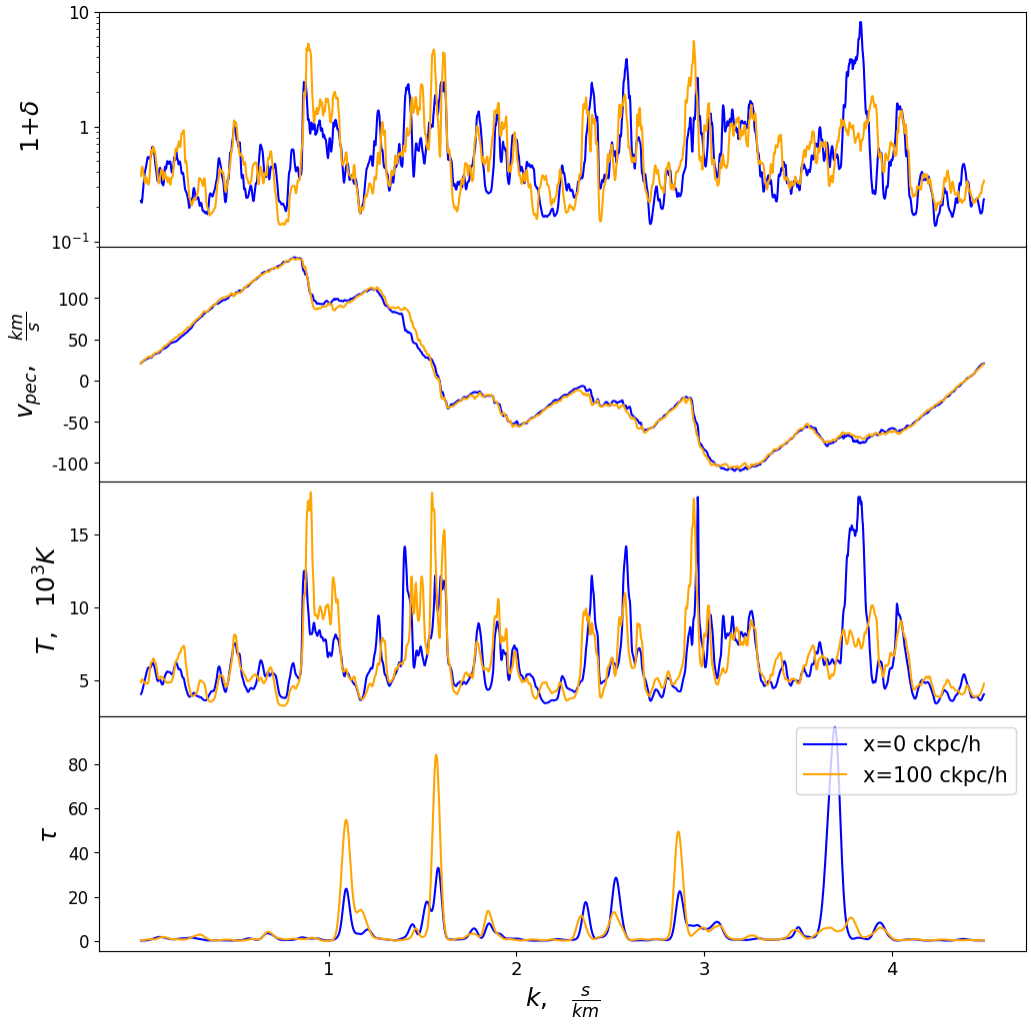


Figure 3.5: *Overdensity $1+\delta$, peculiar velocity v_{pec} (s/km) along LOS, temperature T (K), and optical depth τ for two random parallel LOS separated by $r_\perp = 100$ ckpc h^{-1}. LOS were extracted using* `SPECWIZARD` *from the $z = 4.6$ snapshot of simulation with Early reionization.*

Chapter 4

Results

4.1 Kolmogorov-Smirnov test

The pressure effect is expected to manifest itself at some characteristic pressure smoothing scale. To distinguish between different thermal histories we then need to perform an analysis at all observable scales, looking for the transitions. For this purpose, we decided to use the two-sample **Kolmogorov-Smirnov test**, which is a powerful and general tool.

The Kolmogorov-Smirnov test is a nonparametric test to determine the probability of two samples originating from the same distribution. It can be used only to unbinned distributions of a single variable x. The main advantages of the test are its sensitivity to the shape of cumulative distribution functions (CDFs) and robustness to the shape of the common distribution and reparametrization of x. In addition, the Kolmogorov-Smirnov test handles discontinuity, heterogeneity, and dependence across samples. However, it has limited sensitivity to the tail part of distributions.

The statistics of the test is the maximum absolute difference $D_{n,m}$ (see Figure 4.1) between CDFs $S_{1,n}(x)$ and $S_{2,m}(x)$ of the two samples:

$$D_{n,m} = \sup_x |S_{1,n}(x) - S_{2,m}(x)|, \tag{4.1}$$

where n and m are sample sizes. The higher $D_{n,m}$, the lower the probability that samples come from the same distribution.

The p-value corresponding to the observed difference $D_{n,m}$ is calculated from the probability distribution of $D_{n,m}$ for the null hypothesis (both samples were derived from the same distribution). The approximate expression for large effective sample size $n_{eff} = \frac{nm}{n+m}$ is provided in Stephens

(1970) [75]:

$$\mathrm{P}(D > D_{n,m}) = 1 - P_{KS}\big(d \leq D_{n,m}[\sqrt{n_{eff}} + 0.12 + 0.11/\sqrt{n_{eff}}]\big), \quad (4.2)$$

where $P_{KS}(X \leq x) = \frac{\sqrt{2\pi}}{x} \sum_{k=1}^{\infty} \exp\Big(-\frac{(2k-1)^2\pi^2}{8x^2}\Big)$ is the CDF of Kolmogorov-Smirnov distribution. This equation shows that the test becomes more sensitive to the deviations from the null hypothesis while adding more data points due to the increase in CDF accuracy.

To interpret the results of this test, it is more convenient to use the quantity $c(\alpha)$, where α is the level of rejection. It is related to $D_{n,m}$ as:

$$c(\alpha) = D_{n,m}\sqrt{\frac{nm}{n+m}}. \quad (4.3)$$

The larger the α, the more probable is the null hypothesis. The rejection level α is extracted from the relation $c(\alpha) = \sqrt{-\frac{1}{2}\ln\frac{\alpha}{2}}$ [76]. Some values are presented in Table 4.1:

α	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001
$c(\alpha)$	1.073	1.138	1.224	1.358	1.48	1.628	1.731	1.949

Table 4.1: *Example values of $c(\alpha)$ and corresponding rejection levels α. Credit: Wikipedia.*

In our study, we calculated $c(\alpha)$ for two sets of phase differences at each scale k and analyzed the resulting function.

4.2 Analysis

The main idea is to apply the Kolmogorov-Smirnov test for different thermal histories and parameters characterizing IGM and find scales, where there are transitions in $c(\alpha)$. Thermal histories are considered distinguishable when those scales exist and there is no difference while changing $\{T_0, \gamma\}$ and accounting for the effects of thermal broadening and peculiar velocities, that are present in real observational data.

To verify the validity of our results, we first performed a test to estimate the mean level of rejection in our 1000 spectra sample. The reason is the small sample size, resulting in some statistical fluctuations that may cause deviations from $\alpha = 1$ even for data with identical parameters. For this

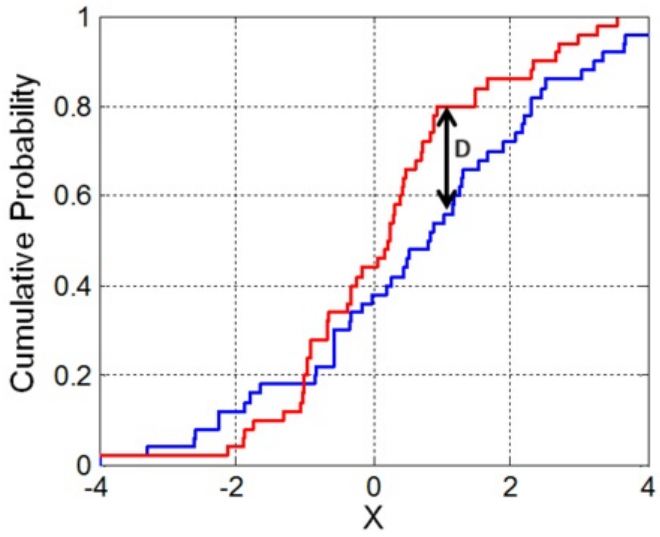


Figure 4.1: *Illustration to the explanation of the Kolmogorov-Smirnov test. Adapted from Wikipedia.*

purpose, we split our sample of θ into two subsamples of equal size and applied the Kolmogorov-Smirnov test to them. Then, we set the resulting $c(\alpha)$ as the maximum value accepting the null hypothesis (subsamples came from the same distribution).

Figure 4.2 shows the results of this preliminary test. For any choice of $\{T_0, \gamma\}$ and for both thermal histories, the $c(\alpha)$ lies in 1σ confidence interval, indicating that $\alpha \geq 0.32$. Moreover, because the subsample size is half the total sample, we expect smaller $c(\alpha)$ and smoother peaks in the final results. It means that there is no statistical difference between those two subsamples, and we can proceed with our analysis.

4.3 Independency of thermal broadening

In this section, we present our findings about the impact of varying thermal parameters characterizing IGM. Our work is based on a similar analysis by Rorai et al. (2013) [56], which uses PDFs of phase differences as the main statistics. The goal is to compare the Kolmogorov-Smirnov test results with those from the paper.

For our purposes, we took a $\{T_0, \gamma\}$ grid described in Section 3.4 and generated θ samples for both thermal histories using the algorithm outlined in A. Then we applied the Kolmogorov-Smirnov test varying either

mean IGM temperature T_0 or slope of temperature-density relation γ to assess their individual impacts.

Figures 4.3 and 4.4 show the results for different T_0 and γ, respectively. The observed differences lie in the 1σ region when peculiar velocities are excluded, indicating minimal statistical difference. Adding peculiar velocities increases differences noticeably, but even in this case, they are within the 2σ, which corresponds to statistical similarity. Figure 4.5 shows examples of PDF(θ), corresponding to Figure 4.3. It demonstrates that PDFs are almost identical, confirming the conclusion from [56].

Remarkably, changes in the mean temperature T_0 cause more significant differences than changes in slope γ. This can be explained by their contributions to the FGPA (2.6), where the slope is present only in the exponent, while the temperature is in the base of the power.

Overall, our results agree with those of Rorai et al. (2013) [56]. However, there are some slight differences. Firstly, they used a dark-matter-only simulation and mimicked the pressure effect by smoothing dark matter distribution with the cubic kernel. This method is computationally efficient but lacks the precision of the full hydrodynamical simulation used in this work. Secondly, they argued that thermal broadening does not affect phases at all because it is a convolution with a Gaussian kernel, which is not true in general. The temperature and the peculiar velocities vary along the LOS, causing small fluctuations in the exponent. Thirdly, there is a non-linear relation between flux and optical depth, which can distort the phases. Lastly, the most noticeable impact comes from peculiar velocities, which were not considered in [56]. This effect is not well understood and requires further investigation.

4.4 Distinguishing between thermal histories

In this section, we discuss our results considering different thermal histories. Using the same methodology as before, we compare samples with identical temperature-density relation parameters but with different pressure smoothing (reionization history) from Section 3.4.

Figure 4.6 shows the results of the Kolmogorov-Smirnov test for some sets of thermal parameters $\{T_0, \gamma\}$. The most significant difference is observed at small separations, comparable to the pressure smoothing scale. It exceeds 3σ, clearly indicating that this method can distinguish between various thermal histories. Important to note that this conclusion does not depend on the influence of thermal Doppler broadening and peculiar velocities.

Version of February 16, 2024 – Created March 5, 2024 - 09:37

The prominent feature of these plots is the presence of two distinct peaks when peculiar velocities are ignored. One peak is observed at larger scales ($k \sim 0.05$ s/km), while the other corresponds to smaller scales ($k \sim 0.35$ s/km). The small-scale effect can be explained by the fact that higher temperatures, which correspond to earlier reionization histories, cause a greater expansion of Hydrogen clouds. Then the structures of varying shapes and orientations are expanded more and have a higher probability of intersecting both LOS. The large-scale effect originates from a faster transition to a uniform distribution, defined by the minimum size of Hydrogen clouds. This scale is larger for hotter thermal histories, implying the smaller transition k. Interestingly, the second peak vanishes when peculiar velocities are taken into account.

Figure 4.7 supports statements above, representing phase difference PDFs at $r_\perp = 50$ ckpc without accounting for peculiar velocities. At larger scales, the hotter thermal history has more correlated spectra. Conversely, correlations disappear rapidly at smaller scales.

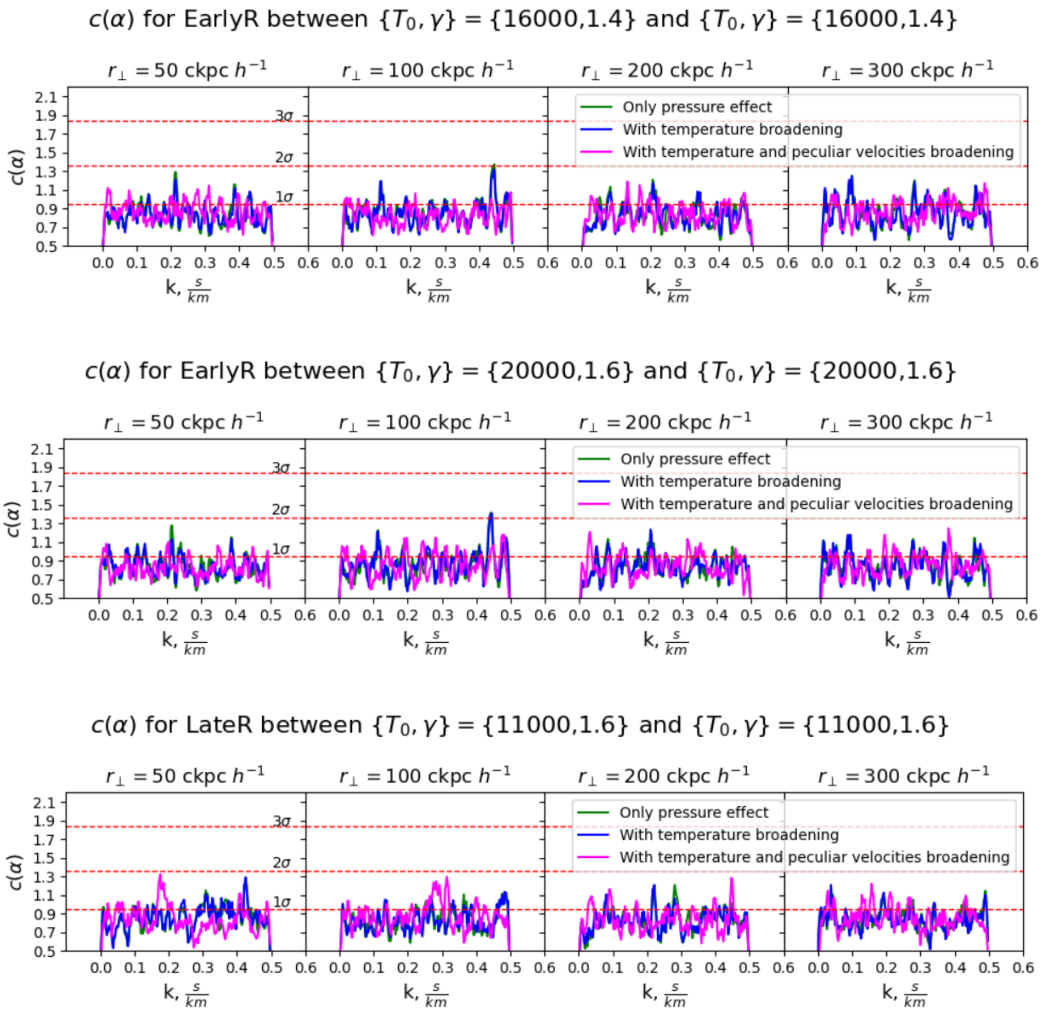


Figure 4.2: *Results of Kolmogorov-Smirnov test for 500 LOS subsamples, represented by $c(\alpha)$ for studied values of k. Each panel corresponds to different $\{T_0, \gamma\}$ parameters and thermal histories and includes various effects (pressure smoothing, temperature broadening, peculiar velocities). Red dashed lines represent confidence levels in σ (1σ: 68%, 2σ: 95%, 3σ: 99.7%).*

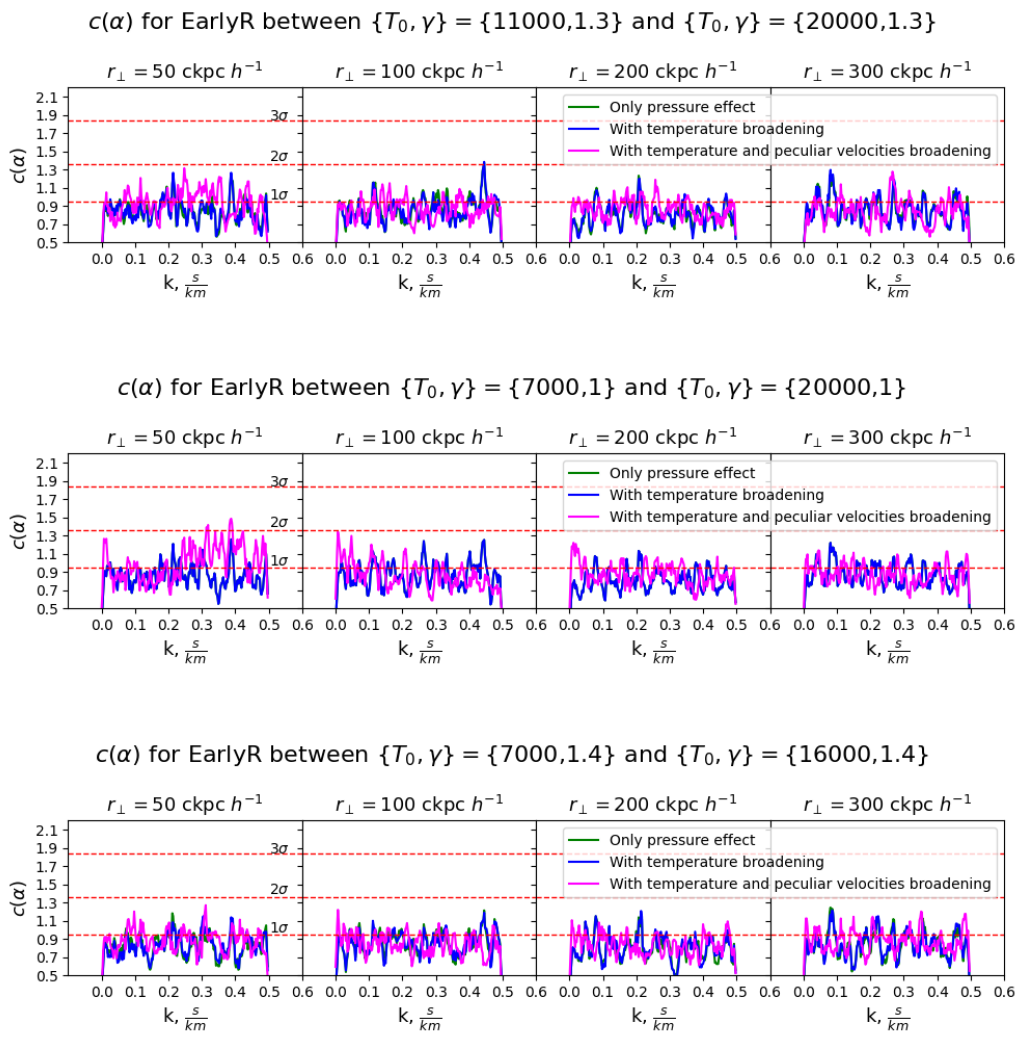


Figure 4.3: *Results of Kolmogorov-Smirnov test for varying mean IGM temperatures T_0 for Early reionization, represented by $c(\alpha)$ for studied values of wavenumbers k. Each panel corresponds to different sets $\{T_0, \gamma\}$ of parameters characterizing thermal state of IGM and includes various effects (pressure smoothing, temperature broadening, peculiar velocities). Red dashed lines represent confidence levels in σ (1σ: 68%, 2σ: 95%, 3σ: 99.7%).*

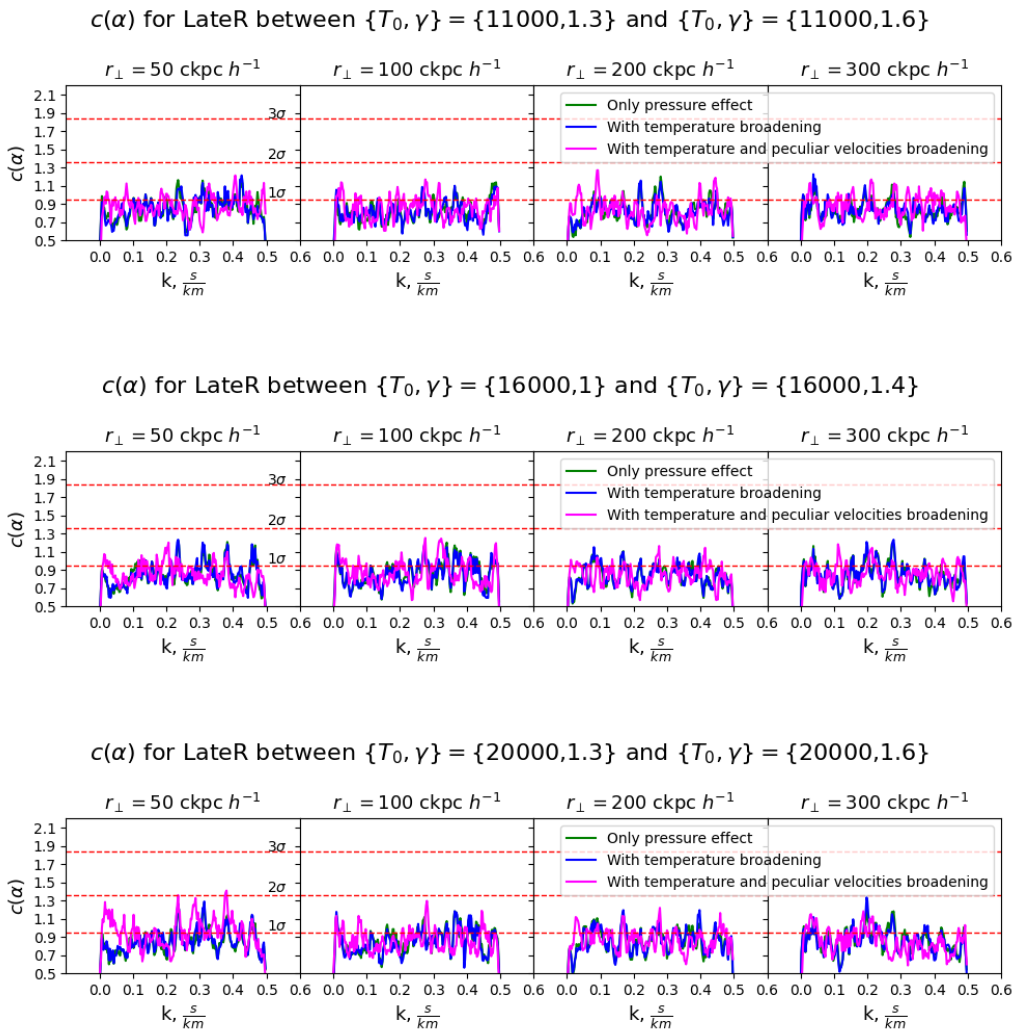


Figure 4.4: *Results of Kolmogorov-Smirnov test for varying slope γ of temperature-density relation for Late reionization, represented by $c(\alpha)$ for studied values of wavenumbers k. Each panel corresponds to different sets $\{T_0, \gamma\}$ of parameters characterizing thermal state of IGM and includes various effects (pressure smoothing, temperature broadening, peculiar velocities). Red dashed lines represent confidence levels in σ (1σ: 68%, 2σ: 95%, 3σ: 99.7%).*

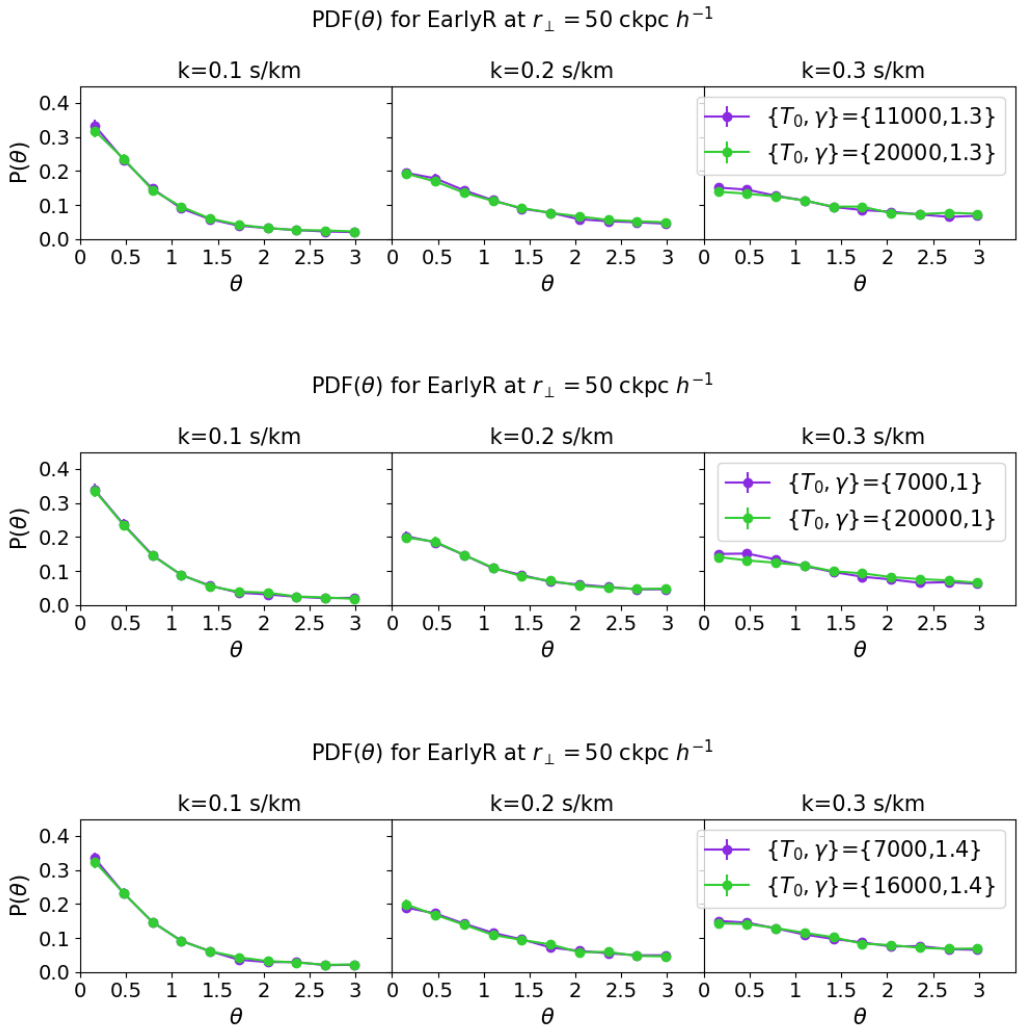


Figure 4.5: *Examples of phase difference probability density functions $P(\theta)$ for Early reionization at $r_\perp = 50\,ckpc\,h^{-1}$ for different sets $\{T_0, \gamma\}$ of parameters characterizing thermal state of IGM from the Figure 4.3 and accounting for effects of pressure smoothing, temperature broadening, and peculiar velocities. To smooth statistical fluctuations, PDFs were calculated in bins containing 10 points in k-space.*

Version of February 16, 2024 – Created March 5, 2024 - 09:37

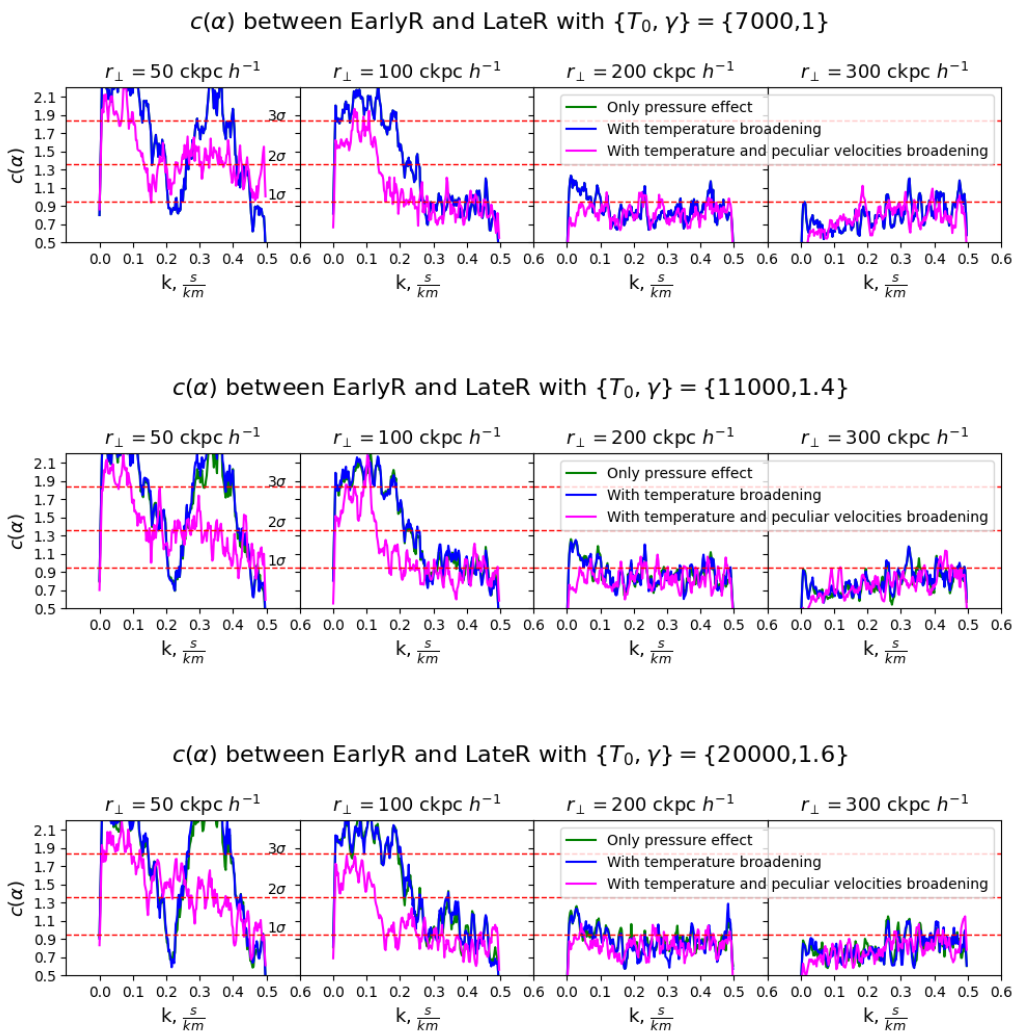


Figure 4.6: *Results of Kolmogorov-Smirnov test for different thermal histories (Early and Late reionization), represented by $c(\alpha)$ for studied values of wavenumbers k. Each panel corresponds to different sets $\{T_0, \gamma\}$ of parameters characterizing thermal state of IGM and includes various effects (pressure smoothing, temperature broadening, peculiar velocities). Red dashed lines represent confidence levels in σ (1σ: 68%, 2σ: 95%, 3σ: 99.7%).*

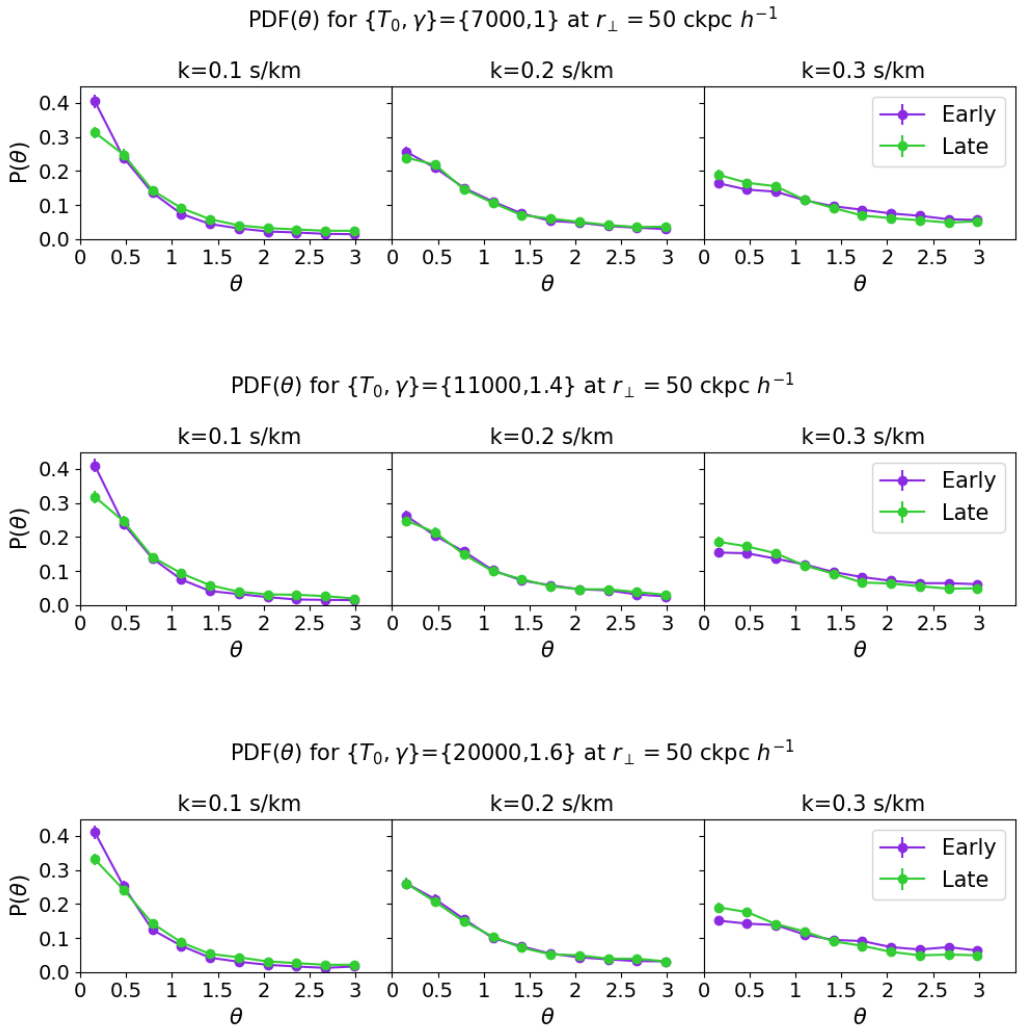


Figure 4.7: *Examples of phase difference probability density functions $P(\theta)$ for Early and Late reionizations at $r_\perp = 50$ ckpc h^{-1} for different sets $\{T_0, \gamma\}$ of parameters characterizing thermal state of IGM from the Figure 4.6 and without effects of temperature broadening and peculiar velocities. To smooth statistical fluctuations, PDFs were calculated in bins containing 10 points in k-space.*

Chapter 5

Summary

Most of the baryonic matter in the Universe is located in IGM, which is a suitable environment for probing small-scale structures. During its evolution, IGM experiences different phases, such as cooling, ionization, and heating. The key moment in the history of IGM is the Epoch of Reionization, which still possesses a lot of mysteries.

Two main approaches were developed to study IGM, namely the 21-cm line and the Lyman-α forest. In this thesis, we focus on the latter one, which is capable of probing a wide range of cosmic gas temperatures, densities, and ionization fractions at $z \sim 2-5$. The longitudinal FPS of Lyman-α forest exhibits a cut-off at small scales, that can be explained by Hydrogen pressure smoothing, thermal Doppler broadening, peculiar velocities along the LOS, and warm dark matter. The first one is the most important for studying EoR. However, it is a non-trivial task to separate it from the other effects.

In this thesis, we studied correlations between Lyman-α forests of close quasar pairs to distinguish different thermal histories of the Universe. The goal was to develop a method that is sensitive to 3D effects (pressure smoothing), but independent of 1D effects (thermal Doppler broadening and peculiar velocities). For this purpose, we studied distributions of phase difference θ for different thermal histories, parameters characterizing IGM $\{T_0, \gamma\}$, LOS separations $r_\perp$, wavenumbers k, and accounting for different effects. To analyze the difference between PDF(θ) we used the Kolmogorov-Smirnov test.

The method we developed is insensitive to the thermal state of IGM, characterized by values of mean IGM temperature T_0 and slope γ of the temperature-density relation. At the same time, we can distinguish different reionization histories. This result agrees with Rorai et al. (2013) [56],

where the close quasar pairs approach was introduced.

In addition, we encountered a notable feature of two peaks in the outcome of the Kolmogorov-Smirnov test between different thermal histories when the separations between LOS are $\lesssim \lambda_F$ and peculiar velocities are not accounted for. This phenomenon can be explained by the change in the influence of pressure smoothing at small and large scales. For a bigger pressure effect, structures parallel to the LOS are more likely to be penetrated by both LOS, explaining the large-scale effect. However, the higher pressure effect also corresponds to a faster transition to the uniform PDF(θ), causing a small-scale peak. Moreover, the second peak vanishes when accounting for peculiar velocities.

This thesis does not exhaust all the potential of the close quasar pairs. Our next step is to add more thermal histories to probe the sensitivity of the method. In addition, we are going to explore the feasibility of distinguishing scenarios with WDM.

Version of February 16, 2024 – Created March 5, 2024 - 09:37

Appendices

Appendix A

Phase difference calculation

Here we describe the algorithm used to compute the phase difference (2.19) between two parallel spectra.

1. Extract overdensity $\delta(v)$ and peculiar velocities $v_{pec}(v)$ for given spectra from the simulation spashot.

2. Calculate temperature inserting chosen parameters $\{T_0, \gamma\}$ in temperature-density relation (1.31).

3. Calculate density of neutral Hydrogen n_{HI} at the redshift z_{snap} of the snapshot:

$$n_{HI}(z_{snap}, v) = \frac{\alpha_R(T(v)) n_H^2(z_{snap}, v)}{\Gamma_\gamma(z_{snap})} \frac{1 - Y_p/2}{1 - Y_p}, \tag{A.1}$$

$$n_H(z_{snap}, v) = \frac{3H_0\Omega_b(1 - Y_p)}{8\pi G m_p}(1 + \delta(v))(1 + z_{snap})^3, \tag{A.2}$$

   where we used $\alpha_R(T) = 1.269 \cdot 10^{-13} \frac{(2T_{HI}/T)^{1.503}}{(1+(2T_{HI}/0.522T)^{0.470})^{1.923}}$ with the Hydrogen ionization threshold $T_{HI} = 157807$ K from Hui et al. (1997) [44] and $\Gamma_\gamma(z)$ is adapted from Onorbe et al. (2017) [71].

4. Calculate optical depth τ with and without chosen effects (temperature broadening T and peculiar velocities v). The code part A.1 shows how we computed τ accounting for Doppler broadening and peculiar velocities.

```
# Calculate optical depth for pTv case
# Given:
# TemperatureData - array of temperatures
```

```
# nHIdata - array of neutral Hydrogen number densities
# Vpdata - array of peculiar velocities
# N - number of pixels in each spectrum
# DeltaV - velocity interval (pixel size)
# KtoGeV, mH, n0(z_snapshot) - given quantities

import numpy as np

def TauWithTempandVp(nHIdata,TemperatureData,Vpdata):
    btab=c*(2*KtoGeV*TemperatureData/mH)**0.5
    nHI_over_n0 = nHIdata/n0(z_snapshot)
    tau = np.zeros((N, N))
    x = np.arange(-N//2, N//2)+1
    for k in range(N):
        erffunc = 1/2*(erf((x+0.5)*DeltaV/btab[k]-
        Vpdata[k]/btab[k]) - erf((x-0.5)*DeltaV/btab[k]-
        Vpdata[k]/btab[k]))
        dk = N//2 - k
        tautab = np.zeros(N)
        if dk == 0:
            tautab = nHI_over_n0*erffunc
        else:
            tautab[-dk:] = nHI_over_n0[:dk]*erffunc
            tautab[:-dk] = nHI_over_n0[dk:]*erffunc
        tau[:,k] = np.copy(tautab)
    return np.sum(tau, axis = 0)

```

Listing A.1: *Calculation of optical depth τ accounting for temperature Doppler broadening and peculiar velocities*

5. Calculate transmission flux $F(v)$ as $F(v) = \exp(-\tau)$.

6. Perform Fourier transform from Hubble velocity v space to wavenumbers k space $F(v) \rightarrow F(k)$ for each spectra. At this step, we used the discrete Fourier transform:

$$F[k] = \frac{1}{N} \sum_{v=0}^{N-1} e^{2\pi i \frac{kv}{N}} F[v], \tag{A.3}$$

   implemented in Scipy as an Inverse Fast Fourier Transform.

7. Calculate the difference between flux phases θ for each pair of parallel LOS separated by some $r_\perp$ and normalize it to fit in the domain $[0, \pi]$.

Version of February 16, 2024 – Created March 5, 2024 - 09:37

Bibliography

[1] Renyue Cen and Jeremiah P. Ostriker. Where are the baryons? *Astrophys. J.*, 514:1–6, 1999.

[2] J. Colin Hill, Simone Ferraro, Nick Battaglia, Jia Liu, and David N. Spergel. Kinematic Sunyaev-Zel'dovich Effect with Projected Fields: A Novel Probe of the Baryon Distribution with Planck, WMAP, and WISE Data. *Phys. Rev. Lett.*, 117(5):051301, 2016.

[3] Thais Lemos, Rodrigo S. Gonçalves, Joel C. Carvalho, and Jailson S. Alcaniz. Cosmological model-independent constraints on the baryon fraction in the IGM from fast radio bursts and supernovae data. *Eur. Phys. J. C*, 83(2):138, 2023.

[4] Zhengxiang Li, He Gao, Jun-Jie Wei, Yuan-Pei Yang, Bing Zhang, and Zong-Hong Zhu. Cosmology-independent estimate of the fraction of baryon mass in the IGM from fast radio burst observations. *Astrophys. J.*, 876(2):146, 2019.

[5] Masataka Fukugita and P. James E. Peebles. The Cosmic energy inventory. *Astrophys. J.*, 616:643–668, 2004.

[6] Renyue Cen and Jeremiah P. Ostriker. Where are the baryons? 2. feedback effects. *Astrophys. J.*, 650:560–572, 2006.

[7] Matthew McQuinn. The Evolution of the Intergalactic Medium. *Ann. Rev. Astron. Astrophys.*, 54:313–362, 2016.

[8] Antonella Garzilli, Andrii Magalich, Tom Theuns, Carlos S. Frenk, Christoph Weniger, Oleg Ruchayskiy, and Alexey Boyarsky. The Lyman-α forest as a diagnostic of the nature of the dark matter. *Mon. Not. Roy. Astron. Soc.*, 489(3):3456–3471, 2019.

[9] Jason Tumlinson, Mark L. Giroux, and J. Michael Shull. Probing the first stars with hydrogen and helium recombination emission. *Astrophys. J. Lett.*, 550:L1, 2001.

[10] Daniel Schaerer. The Massive star IMF. In *IAU Symposium 212: A Massive Star Odyssey from Main Sequence to Supernova*, 8 2002.

[11] Aparna Venkatesan and James W. Truran. The ionizing efficiency of the first stars. *Astrophys. J. Lett.*, 594:L1–L4, 2003.

[12] Nickolay Y. Gnedin and Lam Hui. Probing the universe with the Lyman alpha forest: 1. Hydrodynamics of the low density IGM. *Mon. Not. Roy. Astron. Soc.*, 296:44–55, 1998.

[13] Lam Hui and Zoltan Haiman. The thermal memory of reionization history. *The Astrophysical Journal*, 596(1):9, 2003.

[14] Girish Kulkarni, Joseph F Hennawi, Jose Onorbe, Alberto Rorai, and Volker Springel. Characterizing the pressure smoothing scale of the intergalactic medium. *The Astrophysical Journal*, 812(1):30, 2015.

[15] Avery A. Meiksin. The Physics of the Intergalactic Medium. *Rev. Mod. Phys.*, 81:1405–1469, 2009.

[16] Piero Madau, Avery Meiksin, and Martin J. Rees. 21-cm tomography of the intergalactic medium at high redshift. *Astrophys. J.*, 475:429, 1997.

[17] Jonathan R. Pritchard and Abraham Loeb. 21-cm cosmology. *Rept. Prog. Phys.*, 75:086901, 2012.

[18] Lars Hernquist, Neal Katz, David H. Weinberg, and Jordi Miralda-Escude. The Lyman alpha forest in the cold dark matter model. *Astrophys. J. Lett.*, 457:L51, 1996.

[19] Jonathan P. Gardner et al. The James Webb Space Telescope. *Space Sci. Rev.*, 123:485, 2006.

[20] David J. Bacon et al. Cosmology with Phase 1 of the Square Kilometre Array: Red Book 2018: Technical specifications and performance forecasts. *Publ. Astron. Soc. Austral.*, 37:e007, 2020.

[21] Chris L. Carilli and S. Rawlings. Science with the Square Kilometer Array: Motivation, key science projects, standards and assumptions. *New Astron. Rev.*, 48:979, 2004.

[22] Jonathan R. Pritchard, Abraham Loeb, and J. Stuart B. Wyithe. Constraining Reionization using 21-cm Observations in Combination with CMB and Lyman-Alpha Forest Data. *Mon. Not. Roy. Astron. Soc.*, 408:57, 2010.

[23] Pieter Daniel Meerburg, Cora Dvorkin, and David N. Spergel. Probing Patchy Reionization through τ-21 cm Correlation Statistics. *Astrophys. J.*, 779:124, 2013.

[24] Volker Bromm, Naoki Yoshida, Lars Hernquist, and Christopher F. McKee. The formation of the first stars and galaxies. *Nature*, 459:49–54, 2009.

[25] Brant E. Robertson, Richard S. Ellis, James S. Dunlop, Ross J. McLure, and Daniel P. Stark. Early star-forming galaxies and the reionization of the Universe. *Nature*, 468:49, 2010.

[26] Romeel Davé. Simulations of the intergalactic medium. In *Symposium-International Astronomical Union*, volume 216, pages 251–265. Cambridge University Press, 2005.

[27] Piero Madau, Martin J Rees, Marta Volonteri, Francesco Haardt, and S Peng Oh. Early reionization by miniquasars. *The Astrophysical Journal*, 604(2):484, 2004.

[28] Eric Wim Flesch. The million quasars (milliquas) catalogue, v8. *arXiv preprint arXiv:2308.01505*, 2023.

[29] Rajat M. Thomas and Saleem Zaroubi. Time-evolution of ionization and heating around first stars and miniquasars. *Mon. Not. Roy. Astron. Soc.*, 384:1080, 2008.

[30] Ruben Salvaterra, Francesco Haardt, and Andrea Ferrara. Cosmic backgrounds from miniquasars. *Mon. Not. Roy. Astron. Soc.*, 362:L50, 2005.

[31] Mark Dijkstra, Zoltan Haiman, and Abraham Loeb. A Limit from the x-ray background on the contribution of quasars to reionization. *Astrophys. J.*, 613:646–654, 2004.

[32] Alexander Kashlinsky, R. G. Arendt, John C. Mather, and S. H. Moseley. Tracing the first stars with fluctuations of the cosmic infrared background. *Nature*, 438:45–50, 2005.

[33] Saleem Zaroubi. The Epoch of Reionization. 6 2012.

[34] Andrea Ferrara and Stefania Pandolfi. Reionization of the Intergalactic Medium. *Proc. Int. Sch. Phys. Fermi*, 186:1–57, 2014.

[35] George D. Becker, James S. Bolton, Piero Madau, Max Pettini, Emma V. Ryan-Weber, and Bram P. Venemans. Evidence of patchy hydrogen reionization from an extreme Lyα trough below redshift six. *Mon. Not. Roy. Astron. Soc.*, 447:3402, 2015.

[36] James E. Gunn and Bruce A. Peterson. On the Density of Neutral Hydrogen in Intergalactic Space. *Astrophys. J.*, 142:1633, 1965.

[37] Robert H. Becker et al. Evidence for Reionization at Z $\sim$ 6: Detection of a Gunn-Peterson trough in a Z = 6.28 Quasar. *Astron. J.*, 122:2850, 2001.

[38] Richard L. White, Robert H. Becker, Xiao-Hui Fan, and Michael A. Strauss. Probing the ionization state of the universe at Z $\gtrsim$ 6. *Astron. J.*, 126:1, 2003.

[39] R. Adam et al. Planck intermediate results. XLVII. Planck constraints on reionization history. *Astron. Astrophys.*, 596:A108, 2016.

[40] Dhiraj Kumar Hazra and George F Smoot. Witnessing the reionization history using cosmic microwave background observation from planck. *Journal of Cosmology and Astroparticle Physics*, 2017(11):028, 2017.

[41] Michael Shull, Kevin France, Charles Danforth, Britton Smith, and Jason Tumlinson. Hubble/COS Observations of the Quasar HE 2347-4342: Probing the Epoch of He II Patchy Reionization at Redshifts z = 2.4-2.9. *Astrophys. J.*, 722:1312–1324, 2010.

[42] Phillip James Edwin Peebles. *The large-scale structure of the universe*, volume 98. Princeton university press, 2020.

[43] Girish Kulkarni, Joseph F. Hennawi, Jose Oñorbe, Alberto Rorai, and Volker Springel. Characterizing the Pressure Smoothing Scale of the Intergalactic Medium. *Astrophys. J.*, 812:30, 2015.

[44] Lam Hui and Nickolay Y Gnedin. Equation of state of the photoionized intergalactic medium. *Monthly Notices of the Royal Astronomical Society*, 292(1):27–42, 1997.

[45] JD Wagenveld, A Saxena, KJ Duncan, HJA Röttgering, and M Zhang. Revealing new high-redshift quasar populations through gaussian mixture model selection. *Astronomy & Astrophysics*, 660:A22, 2022.

[46] Adam D. Myers et al. The SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Quasar Target Selection. *Astrophys. J. Suppl.*, 221(2):27, 2015.

[47] Isabelle Pâris et al. The Sloan Digital Sky Survey Quasar Catalog: Fourteenth data release. *Astron. Astrophys.*, 613:A51, 2018.

[48] Alex Krolewski et al. Constraining primordial non-Gaussianity from DESI quasar targets and Planck CMB lensing. 5 2023.

[49] Andrii Magalich. *Probing the properties of dark matter particles with astrophysical observations*. PhD thesis, Leiden U., 2019.

[50] Antonella Garzilli, Tom Theuns, and Joop Schaye. The broadening of Lyman-α forest absorption lines. *Mon. Not. Roy. Astron. Soc.*, 450(2):1465–1476, 2015.

[51] Rupert A. C. Croft, David H. Weinberg, Neal Katz, and Lars Hernquist. Recovery of the power spectrum of mass fluctuations from observations of the Lyman alpha forest. *Astrophys. J.*, 495:44, 1998.

[52] Alexey Boyarsky, Julien Lesgourgues, Oleg Ruchayskiy, and Matteo Viel. Lyman-alpha constraints on warm and on warm-plus-cold dark matter models. *JCAP*, 05:012, 2009.

[53] Matteo Viel, George D. Becker, James S. Bolton, and Martin G. Haehnelt. Warm dark matter as a solution to the small scale crisis: New constraints from high redshift Lyman-α forest data. *Phys. Rev. D*, 88:043502, 2013.

[54] Julien Baur, Nathalie Palanque-Delabrouille, Christophe Yeche, Alexey Boyarsky, Oleg Ruchayskiy, Éric Armengaud, and Julien Lesgourgues. Constraints from Ly-α forests on non-thermal dark matter including resonantly-produced sterile neutrinos. *JCAP*, 12:013, 2017.

[55] J. S. Bolton, M. Viel, T. S. Kim, M. G. Haehnelt, and R. F. Carswell. Possible evidence for an inverted temperature-density relation in the intergalactic medium from the flux distribution of the Lyman-alpha forest. *Mon. Not. Roy. Astron. Soc.*, 386:1131–1144, 2008.

[56] Alberto Rorai, Joseph F. Hennawi, and Martin White. A New Method to Directly Measure the Jeans Scale of the Intergalactic Medium Using Close Quasar Pairs. *Astrophys. J.*, 775:81, 2013.

[57] Soumak Maitra, Raghunathan Srianand, Patrick Petitjean, Hadi Rahmani, Prakash Gaikwad, Tirthankar Roy Choudhury, and Christophe Pichon. Three- and two-point spatial correlations of intergalactic medium at $z \sim 2$ using projected quasar triplets. *Mon. Not. Roy. Astron. Soc.*, 490(3):3633–3653, 2019.

[58] Alberto Rorai, Joseph F. Hennawi, Jose Oñorbe, Martin White, J. Xavier Prochaska, Girish Kulkarni, Michael Walther, Zarija Lukić, and Khee-Gan Lee. Measurement of the small-scale structure of the intergalactic medium using close quasar pairs. *Science*, 356:418, 2017.

[59] Minghao Yue, Xiaohui Fan, Jinyi Yang, and Feige Wang. A candidate kiloparsec-scale quasar pair at z= 5.66. *The Astrophysical Journal Letters*, 921(2):L27, 2021.

[60] Mark Vogelsberger, Federico Marinacci, Paul Torrey, and Ewald Puchwein. Cosmological Simulations of Galaxy Formation. *Nature Rev. Phys.*, 2(1):42–66, 2020.

[61] Ya B Zel'Dovich. Gravitational instability: An approximate theory for large density perturbations. *Astronomy and astrophysics*, 5:84–89, 1970.

[62] Josh Barnes and Piet Hut. A hierarchical o (n log n) force-calculation algorithm. *nature*, 324(6096):446–449, 1986.

[63] Tomoaki Ishiyama, Keigo Nitadori, and Junichiro Makino. 4.45 pflops astrophysical n-body simulation on k computer–the gravitational trillion-body problem. In *SC'12: Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, pages 1–10. IEEE, 2012.

[64] Ewald Puchwein et al. The Sherwood–Relics simulations: overview and impact of patchy reionization and pressure smoothing on the intergalactic medium. *Mon. Not. Roy. Astron. Soc.*, 519(4):6162–6183, 2023.

[65] George D Becker, Frederick B Davies, Steven R Furlanetto, Matthew A Malkan, Elisa Boera, and Craig Douglass. Evidence for large-scale fluctuations in the metagalactic ionizing background near redshift six. *The Astrophysical Journal*, 863(1):92, 2018.

[66] Volker Springel. The Cosmological simulation code GADGET-2. *Mon. Not. Roy. Astron. Soc.*, 364:1105–1134, 2005.

[67] Roman Scoccimarro, Lam Hui, Marc Manera, and Kwan Chuen Chan. Large-scale Bias and Efficient Generation of Initial Conditions for Non-Local Primordial Non-Gaussianity. *Phys. Rev. D*, 85:083002, 2012.

[68] Antony Lewis, Anthony Challinor, and Anthony Lasenby. Efficient computation of CMB anisotropies in closed FRW models. *Astrophys. J.*, 538:473–476, 2000.

[69] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 594:A13, 2016.

[70] Francesco Haardt and Piero Madau. Radiative transfer in a clumpy Universe. 2. The Ultraviolet extragalactic background. *Astrophys. J.*, 461:20, 1996.

[71] Jose Oñorbe, Joseph F Hennawi, and Zarija Lukić. Self-Consistent Modeling of Reionization in Cosmological Hydrodynamical Simulations. *Astrophys. J.*, 837(2):106, 2017.

[72] R. Adam et al. Planck 2015 results. I. Overview of products and scientific results. *Astron. Astrophys.*, 594:A1, 2016.

[73] Tom Theuns, Anthony Leonard, George Efstathiou, F. R. Pearce, and P. A. Thomas. P**3M-SPH simulations of the lyman-alpha forest. *Mon. Not. Roy. Astron. Soc.*, 301:478–502, 1998.

[74] Gabriel Altay and Tom Theuns. Urchin: a reverse ray tracer for astrophysical applications. *Monthly Notices of the Royal Astronomical Society*, 434(1):748–764, 2013.

[75] Michael A Stephens. Use of the kolmogorov–smirnov, cramer–von mises and related statistics without extensive tables. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 32(1):115–122, 1970.

[76] M Donald MacLaren. The art of computer programming. volume 2: Seminumerical algorithms (donald e. knuth). *SIAM Review*, 12(2):306–308, 1970.